ESTIMATORS FOR HANDLING MULTICOLLINEARITY PROBLEMS IN LINEAR REGRESSION MODEL WITH NORMALLY AND UNIFORMLY DISTRIBUTED REGRESSORS

BELLO, A.H.
Department of Statistics, School of Physical Sciences, Federal University of Technology, Akure, Nigeria

ABSTRACT

The assumptions of the classical linear regression model are hardly satisfied in real life situation. Violation of the assumption of independent explanatory variables and error terms in linear regression model leads to the problems of multicollinearity and autocorrelation respectively. Estimators to handle each problem have been separately developed by different authors. Moreover, in practice, these two problems do co-exist but estimators to handle them jointly are rare. Consequently, this research proposed and validate two estimators, Feasible Ordinary Ridge Estimators (FORE) and Feasible Generalized Ridge Estimators (FGRE), to handle the problems of multicollinearity and autocorrelation separately. The existing and proposed estimators were categorized into five (5) groups namely: One–Stage Estimators (OSE), Two–Stage Estimators (TSE), Feasible Generalized Least Square Estimators (FGLSE), Two-Process Estimators (TPE) and Modified Ridge Estimators (MRE). Monte Carlo experiments were conducted one thousand (1000) times on a linear regression model exhibiting different degrees of multicollinearity ($\lambda =$ 0.4, 0.6, 0.8, 0.95 and 0.99) with both normally and uniformly distributed regressors and autocorrelation ($\rho = 0.4, 0.8, 0.95$ and 0.99) at six sample sizes (n =10, 20, 30, 50, 100 and 250). In this study our autocorrelation is set to Zero (ρ = 0). Finite sampling properties of estimators namely; Bias (BAS), Mean Absolute Error (MAE), Variance (VAR) and most importantly Mean Square Error (MSE) of the estimators were evaluated, examined and compared at each specified level of multicollinearity, autocorrelation and sample sizes. These were done by ranking the estimators on the basis of their performances according to the criteria so as to determine the best estimator. Results of the investigation when multicollinearity alone was in the model revealed that the best estimator is in the category of One-Stage Estimator (OSE). With normally distributed regressor, the best estimator is generally the existing estimator OREKBAY except under the bias criterion. At this instance, the estimators FGLSE, ML and CORC are the best. Also, with uniformly distributed regressor, it was observed that the best estimator under all criteria is the existing estimator OREKBAY except under the bias criterion. At this instance the OLSE and FGLSE - ML are the best.

Keywords: Multicollinearity, Autocorrelation, Estimator, Regressors and TSP- Time Series Processor
Correspondence: habello@futa.edu.ng, 234-07033420672

INTRODUCTION

Multicollinearity is one of the important problems in multiple regression analysis. It is usually regarded as a problem arising as a result of the violation of the assumption that explanatory variables are linearly independent. However, just satisfaction of this assumption does not preclude the possibility of an approximate linear dependence among the explanatory variables, and hence the problem of multicollinearity. Though no precise definition of multicollinearity has been firmly established in the literature, multicollinearity is generally agreed to be present if there is an approximate degree of linear relationship among some of the predictor variables in the data. Bowerman and Connell [1] stated that the term multicollinearity refers to a situation in which there is an exact (or nearly exact) linear relation among two or more of the explanatory variables. Exact relations may arise by mistake or lack of understanding. Multicollinearity can also be defined in the concept of orthogonality. When the predictors are orthogonal or uncorrelated, all eigenvalues of the design matrix are equal to one and the design matrix is full rank. If at least one eigenvalue is different from one, especially when equal to zero or near zero, then non-orthogonality exists, meaning that multicollinearity is present. Multicollinearity can lead to increasing complexity in the research results, thereby posing difficulty for researchers to provide interpretation Charterjee *et al.* [2]. In theory, there are two extremes; Perfect Multicollinearity and No Multicollinearity. In practice, data typically are somewhere between those extremes. Thus, multicollinearity is a matter of degree. The real issue is to determine the point at which the degree of multicollinearity becomes “harmful”. The econometric literature typically takes the theoretical position that predictor variables are not collinear in the population. Hence, any observed multicollinearity in empirical data is considered as a sample based “problem” rather than as representative of the underlying population relationship Kmenta [3]. In many marketing research situations, however, it is unrealistic to assume that predictor variables are always strictly orthogonal at the population level especially when one is working with behavioral constructs. Regardless of whether multicollinearity in data is assumed to be a sampling problem or true reflection of population relationships, it must be looked into when data are analyzed using regression analysis because it has several potential undesirable consequences on the parameter estimates. When multicollinearity is a problem, parameter estimates have wrong signs when compared with theoretical knowledge and variables have insignificant coefficients. The regression coefficients, though determinate when multicollinearity is imperfect, possess large standard errors which imply that the coefficients cannot be estimated with great precision. Hawking and Gujarati [4, 5] Various other estimation methods have been developed to tackle this problem. These estimators include Stein estimator introduce by James [6], Ridge Regression estimator developed by Hoerl and Kennard [7]. Estimator based on Principal Component Regression suggested by Massy, Manfield and Paris [8 – 10], and method of Partial Least Squares developed by Wold [11]. Ayinde *et al* [12] examined the performances of estimators based on principal component regression, ML and Cochrane Orcutt.

The term autocorrelation may be defined as "correlation between members of series of observations ordered in time [as in time series data] or space [as in cross-sectional data].'' In the regression context, the classical linear regression model assumes that such autocorrelation does not exist in the disturbances u_i . Symbolically, $E(u_i u_j) = 0 \quad i \neq j$. The classical model assumes that the disturbance term relating to any observation is not influenced by the disturbance term relating to any other observation. For example, if we are dealing with quarterly time series data involving the regression of output on labor and capital inputs and if there is a labor strike affecting output in one quarter, there is no reason to believe that this disruption will be carried over to the next quarter. That is, if output is lower this quarter, there is no reason to expect it to be lower next quarter. Similarly, if we are dealing with cross-sectional data involving the regression of family consumption expenditure on family income, the effect of an increase of one family's income on its consumption expenditure is not expected to affect the consumption expenditure of another family. However, if there is such a dependence, we have autocorrelation Formby *et al.* [13].

The inefficiency of Ordinary Least Square estimator to estimate the parameters of linear regression model in the presence of autocorrelation led to the development of Generalized Least Squares (GLS) estimator. It requires the true autocorrelation value to be known which is not often so. Using the estimated autocorrelation value leads to the development of Feasible Generalized Least Squares (FGLS) estimators. Thus, studying the finite sample properties of these estimators becomes very imperative. This seems to be very difficult analytically [14]. However, Monte Carlo approach is often utilized to accomplish this task. Cochrane and Orcutt [15], an economist, observed that the presence of autocorrelated error terms requires some modifications for the OLS estimator to be used. Their suggestion involved an autoregressive transformation of the series involved and that the quasi first differences of such series should be used. Kadiyala [16] observed that the transformation suggested by [15] can lead to a less efficient estimator. He therefore suggested that the addition of one weighted observation to CORC procedure may give a better estimator practically without any extra cost. Rao and Griliches [17] did one of earliest Monte Carlo investigations on this study. The estimators they examined include OLS, CORC, Two-Step estimators based on Durbin $\hat{\rho}$ and Prais–Winsten. Their major conclusions were; the OLS estimator is less efficient than all other methods considered for moderate and high values of ($|\rho| > 0.3$) and there is a definite gain from using feasible generalized least squares when $|\rho| \geq 0.3$ and little loss from using such methods otherwise.

The work of [17] was revisited by Spitzer [18]; but in addition to the estimators considered by [17], he examined the performance of ML estimator. Situations where the two problems exist together in a data set are not uncommon. Therefore, this paper attempts to investigate the performances of the proposed estimators when multicollinearity alone is in the model through Monte Carlo studies.

METHODOLOGY

Estimators for handling multicollinearity generalized ridge estimator

The generalized ridge estimator of β is given as $\widehat{\beta_k} = (X'X + KI)^{-1}X'Y$ where K is a diagonal matrix with non negative diagonal elements $(K_1, K_2, \dots, K_p)$. The $K_i (i = 1,2, \dots, p)$ need not be equal. The optimum value of K had been obtained by Hoerl *et al.* [19] as:

$$K_i = \frac{\sigma^2}{\alpha_i^2} \quad i = 1,2,3, \dots, p. \qquad (1)$$

Since σ^2 and α_i^2 are generally unknown, the K_i needs to be estimated. [19] suggested the replacement of σ^2 and α_i^2 by their corresponding unbiased estimators $\hat{\sigma}^2$ and $\hat{\alpha}_i^2$. Therefore, $\hat{K}_i = \frac{\hat{\sigma}^2}{\hat{\alpha}_i^2}$ where $\hat{\sigma}^2 = \frac{\sum_{i=1}^{n} e_i^2}{n-p}$. (The estimated error variance from OLS estimation) and α_i^2 is the regression coefficient from OLS estimation.

The ordinary ridge estimator

The Ordinary ridge regression (ORR) estimator requires a fixed value of the ridge Parameter, K. Several Ks have also been proposed by authors including [7] and that of Ayinde *et al.* [20] respectively given as:

$$\hat{K}_{HK} = \frac{\hat{\sigma}^2}{\mathrm{Max}(\hat{\alpha})^2} \quad (2)$$

$$\hat{K}_{LA} = \frac{\hat{\sigma}^2}{[\mathrm{Max}(\hat{\alpha})]^2} \quad (3)$$

Sclove [21] suggested an empirical K-Bayesian Ridge Parameter given as:

$$\hat{K}_{BAY} = \frac{(SSR/n-p)}{(\sum y2 \ - SSR/\ \mathrm{trace}(X'X))} \quad (4)$$

Where SSR = Sum of Square of Regression

Estimators for handling autocorrelation

Consider the Linear Regression Model with Autoregressive of order 1, AR (1) given as:

$$y_t = \beta + \beta_1 x_{1t} + \beta_2 x_{2t} + \ldots + \beta_k x_{k-t} + u_t \quad (5)$$

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

Therefore, the variance – covariance matrix becomes:

$$E(UU') = \sigma^2\Omega = \sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{n-2} & \rho^{n-1} \\ \rho & 1 & \rho & \cdots & \rho^{n-3} & \rho^{n-2} \\ \rho^2 & \rho & 1 & \cdots & \rho^{n-4} & \rho^{n-3} \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ \rho^{n-2} & \rho^{n-3} & \rho^{n-4} & \ldots & 1 & \rho \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \cdots & \rho & 1 \end{bmatrix} \quad (6)$$

and $\sigma^2 = \sigma_u^2 = \frac{\sigma_v^2}{(1-\rho^2)}$,

and the inverse of Ω is given as:

$$\Omega^{-1} \frac{1}{1-\rho^2} \begin{bmatrix} 1 & -\rho & 0 & \cdots & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & \cdots & 0 & 0 \\ 0 & -\rho & 1+\rho^2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1+\rho^2 & -\rho \\ 0 & 0 & 0 & \cdots & -\rho & 1 \end{bmatrix} \quad (7)$$

We now search for a suitable transformation matrix P^*. If we consider first an (n-1) x n matrix P^* defined as:

$$P^* = \begin{bmatrix} -\rho & 1 & 0 & \cdots & 0 & 0 \\ 0 & -\rho & 1 & \cdots & 0 & 0 \\ 0 & 0 & -\rho & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -\rho & 1 \end{bmatrix} \quad (8)$$

Multiplying then shows that $P^{*\prime}P^*$ gives an n x n matrix which apart from a proportional constant is identical with Ω^{-1} except for the first elements in the leading diagonal.

Now if we consider an n x n matrix P obtained from P^* by adding a new row to the first row with $\sqrt{1-\rho^2}$ in the first position and zero elsewhere, that is

$$P = \begin{bmatrix} (1-\rho^2)^{\frac{1}{2}} & 0 & 0 & \cdots & 0 & 0 \\ -\rho & 1 & 0 & \cdots & 0 & 0 \\ 0 & -\rho & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -\rho & 1 \end{bmatrix} \quad (9)$$

The difference between P^* and P lies only in the treatment of the first sample observation. P^* is much easier to use provided we are prepared to lose information on the first observation. However, when n is large, the difference is negligible, but in small sample, the difference can be major. Moreover, there is need for Ω or more precisely ρ to be known for the GLS via the transformation matrix P^* and P to be used. This is not often the case; we resort to estimating Ω by $\hat{\Omega}$ to have Feasible Generalized Least Squares estimator (FGLSE). There are several ways of consistently estimating ρ, after which either the P^* or P transformation matrix can be used.

The estimators are: 1. Cochrane and Orcutt Estimator (CORC): Iterative Procedure 2. Maximum Likelihood Estimator (ML)

The Proposed Estimators: The proposed estimators were classified into three (3) groups and are presented as follows:

1. **Two – Stage Estimators (TSE):** The discussion requiring an (n x n) or (n-1 x n) non-singular Matrix P for GLSM transformation, the true autocorrelation value (ρ) was used to obtain the Matrix P to transform the data. Thus, the following estimators are proposed. The algorithms required are as follows:

   (N-1) - Autocorrelation Corrected Generalized Least Square Estimators [(N-1)-AUTOCOGLSE]:

   This estimator requires using (n-1) x n matrix P to transform the data. It can also be referred to as Generalized Least Square Estimator. The procedures are as follows:

   i. Transform the data using true autocorrelation value.

   ii. Apply OLS estimator on the transformed data to obtain the estimates of the parameters of the regression model.

   NOTE: Since these estimators require the use of **true autocorrelation values**, therefore, they do not find relevance in practice.

2. **Two – Process Estimators (TPE):** In Practice, the true autocorrelation value is not known. These estimators are proposed by using the estimated autocorrelation value resulting from

known Feasible Generalized Least Square Estimator to transform the data before any other estimator is applied. The estimators and algorithms required are as follows:

3. Modified Ridge Estimators:

The ridge estimators discussed in section (2.3) require estimation of

$$\hat{K}_i = \frac{\hat{\sigma}^2}{\hat{\alpha}_i^2}$$

where,

$\hat{\sigma}^2$ is the Mean Square Error based on OLS estimation

$\hat{\alpha}_i^2$ is the regression coefficient i based on OLS estimation, i = 1,2.........p.

Now, since there is autocorrelation problem, using OLS estimator is inappropriate. It underestimates the variance of the regression co-efficient and the MSE is biased and inconsistent of σ^2 (Johnston, 1984). The proposed estimators used the appropriate estimators, CORC and ML, to obtain MSE and the regression coefficients. These result into the following proposed estimators. The algorithms required are as follows:

Model formulation

Consider the linear regression model given as:

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \beta_3 x_{3t} + \mu_t \quad (10)$$

Where, $\mu_t = \rho\mu_{t-1} + \varepsilon_{t,}\ |\rho| < 1$,
$t = 1,2,,\ldots\ldots\ldots\ldots n, \quad \varepsilon_t \sim N\,(0, \sigma^2)$.

The regressors are fixed and exhibit different degree of multicollinearity.

The Monte Carlo Experiments

The experiment was replicated (R) one thousand time (1000) and at sample sizes of
n = 10,20,30,50,100 and 250.

Generation of Data: Generation of the Explanatory Variables

a. Correlated Normally distributed Variables

The equations provided and used by [22] were used to generate normally distributed random variables with specified inter-correlation. With p = 3, the equations are:

$$X_1 = \mu_1 + \sigma_1 z_1 \quad (11)$$

$$X_2 = \mu_2 + \lambda_{12}\sigma_{12} + \sqrt{m_{22} z_2} \quad (12)$$

$$X_3 = \mu_3 + \lambda_{13}\sigma_3 z_1 + \frac{m_{23}}{\sqrt{m_{22}}} + \sqrt{n_{33} z_3} \quad (13)$$

Where,

$m_{22} = \sigma^2(1 - \lambda_{12}^2)$

$m_{23} = \sigma_2\sigma_3(\lambda_{23} - \lambda_{12}\lambda_{13})$

$m_{33} = \sigma^2(1 - \lambda_{13}^2)$

and $n_{33} = m_{33} - \frac{m_{23}^2}{m_{22}}$; and $z_i \sim N(0,1)$, $i = 1 , 2 , 3$

By these equations, the inter-correlation matrix has to be positive definite among the independent variables.

In this study,

$\lambda_{12} = \lambda_{13} = \lambda_{23} = \lambda =$ 0.4, 0.6, 0.8, 0.95 and 0.99; $and\ x_i \sim N(0,1)$, $i = 1 , 2 , 3$

b. Correlated Uniformly Distributed Variables

Using the generated correlated normally distributed variables above, $x_i \sim N(0,1)$, $i = 1 , 2 , 3$; the study further utilized the properties of random variables that cumulative distribution function of normal distribution produce U(0,1) without affecting the correlation among variables to generate correlated uniform distributed variables $x_i \sim U(0,1)$, $i = 1 , 2 , 3$ [23]

c. Generation of the error term

The error terms were generated by using the distributional properties of the autocorrelation error terms of AR (1) model given as:

$$u_t \sim N\left(0, \frac{\sigma_e^2}{(1-\rho^2)}\right) \quad (14)$$

Thus, assuming the model start from infinite past, the error terms were generated as follows:

$$u_1 = \frac{\varepsilon_1}{\sqrt{1-\rho^2}} \quad (15)$$

$$u_t = \rho u_{t-1} + \varepsilon_{t,} \quad t = 2,3,4 \ldots\ldots\ldots. n \quad (16)$$

Autocorrelation value (ρ) is varied from 0.4, 0.8, 0.95, and 0.99 but in this study it is set to be zero (ρ = 0).

d. Generation of Dependent Variable

The model parameter values were taken as $\beta_0 = 0$, β_1= 0.8, $\beta_2 = 0.1$ and $\beta_3 = 0.6$. Thus, the dependent variable was also determined. Data were therefore generated for all the specifications of different combinations of multicollinearity, autocorrelation and sample size; a total of one hundred and twenty (120) different combinations (6x5x4) all together.

Estimation methods used in the study

They were categorized into five groups as follows:

1. One – Stage Estimator (OSE): These are existing estimation techniques that require only one stage in their estimation procedures. They are given as follows:

i. Ordinary Least Squares Estimator (OLSE)
ii. Generalized Ridge Estimator (GRE)
iii. Ordinary Ridge Estimator with K – Bayesian (OREKBAY)

iv. Ordinary Ridge Estimator with K-Lukman and Ayinde (OREKLA)

2. Two -Stage Estimators (TSE): These are:

i. (N-1) -Autocorrelation Corrected Generalized Least Square Estimators (N-1) –AUTOCOGLSE]:

ii. (N-1) - Autocorrelation Corrected Generalized Ridge Estimators [(N-1) –AUTOCOGRE]: and others remaining estimators

3. Feasible Generalized Least Square Estimators (FGLSE): These are existing estimators:

i. Corchran Orcutt Estimator (CORC)

ii. Maximum Likelihood Estimator (ML)

4. Two- Process Estimators (TPE): These are:

i. (N-1) - Autocorrelation Corrected Feasible Generalized Least Square Estimators [(N-1) -AUTOCOFGLSE-CORC]:

ii. (N-1) - Autocorrelation Corrected Feasible Generalized Ridge Estimators [(N-1) -AUTOCOFGRE-CORC]: and others remaining estimators

5. Modified Ridge Estimators (MRE): These are:

i. Cochrane Orcutt Modified Generalized Ridge Estimator (CORCMGRE)

ii. Cochrane Orcutt Modified Ordinary Ridge Estimator-KBAY (CORCMORE-KBAY) and others remaining estimators

Criteria for evaluation, examination and comparison of the estimators

Evaluation, examination and comparison of the estimators were done based on their finite sampling properties. These include Bias closest to zero (BAS), Mean Absolute Error (MAE), Variance (VAR) and Mean Square Error (MSE) defined as follows:

i. $\hat{\beta}_i = \frac{1}{1000}\sum_{j=1}^{1000}\hat{\beta}_{ij}$ (17)

ii. $BAS(\hat{\beta}_i) = \left|\frac{1}{1000}\sum_{j=1}^{1000}(\hat{\beta}_{ij} - \beta_i)\right|$ (18)

iii. $MAE(\hat{\beta}_i) = \frac{1}{1000}\sum_{j=1}^{1000}\left|\hat{\beta}_{ij} - \beta_i\right|$ (19)

iv. $VAR(\hat{\beta}_i) = \frac{1}{1000}\sum_{j=1}^{1000}(\hat{\beta}_{ij} - \hat{\beta}_i)^2$ (20)

v. $MSE(\hat{\beta}_i) = Bias(\hat{\beta}_i)^2 + Var(\hat{\beta}_i)$ (21)

RESULTS AND DISCUSSIONS

Results when there is multicollinearity alone in the model

Table 1: Summary of the best estimators on the basis of criteria at different sample sizes with normally and uniformly distributed regressors when multicollinearity alone is in the model

	NORMAL					UNIFORM				
N	**BIAS**	**MAE**	**VAR**	**MSE**	**OVERALL**	**BIAS**	**MAE**	**VAR**	**MSE**	**OVERALL**
10	(N-1)-AUTOCO FGLSE-CORC/CORC (2) **N-AUTOCO FGLSE-ML/ML (3)**	GRE (2) **OREKBAY (3)**	GRE (2) **OREKBAY (3)**	GRE (2) **OREKBAY (3)**	(N-1)-AUTOCO FGLSE-CORC/CORC (2) N-AUTOCO FGLSE-ML/ML (3) GRE (6) **OREKBAY (9)**	**OLSE (4)** N-AUTOCO FGLSE-ML/ML (1)	**OREKBAY (3)** OREKLA (1) MLMGRE (1)	**GRE (4)** OREKBAY (1)	**OREKBAY (3)** OREKLA (1) GRE (1)	OLSE (4) N-AUTOCO FGLSE-ML/ML (1) GRE (5) **OREKBAY (7)** OREKLA (2) MLMGRE (1)
20	**N-AUTOCO FGLSE-ML/ML (3)** OLSE (2)	**OREKBAY (5)**	**OREKBAY (5)**	**OREKBAY (5)**	N-AUTOCO FGLSE-ML/ML (3) OLSE (2) **OREKBAY (15)**	**OLSE (5)**	**OREKBAY (4)** CORCMORE-KLA (1)	**OREKBAY (4)** CORCMORE-KLA (1)	**OREKBAY (4)** CORCMORE-KLA (1)	**OREKBAY (12)** OLSE (5) CORCMORE-KLA (3)
30	**N-AUTOCO FGLSE-ML/ML (5)**	OREKLA (1) **OREKBAY (4)**	OREKLA (1) **OREKBAY (4)**	**OREKBAY (5)**	N-AUTOCO FGLSE-ML/ML (5) **OREKBAY (13)** OREKLA (2)	**N-AUTOCO FGLSE-ML/ML (5)**	CORCMORE-KBAY (1) **OREKBAY (4)** GRE (1)	CORCMORE-KBAY (1) N-AUTOCOGRE-ML (1) GRE (1) **OREKBAY (2)**	CORCMORE-KBAY (1) **OREKBAY (3)** GRE (1)	N-AUTOCO FGLSE-ML/ML (5) **OREKBAY (9)** CORCMORE-KBAY (3) N-AUTOCOGRE-ML (1) GRE (2)

50	**(N-1)-AUTOCO FGLSE-CORC/CORC (5)**	GRE (2) **OREKBAY (3)**	GRE (1) **OREKBAY (4)**	GRE (2) **OREKBAY (3)**	(N-1)-AUTOCO FGLSE-CORC/CORC (5) **OREKBAY (10)** GRE (5)	**N-AUTOCO FGLSE-ML/ML (5)**	GRE (1) **OREKBAY (2) CORCMORE-KLA (2)**	**OREKBAY (3)** CORCMORE-KLA (1) GRE (1)	**GRE (2) OREKBAY (2)** CORCMORE-KLA (1)	N-AUTOCO FGLSE-ML/ML (5) **OREKBAY (7)** GRE (4) CORCMORE-KLA (4)
100	(N-1)-AUTOCO FGLSE-CORC/CORC (1) **N-AUTOCO FGLSE-ML/ML (3)** OLSE (1)	**OREKBAY (5)**	GRE (2) **OREKBAY (3)**	**OREKBAY (4)** N-AUTOCOFOREKBAY (1)	(N-1)-AUTOCO FGLSE-CORC/CORC (1) N-AUTOCO FGLSE-ML/ML (3) **OREKBAY (12)** N-AUTOCOFOREKBAY (1) GRE (2) OLSE (1)	**OLSE (5)**	**OREKBAY (5)**	**OREKBAY (3)** GRE (1) N-AUTOCO FOREKBAY (1)	**OREKBAY (5)**	OLSE (5) **OREKBAY (13)** GRE (1) N-AUTOCO FOREKBAY (1)
250	**OLSE (4)** (N-1)-AUTOCO FGLSE-CORC/CORC (1)	**OREKBAY (4)** GRE (1)	**OREKBAY (5)**	**OREKBAY (5)**	OLSE (4) **OREKBAY (14)** GRE (1) (N-1)-AUTOCO FGLSE-CORC/CORC (1)	**OLSE (4)** (N-1)-AUTOCO FGLSE CORC/CORC (1)	**OREKBAY (4)** CORCMORE-KLA (1)	**OREKBAY (4)** CORCMORE-KLA (1)	**OREKBAY (4)** CORCMORE-KLA (1)	OLSE (4) (N-1)-AUTOCO FGLSE CORC/CORC (1) **OREKBAY (12)** CORCMORE-KLA (3)

From Table 1, the following are observed about the best estimator(s) under each criterion.

Graphical Representation of choice Estimaors when Multicollinearity is in the Model with Normally and Uniformly Distributed Regressors.

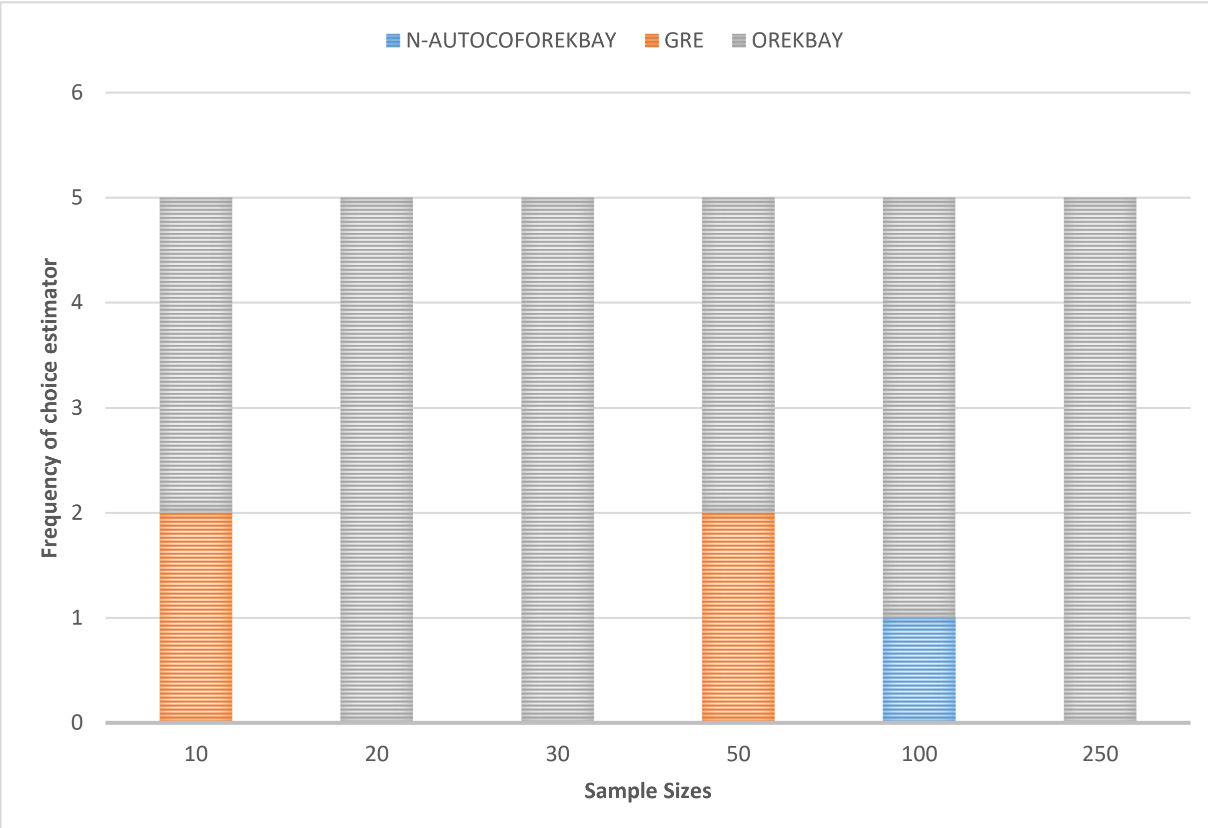

Figure 1A: Compound Bar Chart of Choice Estimators with Normally Distributed Regressors when Multicollinearity alone is in the model (MSE Criterion)

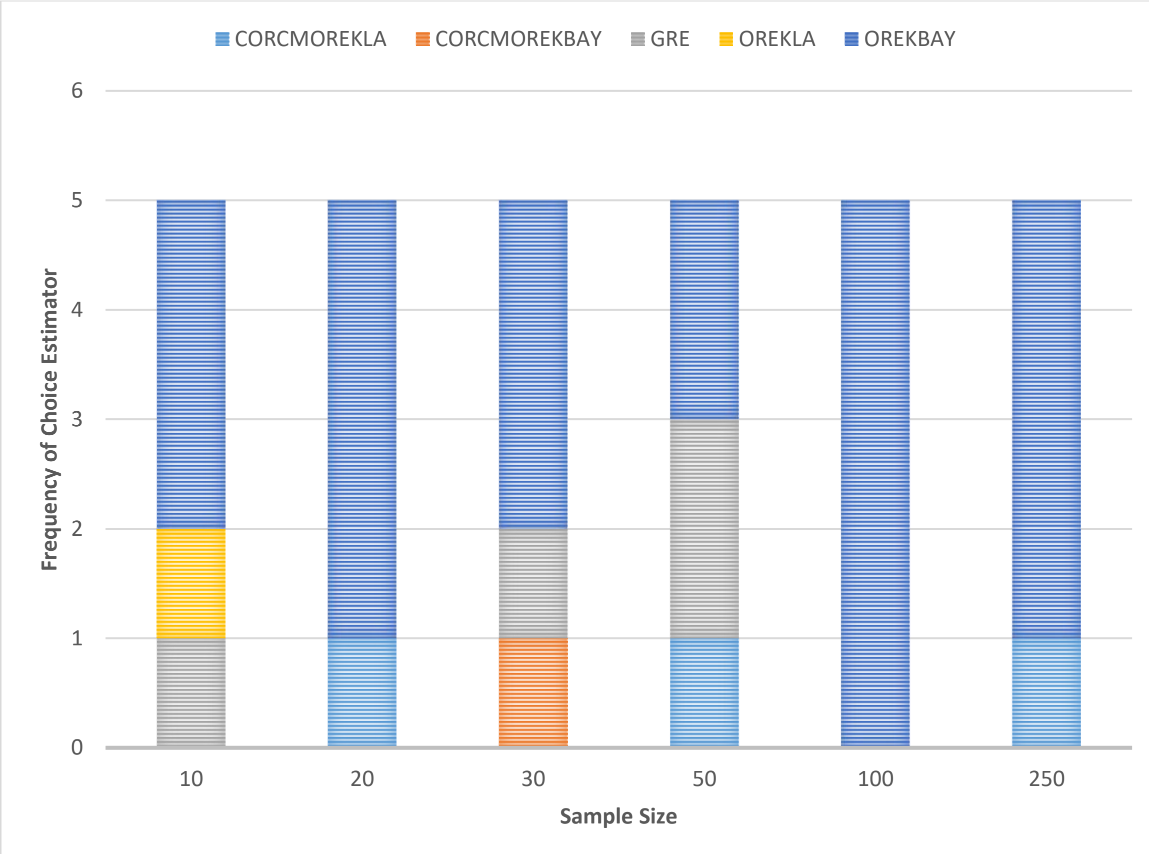

Figure 1B: Compound Bar Chart of Choice Estimators with Uniformly Distributed Regressors when Multicollinearity alone is in the model (MSE Criterion)

Bias

i. With normal regressors, the best estimator is N-AUTOCOFGLSE-ML/ML except when n=50 and n=250. At the formal instance, the (N-1)- AUTOCOFGLSE-CORC /CORC is best while at the latter the OLSE is best.

ii. With uniform regressors, the best estimator is either N-AUTOCOFGLSE-ML/ML or OLSE. The OLSE is best when the sample sizes are small (n=10 and n=20) and large (n=100 and n=250). Moreover, the N-AUTOCOFGLSE-ML/ML is best when the sample sizes are moderate (n=30 and n=50).

iii. These results supported the fact that OLSE is unbiased as long as multicollinearity is not perfect.[5] and [14]

iv. It should be noted that N-AUTOCOFGLSE-ML and (N-1)-AUTOCOFGLSE- CORC are among the newly proposed estimators.

Variance, Mean Absolute Error and Mean Square Error

i. With the normal regressors, the best estimator is OREKBAY especially under the mean square error criterion. However, the GRE (when the sample size is small, n=10) and medium, n=50) and N-AUTO COFORE KBAY (when the sample size is large, n=100) also perform well. Furthermore, it also performs well under the variance and mean absolute error criteria when the sample sizes are large, n=100 for variance criterion, and n=250 for mean absolute error criterion. Moreover, the OREKLA also does well under variance (n=100) and mean absolute error (n=250) criteria. The frequency of each being chosen as a choice estimator is presented pictorially in Figure 1A.

ii. With uniform regressors, the best estimator is the OREKBAY except under the variance criterion where the GRE is best when the sample size is small (n=10). Moreover, under the mean square error criterion, CORCMOREKLA, CORCMOREKBAY, OREKLA and GRE also perform well occasionally. In addition, CORCMOREKLA is also best under variance criterion when the sample size is medium (n=50). Moreover, other estimators that do well include MLMGRE, N-AUTOCOGRE-ML and N-AUTOCOFOREKBAY. The frequency of each estimator being chosen as a choice estimator is presented pictorially in Figure 1B.

iii. It should be noted that MLGRE, CORCMOREKLA, CORCMOREKBAY, N-AUTOCOGRE-ML and N-AUTOCOFOREKBAY that also do well are among the newly proposed estimators.

iv. Although OLSE is unbiased but these results agreed with the finding that the Ridge Estimators are more efficient than the OLSE in the presence of multicollinearity.[14]

CONCLUSION

It can be observed that the results of OREKBAY and N-AUTOCOOREKBAY, GRE and N-AUTOCOGRE are the same. This is because there is no autocorrelation problem in the results being presented. In the presence of multicollinearity with normally distributed regressor, the best estimator is generally the existing estimator OREKBAY except under the bias criterion. At this instance, the estimators FGLSE, ML and CORC are the best. Also, with uniformly distributed regressor, it was observed that the best estimator under all criteria is the existing estimator OREKBAY except under the bias criterion. At this instance the OLSE and FGLSE - ML are the best.

REFERENCES

1. BOWERMAN, B.L. & O‟ CONNELL, R.T. (2006). Linear Statistical Models and Applied Approach, *Boston. PWS-KENT Publishing. 4th Edition.*
2. CHARTTERJEE, S., HADI, A.S. & PRICE, B. (2000). Regression by Example, 3rd Edition. *A Wiley-Interscience Publication, John Wiley and Sons.*
3. KMENTA, J. (1986). Elements of econometrics. *Macmillan Publishing Co. New York*.
4. HAWKING, R.R. & PENDLETON, O.J. (1983). The regression dilemma. *Commun Stat.- Theo. Meth.***12:** 497-527.
5. GUJARATI, D.N. & PORTER, D.C. (2009). Basic Econometrics. *Mc Graw-Hill, New York. 5th Edition*
6. JAMES, S. (1956). Inadmissibility of the usual estimator for the mean of a Multivariate Distribution. *Proc. Third Berkelysymp.math.statist.prob*. 197-206
7. HOERL, A.E. & KENNARD, R.W. (1970). Ridge regression biased estimation for non-orthogonal problems, *Technometrics,* **8:** 27 – 51.
8. MASSEY, W.F. (1965). Principal Component Regression in exploratory statistical research. *Journal of the American Statistical Association,* **60:** 234– 246
9. MANSFIELD, W. & GUNST, A. (1977). An Analytic Variable Selection Technique for Principal Component Regression. *App/Stat,* **26(**1): 34 – 40
10. PARIS, Q. (2001a). Multicollinearity and Maximum Entropy Leuven Estimator. *Economic Bulletin***,** **3(**11**):** 1-9
11. WOLD, H. (1966). Estimation of principal component and related model by iterative Least Squares. In P.R. Krishnainh [ed] *Multivariate Analysis. New York Academic Press*, 391-420
12. AYINDE, K., APATA, E.O. & ALABA, O.O. (2012). Estimators of Linear Regression Model and Prediction under some Assumptions Violation. *Open Journal of Statistics,* **2:** 534 – 546.
13. FORMBY, T.B., HILL, R.C. & JOHNSON, S.R. (1984). Advance Econometric Methods. *Springer-Verlag, New York, Berlin, Heidelberg, London, Paris, Tokyo.* 2nd Edition.
14. AYINDE, K. & IPINYOMI, R.A. (2007). A comparative study of the OLS and some GLS estimators when normally distributed regressors are stochastic. *Trend in Applied Sciences Research,* **2(4):** 354-359.
15. COCHRANE, D. & ORCUTT, G.H. (1949). Application of Least Square to relationship containing autocorrelated error terms. *Journal of American Statistical Association,* **44:** 32–61.
16. KADIYALA, K.R. (1968). Testing for the independent of Regression Disturbances, *ECTRA,* **38:** 71-117
17. RAO, P. & GRILICHES, Z. (1969). Small sample properties of several two-stage regression methods in the context of autocorrelation error. *Journal of American Statistical Association,* **64:** 251 – 272.
18. SPITZER, J.J. (1979). A Monte Carlo investigation of the (Box-Cox) transformation in small samples. *Journal of the America Statistical Association.* **73**: 488-495
19. HOERL, A.E., KENNARD, R.W. & BALDWIN, K.F. (1975). Ridge Regression: Some Simulation. *Journal of Communication in Statistics.*

20. AYINDE, K., LUKMAN, A.F. & AROWOLO, O.T. (2015). Combined Parameters Estimation methods of Linear Regression Model with Multicollinearity and Autocorrelation. *Journal of Asian Scientific Research,* **5(**5**)**: 243-250.
21. SCLOVE, S. (1973). Improved estimators for coefficient in Linear regression. *Journal of American Statistical Association* **63,** 596-606
22. AYINDE, K**. &** ADEGBOYE, O.S. (2010). Equations for generating normally distributed random variables with specified intercorrelation. *Journal of Mathematical Sciences,* **21(**2): 183–203**.**
23. SCHUMANN, E. (2009). Generating Correlated Uniform Variate. *http:// comisef.wikidot.com / tutorial: correlateduniformvariate.*