International Journal of Sciences: Basic and Applied Research (IJSBAR)

ISSN 2307-4531

http://gssrr.org/index.php?journal=JournalOfBasicAndApplied

Combined Estimators as Alternative to Ordinary Least Square Estimator

Kayode Ayinde[a]*

[a] *Department of Statistics, Ladoke Akintola University of Technology, P.M.B.4000, Ogbomoso,Oyo State, Nigeria*

[a]*Email:bayoayinde@yahoo.com*

Abstract

The Ordinary Least Square (OLS) estimator of the classical linear regression model is Best Linear Unbiased Estimator (BLUE) provided the assumptions of the model are not violated. In this paper, attempt is made to combine some Feasible Generalized Least Square (FGLS) estimators with the estimator based on Principal Component (PC) Analysis and compare their finite sampling properties and goodness-of-fit statistics with that of the OLS estimator through Monte Carlo Simulation study. Using both normally and uniformly distributed variables as regressors, results show that the estimators perform better and similar with increased sample sizes and that the results from normally distributed variables are much better on the basis of the criteria. The OLS estimator remains the most efficient and the combined estimators compete favorably with the estimator (OLS) especially when the sample size is large. The combined estimators are frequently more efficient than their separate counterpart estimator, asymptotically equivalent and best in terms of their goodness-of-statistics estimates. Thus, the combined estimators have the advantage of being used better for prediction. Numerical example supports findings.

Keywords: OLS Estimator; Combined Estimator; Sampling Properties; Goodness –of – fit Statistics.

1. Introduction

The Ordinary Least Square (OLS) estimator developed several years ago is a well known estimator for parameter estimation of the Classical Linear Regression Model. When the assumptions of the model are intact, the estimator is Best Linear Unbiased Estimator (BLUE) and computationally simple to use [1, 2, 3, 4.].Among these assumptions are the assumptions that regressors and error terms being assumed to be independent.

The violation of the assumption of independence of regressors leads to multicollinearity while that of error terms results into autocorrelation. Various estimation methods have separately been developed to tackle these problems. Estimators which include the Ridge Regression estimator developed by Hoerl [5] and Hoerl and Kennard [6], Estim-

----

* Corresponding author. Tel.: +234-803-585-0519;
E-mail address: kayinde@lautech.edu.ng

tor based on Principal Component Regression suggested by Massy [7], Marquardt [8] and method of Partial Least Squares developed by Hermon Wold in the 1960s [9,10, 11] are either developed or adopted to tackle the problem of multicollinearity. Furthermore, in order to compensate for the lost of efficiency which the OLS estimator often provides when used to analysis data with autocorrelated error terms, several Feasible Generalized Linear Estimators which include estimator provided by Cochrane and Orcutt [12], Paris and Winstern [13], Hildreth and Lu [14], Durbin [15], Theil [16], Beach and Mackinnon [17], and Thornton [18] have also been developed.

The problem of joint existence of multicollinearity and autocorrelation in a data set leads to the development of combined estimators whose performances are being examined and compared with the OLS estimator in this paper. The model is formulated to ensure complete compliance with the assumption of the classical linear regression model for easy of comparison.

2. Materials and Methods

Consider the linear regression model of the form:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 X_{3t} + U_t \tag{1}$$

Where $U_t = \rho U_{t-1} + \varepsilon_t$, $\varepsilon_t \sim N(0, \sigma^2)$, t = 1, 2, 3,...n

The autocorrelation parameter, ρ , in the model is set at zero to ensure compliance with the classical linear regression model in order to ensure comparison of the estimators . The technique adopted for the development of the combined estimator is very much similar to that of the Principal Component Estimator when used to solve multicollinerarity problem. Just like the Principal Component does its estimation using the OLS estimator by regressesing the extracted components (PCs) on the standardized dependent variable, the combined estimators use the FGLS estimator which include the Cochrane and Orcutt (CORC) estimator [12] and the Maximum Likelihood (ML) estimator [13]. Unlike the OLS estimator which results back into the OLS estimator when all the PCs are used, advantageously, since the FGLS estimators require an iterative methodology for its estimation, the proposed combined estimators can use all the possible PCs, the number of which is the same as the number of regressors, for its estimation and may not result back into the FGLS feasible estimators. The PCs are the extracted components of the X matrix of the regressors. They are orthogonal and hence, the validity of non-dependence of the regressors assumption is ensured. Consequently, the parameters of (1) are estimated by the following eleven (11) estimators: OLS, CORC, ML, Principal Components (PC1, PC12), CORC+PC1, CORC+PC12, CORC+PC123, ML+PC1, ML+PC12 and ML+PC123 estimators.

For the Monte-Carlo simulation study, two types of regessors namely, $X_i \sim N(0,1)$ and $X_i \sim U(0,1)$ were used. The parameters of equation (1) were specified and fixed as$\beta_0 = 4$, $\beta_1 = 2.5$, $\beta_2 = 1.8$ and $\beta_3 = 0.6$. Furthermore, the experiment was replicated in 1000 times (R =1000) under eight (8) levels of sample sizes (n =10, 20, 30, 50, 100, 200, 250 and 500). The estimators were evaluated and compared using the finite properties of estimators' criteria and the goodness – of –fit statistics. Mathematically, for any estimator $\hat{\beta}_i$ of β_i i = 0, 1,2,3 the finite sampling properties used which are Absolute Bias (AB) and the Mean Square Error (MSE) are defined as follows:

$$AB(\hat{\beta}_i) = \frac{1}{R} \sum_{j=1}^{R} \left| \hat{\beta}_{ij} - \beta_i \right| \tag{2}$$

$$MSE(\hat{\beta}_i) = \frac{1}{R}\sum_{j=1}^{R}\left(\hat{\beta}_{ij} - \beta\right)^2 \tag{3}$$

Also, the goodness-of-fit statistics used is the Adjusted Coefficient of Variation. The Adjusted Coefficient of Determination of the model was averaged over the numbers of replications. i.e.

$$\bar{R} = \frac{1}{R}\sum_{i=1}^{R} R^2_{Adj(i)} \tag{4}$$

In order to examine the pattern of the number of occurrence of the Adjusted coefficient of Determination, its frequency distribution were obtained as follows: $R^2_{Adj.} < 0.25, \quad 0.25 \le R^2_{Adj.} < 0.5, \; 0.5 \le R^2_{Adj.} < 0.75,$, $0.75 \le R^2_{Adj.} < 1.$ For all these estimators, a computer program was written using Time Series Processor [19] software to evaluate the Absolute Bias, Mean Square Error, and Average Adjusted coefficient of Determination and its frequency distribution. An estimator is best in terms of Absolute Bias and most efficient in terms of Mean Square Error if its value is the minimum. An estimator is best in terms of estimate of goodness-of-fits if its Adjusted Coefficient of Determination is closest to unity. All these were done by written a computer program using the Time Series Processor (TSP) software [19].

3. Results and Conclusion

Out of all the estimators, the results of the best six (6) are presented and discussed using the criteria outlined earlier. These estimators are OLS, CORC, ML, PC12, CORC+PC123 and ML+PC123.

3.1 The Absolute Bias Criterion

The results on the basis of absolute bias of the estimators using normally distributed regressors are summarized in Table 1 while that of uniformly distributed regressors are graphically summarized in Figure 1.

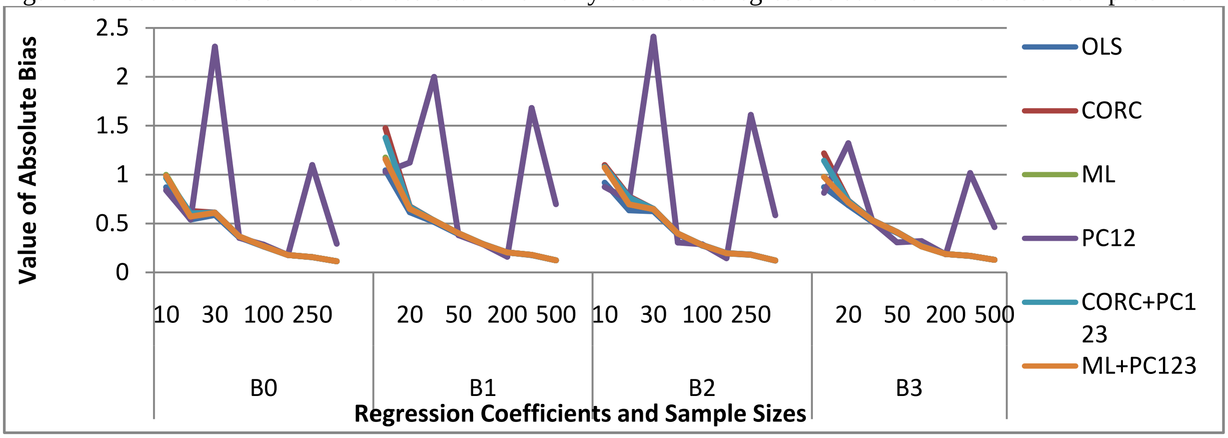
Table 1**:** The Absolute Bias of the Estimators with normally distributed regressors at different levels of sample size

n	Estimator	B0	B1	B2	B3	n	B0	B1	B2	B3
10	OLS	.270780	.316720	.272230	.233600	100	.084187	.083918	.077697	.074691
	CORC	3.52577	.416600	.337370	.349140		.084078	.084837	.079278	.075532
	ML	.293840	.352090	.332430	.265460		.084138	.084757	.079018	.074612
	PC12	.270800	.337400	1.38307	1.12897		.109290	.087985	1.24839	1.20161
	CORC+PC123	.285130	.393250	.335650	.319440		.084162	.084836	.079275	.075518
	ML+PC123	.274860	.348280	.329270	.264470		.084191	.084757	.079018	.074612
20	OLS	.187740	.186070	.187990	.215530	200	.057799	.058331	.055713	.053371
	CORC	.214380	.200740	.243820	.226050		.057825	.058466	.056183	.053401
	ML	.191740	.195360	.203250	.223130		.057837	.058446	.056086	.053420
	PC12	.310110	1.08392	.642150	1.40518		.059498	.446320	.405590	.072946
	CORC+PC123	.194080	.199650	.238870	.225350		.057830	.058465	.056182	.053401
	MLPC123	.189760	.195160	.202790	.222900		.057829	.058446	.056086	.053420
30	OLS	.153490	.149710	.202110	.144300	250	.051443	.050892	.049919	.048572
	CORC	.158030	.154660	.209810	.148270		.051733	.051460	.050513	.048941
	ML	.153830	.154650	.208490	.147950		.051439	.051435	.050349	.048868
	PC12	.255730	1.93013	2.55778	.144400		.137770	1.95705	1.74692	.819080
	CORC+PC123	.153800	.154640	.209690	.148230		.051467	.051459	.050512	.048940
	MLPC123	.153670	.154620	.208450	.147940		.051444	.051435	.050349	.048868

	OLS	.114310	.120310	.114120	.111830		.035125	.035351	.035613	.037140
	CORC	.116190	.123270	.115940	.113990		.035246	.035612	.035755	.037412
5	ML	.114580	.123370	.115940	.113950	5	.035147	.035583	.035722	.037299
0	PC12	.114180	.114760	.086546	.079869	0	.042738	.637030	.576800	.484020
	CORC+PC123	.114460	.123250	.115940	.113980	0	.035148	.035612	.035755	.037412
	MLPC123	.114460	.123360	.115940	.113950		.035138	.035583	.035722	.037299

From Table 1 and Figure 1, it can be seen that apart from PC12 estimator as the sample size increases the estimators' performances are better and much better with normally distributed regressors. The OLS estimator is often best even though all other estimators except the PC12 estimator perform equally well especially when the sample size is large. At small sample sizes, the combined estimators are less biased than their separate counterparts. Summarily, apart from PC12, all the estimators are asymptotically equivalent with the absolute bias criterion.

Figure 1: Absolute Bias of the Estimators with normally distributed regressors at different levels of sample size



3.2 The Mean Square Error Criterion

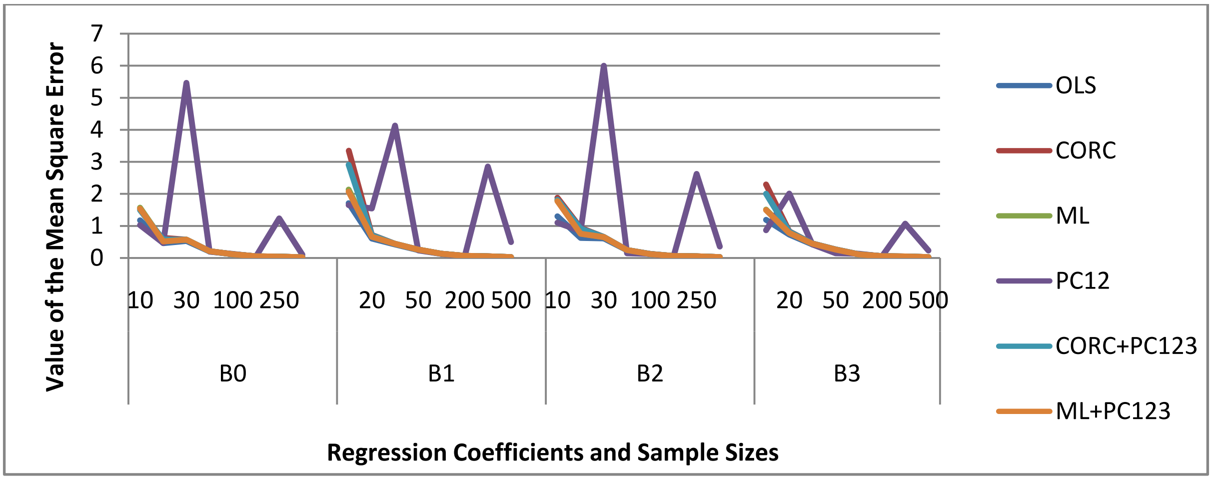
The results of Mean Square Error of the estimators using normally distributed regressors are summarized in Table 2 while that of uniformly distributed regressors are graphically summarized in Figure 2. From Table 2 and Figure 2, it can be observed that apart from the PC12 estimator as the sample size increases the performances of the estimators except that of PC12 are better and much better with normally distributed regressors. The OLS estimator remains the most efficient even though all other estimators except the PC12 estimator perform equally well especially when the sample size is large. The combined estimators are frequently more efficient than their separate counterparts. Summarily, apart from PC12, all the estimators are asymptotically equivalent under the mean square error criterion.

Table 2: The Mean Square Error of the Estimators with normally distributed regressors at different levels of sample size

n	**Estimator**	**B0**	**B1**	**B2**	**B3**	**n**	**B0**	**B1**	**B2**	**B3**
	OLS	.114220	.156860	.115520	.086792		.010880	.011168	.009413	.008838
	CORC	943.580	.270660	.175050	.185280		.010870	.011416	.009768	.008975
1	ML	.137720	.193970	.170900	.113240	1	.010874	.011400	.009739	.008841
0	PC12	.114330	.179330	1.94214	1.30159	0	.018166	.012363	1.56238	1.44776
	CORC+PC123	.127700	.241190	.174440	.154230	0	.010880	.011416	.009768	.008973
	ML+PC123	.119190	.189700	.167660	.112350		.010889	.011400	.009739	.008841
	OLS	.054546	.054806	.055307	.072629		.005204	.005396	.004806	.004366
	CORC	.071430	.063352	.092801	.079128		.005218	.005404	.004863	.004401
2	ML	.057257	.060113	.065291	.077129	2	.005210	.005398	.004859	.004399
0	PC12	.134490	1.20388	.456510	2.00030	0	.005525	.201420	.166940	.007928
	CORC+PC123	.058134	.062627	.088790	.078634	0	.005209	.005404	.004862	.004401

	ML+PC123	.055963	.059997	.064962	.076989		.005209	.005398	.004859	.004399
	OLS	.037200	.035282	.064677	.033006		.004238	.004073	.003874	.003688
	CORC	.039482	.037640	.069025	.034951		.004295	.004137	.003933	.003765
3	ML	.037244	.037568	.068432	.034685	2	.004240	.004132	.003908	.003748
0	PC12	.091811	3.73669	6.56203	.033052	5	.022891	3.83194	3.05396	.674230
	CORC+PC123	.037303	.037629	.068962	.034932	0	.004247	.004137	.003933	.003765
	ML+PC123	.037202	.037553	.068411	.034679		.004242	.004132	.003908	.003748
	OLS	.020846	.023149	.020496	.019290		.001994	.002009	.002039	.002124
	CORC	.021493	.024316	.021028	.020214		.002006	.002040	.002060	.002150
5	ML	.020904	.024256	.021035	.020180	5	.001996	.002037	.002054	.002138
0	PC12	.020821	.021345	.011720	.009935	0	.002859	.406790	.334070	.235810
	CORC+PC123	.020891	.024309	.021027	.020210	0	.001996	.002040	.002060	.002150
	ML+PC123	.020894	.024253	.021033	.020178		.001995	.002037	.002054	.002138

Figure 2: Mean Square Error of the Estimators with normally distributed regressors at different levels of sample size



3.3 Estimate of Goodness–of-fit Statistics (Average Adjusted Coefficient of Determination)

The average coefficient of determination of the estimators and the frequency of their occurrences as earlier mentioned are shown in Table 3 for the normally distributed regressors while that of uniformly regressors are graphically represented in Figure 3. From these representations, the estimates of the goodness-of-fit statistics of all the estimators are much better with normal regressors than the uniform. Apart from PC12 estimator, all other estimators with the explanatory variables account for more than 85% variation of the response variable in the former while ML+PC123 and CORC+PC123 are best in the latter with around 50% of variation of the response variable being accounted for when the sample size is small.

This is further buttressed with the fact that almost all (1000 replications) the estimated adjusted coefficient of determinations is greater than 0.75 using normally distributed regressors. Figure 4 further reveals that the adjusted coefficient of determination of the estimators (except PC12) with uniformly distributed regessors are most frequent between 0.25 and 0.5. Summarily, at each level of sample size, the estimates of the combined estimator are often the best and even that of their separate counterpart, except the PC12, are still better than that of the OLS estimator. Moreover, their estimates are asymptotically the same.

Table 3:Average Adjusted Coefficient of Determination of the Estimators and the Frequency Distribution of their Occurrences

n	Estimator	AVERAGE	FREQUENCY OF OCCURENCES				n	AVERAGE	FREQUENCY OF OCCURENCES			
			$R^2_{Adj.} < 0.25$	$0.25 \le R^2_{Adj.} < 0.5$	$0.5 \le R^2_{Adj.} < 0.75$	$0.75 \le R^2_{Adj.} < 1$			$R^2_{Adj.} < 0.25$	$0.25 \le R^2_{Adj.} < 0.5$	$0.5 \le R^2_{Adj.} < 0.75$	$0.75 \le R^2_{Adj.} < 1$
10	OLS	.853340	0	2	114	884	100	.89800	0	0	.00	1000
	CORC	.875870	0	5	88	907		.89891	0	0	.00	1000
	ML	.864440	0	3	98	899		.89804	0	0	.00	1000
	PC12	.517250	21	418	525	36		.60747	0	1	999	0
	CORC+PC123	.899280	0	2	46	952		.89997	0	0	.00	1000
	ML+PC123	.886120	0	0	55	945		.89910	0	0	.00	1000
20	OLS	.914300	0	0	0	1000	200	.91714	0	0	0	1000
	CORC	.914310	0	0	1	999		.91751	0	0	0	1000
	ML	.914790	0	0	0	1000		.91713	0	0	0	1000
	PC12	.728280	0	1	625	374		.88902	0	0	0	1000
	CORC+PC123	.920000	0	0	0	1000		.91794	0	0	0	1000
	ML+PC123	.920080	0	0	0	1000		.91756	0	0	0	1000
30	OLS	.868930	0	0	5	995	250	.90918	0	0	0	1000
	CORC	.873170	0	0	4	996		.90917	0	0	0	1000
	ML	.869280	0	0	6	994		.90917	0	0	0	1000
	PC12	.161020	932	68	0	.00		.22194	961	39	0	0
	CORC+PC123	.878240	0	0	3	997		.90954	0	0	0	1000
	ML+PC123	.874300	0	0	4	996		.90954	0	0	0	1000.
50	OLS	.908700	0	0	0	1000	500	.91403	0	0	0	1000
	CORC	.910390	0	0	0	1000		.91415	0	0	0	1000
	ML	.908790	0	0	0	1000		.91404	0	0	0	1000
	PC12	.910060	0	0	0	1000		.83528	0	0	0	1000
	CORC+PC123	.912380	0	0	0	1000		.91433	0	0	0	1000
	ML+PC123	.910770	0	0	0	1000		.91421	0	0	0	1000

Figure 3:Average Adjusted Coefficient of Determination of the Estimators with uniformly distributed regressors

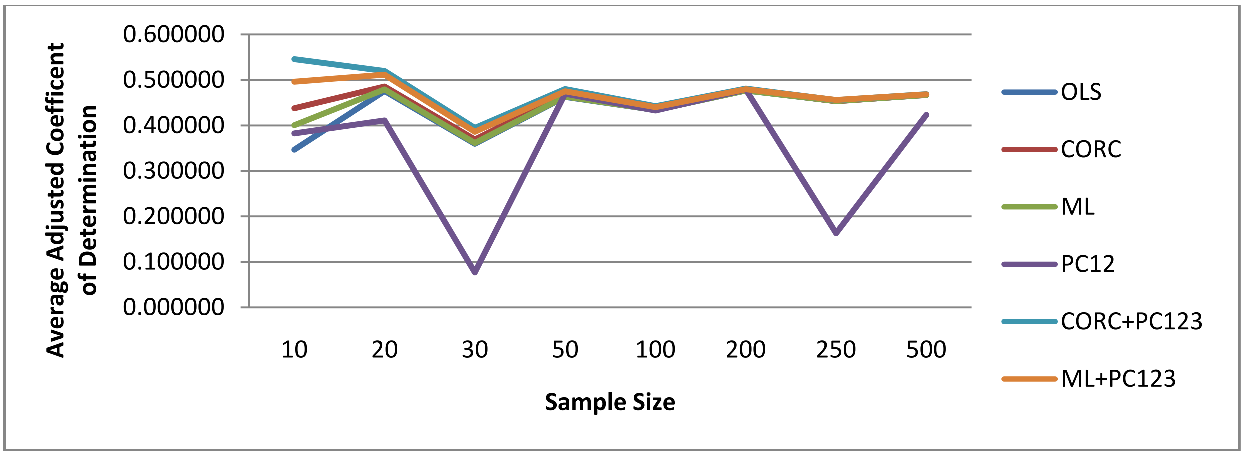
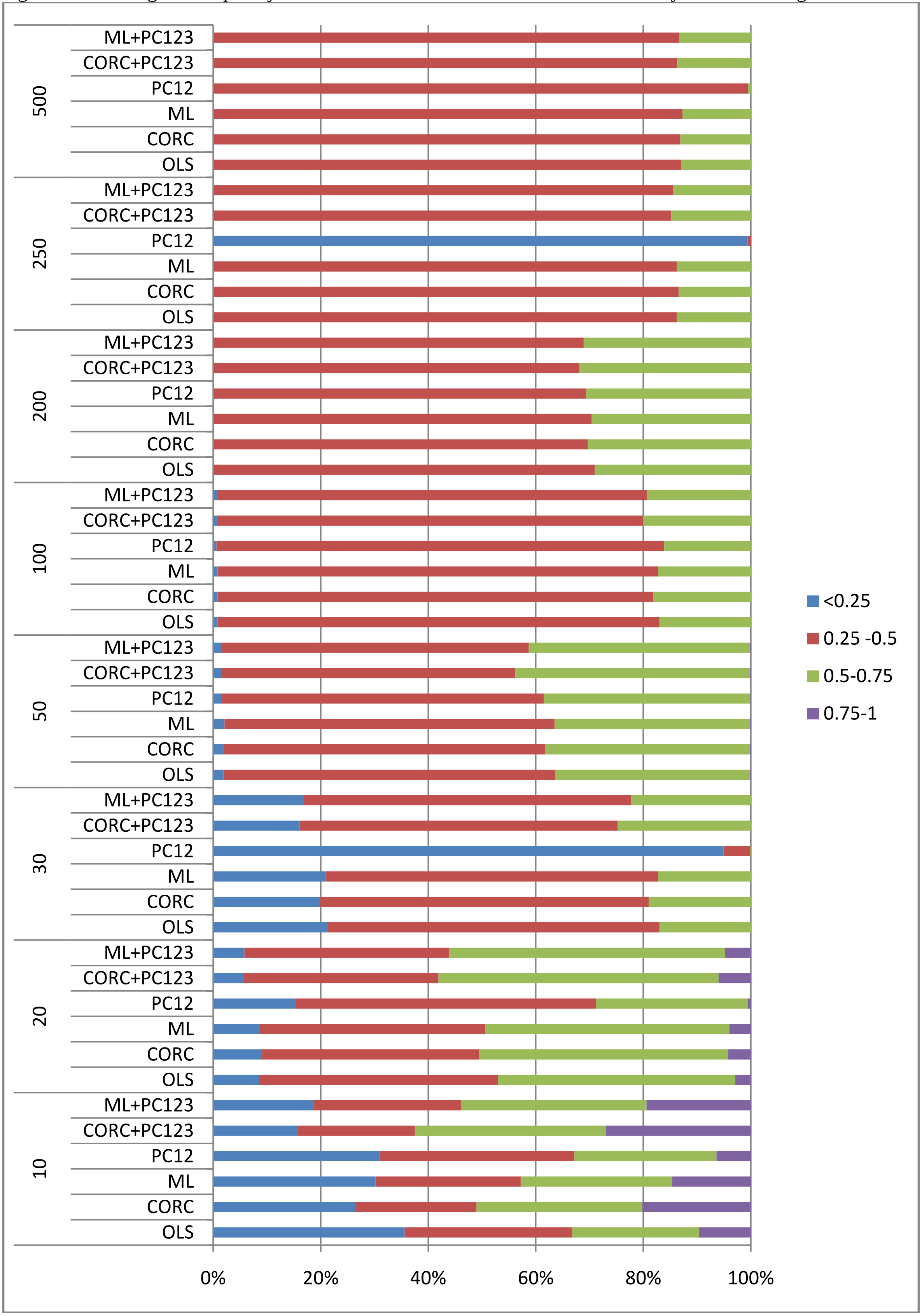


Figure 4:Percentage of frequency of occurrences of the Estimators with uniformly distributed regressors.



3.4 A Numerical Example

Macroeconomic data set comprising of fifteen (15) observations used by Greene [20] to illustrate the computations in a multiple regression using the OLS estimator is also used below to illustrate the findings on the combined estimators. The results are provided in Table 4.

Table 4:Summary of results on the Macroeconomic data provided by Greene [20]

Estimators	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$R^2_{Adj.}$
OLS	**-0.5099 (0.055)**	**-0.0166 (0.0020)**	**0.6718 (0.0552)**	**-0.0023 (0.0012)**	**-0.0003 (0.0012)**	**0.9616**
PC1	0.0921	0.0017	0.0462	0.0025	0.0030	0.5701
PC12	0.0805	0.0027	0.0736	0.0023	-0.0016	0.5909
PC123	0.0567	0.0037	0.1029	-0.0027	0.0007	0.6169
CORC	**-0.5173 (0.0394)**	**-0.0174 (0.0016)**	**0.6823 (0.0395)**	**-0.0018 (0.0009)**	**-0.0006 (0.0009)**	**0.9666**
CORC+PC1	0.1081	0.0014	0.0396	0.0021	0.0026	0.5679
CORC+PC12	0.0892	0.0023	0.0644	0.0022	-0.00063	0.5645
CORC+PC123	0.0602	0.0034	0.0957	-0.0020	0.0011	0.5443
CORC+PC1234	**-0.5167**	**-0.0173**	**0.6807**	**-0.0018**	**-0.0005**	**0.9702**
ML	**-0.5004 (0.0428)**	**-0.0162 (0.0016)**	**0.6639 (0.0423)**	**-0.0023 (0.0008)**	**-0.0005 (0.0010)**	**0.9581**
ML+PC1	0.0975	0.0016	0.0439	0.0024	0.0029	0.6128
ML+PC12	0.0830	0.0025	0.0679	0.0024	-0.0007	0.6176
ML+PC123	0.0612	0.0034	0.0953	-0.0020	0.0011	0.6057
ML+PC1234	**-0.5010**	**-0.0163**	**0.6643**	**-0.0023**	**-0.0005**	**0.9623**

NOTE: The standard error of the regression coefficient estimate is in the parenthesis.

Table 4 clearly shows the CORC, ML, CORC+PC1234 and ML+PC1234 compete very well with the OLS estimator. In fact, none of the separate estimators of the combined estimators produces a better result than the combined estimators with the highest number of principal components in terms of their goodness –of – statistics. It can also be seen that combined estimators with number of component less than the number of predictors do not perform well. This actually buttresses the fact emphasized easier that led to the selection of estimators presented in the simulation study.

4. Conclusion

This study has combined two Feasible Generalized Estimators with the Estimator based on Principal component Analysis and compared the performances of estimators with that of the OLS estimator. The combined estimators perform equally well and compete favourably with OLS estimator in terms of efficiency especially when the sample size is large. Moreover, the goodness-of-fit statistics of the combined estimator are often better than that of the OLS estimator. Thus, the estimators have the advantage of being used when regression analysis is focused on prediction.

References

[1] T. B.Fomby, R. C.Hill and S. R. Johnson.*Advance Econometric Methods*: Springer-Verlag, New York, Berlin, Heidelberg, London, Paris, Tokyo, 1984.

[2] S.Chartterjee, A.S.Hadi and B. Price. *Regression by Example*: 3rd Edition, A Wiley- Interscience Publication, John Wiley and Sons, 2000.

[3] S. Chartterjee and A.S. Hadi. *Regression by Example*: 4th Edition, A Wiley- Interscience Publication, John Wiley and Sons, 2006.

[4] G. S. Maddala.*Introduction to Econometrics*: 3rd Edition, John Willey and Sons Limited, England, 2002.

[5] A. E. Hoerl. Application of ridge analysis to regression problems:*Chemical Engineering Progress,* 58, 54 – 59, 1962.

[6] A. E. Hoerl and R. W.Kennard.Ridge regression biased estimation for non-orthogonal problems. *Technometrics,* 8, 27 – 51,1970.

[7] W. F. Massy.Principal Component Regression in exploratory statistical research.*Journal of the American Statistical Association,* 60, 234 – 246, 1965.

[8] D. W. Marquardt. Generalized inverse, Ridge Regression, Biased Linear Estimation and Non – linear Estimation. *Technometrics*, 12, 591 – 612, 1970.

[9] I. S. Helland. On the structure of partial least squares regression. *Communication is Statistics, Simulations and Computations,* 17, 581 – 607,1988.

[10] I. S. Helland. Partial least squares regression and statistical methods. *Scandinavian Journal of Statistics,* 17, 97 – 114, 1990.

[11] A.Phatak, and S. D. Jony. The geometry of partial least squares. *Journal of Chemometrics*, 11, 311 – 338, 1997.

[12] Cochrane, D. and Orcutt, G.H.: Application of Least Square to relationship containing autocorrelated error terms. *Journal of American Statistical Association,* 44:32 – 61,1949.

[13] S. J. Paris and C. B. Winstein. Trend estimators and serial correlation. *Unpublished Cowles Commision, Discussion Paper,* Chicago, 1954.

[14] C.Hildreth, and J.Y. Lu, Demand relationships with autocorrelated disturbances. Michigan State University. *Agricultural Experiment Statistical Bulletin,* 276 East Lansing, Michigan.1960.

[15] J. Durbin. Estimation of Parameters in Time series Regression Models. *Journal of Royal Statistical Society* B, 22:139 -153, 1960.

[16] H. Theil. *Principle of Econometrics*. New York, John Willey and Sons,1971.

[17] C. M. Beach,and J.S. Mackinnon,A Maximum Likelihood Procedure regression with autocorrelated errors. *Econometrica*, 46: 51 – 57, 1978.

[18] Thornton,D. L. The appropriate autocorrelation transformation when autocorrelation process has a finite past. *Federal Reserve Bank,*St. Louis, 82 – 102, 1982.

[19] *TSP Users'Guide and Reference Manual*. Time Series Processor, New York,2005.

[20] W. H.Greene: *Econometric Analysis*. 5th Edition, Prentice Hall Saddle River, New Jersey 07458, 2003.