Researcher

Websites:
http://www.sciencepub.net
http://www.sciencepub.net/researcher

Emails:
editor@sciencepub.net
marslandresearcher@gmail.com

MARSLAND PRESS
Multidisciplinary Academic Journal Publisher

A SIMULATION STUDY OF MODIFIED RIDGE ESTIMATORS FOR HANDLING MULTICOLLINEARITY PROBLEM

Bello, Abimbola Hamidu

Department of Statistics, School of Physical Sciences, Federal University of Technology, Akure.- Nigeria
habello@futa.edu.ng

Abstract
The study is on Proposed Modified ridge estimators for solving Multicollinearity problems in Linear Regression. A monte Carlo experiment was conducted using TSP 5.0 Statistical Package with sample sizes of 10,20,30,50,100 and 250 and at different multicollinearity levels of 0.2, 0.4, 0.6, 0.8, 0.95 and 0.99 which are categorize as low, moderate and high levels. The data generated were replicated 1000 time with normally and uniformly distributed regressors. The methods employed the use Ordinary Ridge and Generalized Ridge estimators along with CORC and ML estimators. The proposed estimators used the appropriate estimators, CORC and ML, to obtain MSE and the regression coefficients which is now used in developing the proposed estimators. The results of the analysis revealed that among the proposed modified ridge estimators MLMOREKBAY, CORCMOREKBAY, CORCMOREKLA and MLMGRE perform excellently well at different sample sizes and can therefore be used to solve problem of multicollinearity in any dataset. Also, the existing estimator OREKBAY compete favorably well with the proposed estimators in addressing the problem of multicollinearity.
[Bello, Abimbola Hamidu. **A SIMULATION STUDY OF MODIFIED RIDGE ESTIMATORS FOR HANDLING MULTICOLLINEARITY PROBLEM**. *Researcher* 2022;14(10):9-20] ISSN 1553-9865(print); ISSN 2163-8950 (online) http://www.sciencepub.net/researcher. 03. doi:10.7537/marsrsj141022.03.

Keywords: Multicollinearity, CORC - Cochrane-Orcutt Estimator, ML – Maximum Likelihood Estimator, MSE – Mean Square Error and Monte Carlo Experiment.

1.0 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. 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 linear relationship among some of the predictor variables in the data. Bowerman and O ′Connell (2006) 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 (Chartterjee, Hadi and Price, 2000).

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, 1986). 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 Pendleton, 1983; Gujarati and Porter, 2009).

Various authors have worked on estimators for solving multicllinearity among which are: James (1956) which introduced the stein estimator as a method for handling multicollinearity. Massey (1965) also apply Principal Components Regression to handle the problem of multicollinearity by eliminating the model instability and reducing the variances of the regression coefficients. The sample correlation for any pair of components is observed to be zero. Wold (1966) introduced the Partial Least Squares Regression into handling multicollinearity problem. This method is similar to the method of the Principal Components Analysis. However, it utilizes the dependent variable together with the explanatory variables. Hoerl and Kennard (1970). Proposed the ridge estimator for dealing with multicollinearity in a regression model. It is the modification of the OLS that allows biased estimation of the regression coefficients. Also some years later, Hoerl, *et al* (1975) Proposed ridge estimator for dealing with multicollinearity and provided optimal value K of the ridge parameter given as: $\widehat{K}_i = \frac{\hat{\sigma}^2}{\hat{\alpha}_i^2}$; where $\hat{\sigma}^2$ is an unbiased estimator of error-variance from OLS estimation and $\hat{\alpha}_i^2$ is also regression coefficient from OLS estimation

This research work is aimed at developing a modified ridge estimators and compare it with the existing estimators to address the problem of multicollineariy.

2.0 Methodology

The general form of Linear Regression Model is: Y = X β + U (1)

Where,

Y is an (n x 1) vector of observations of the dependent variable. X matrix is an n x (k+1) full rank matrix of observable and fixed values of the explanatory variables. β is a ((k+1) x1) vector of unknown parameters to be estimated. U is (n x 1) vector of random error. The parameter estimate is defined as:

$$\hat{\beta} = (X^1X)^{-1}X^1Y \tag{2}$$

The Ridge Estimators

Ridge regression is a method of biased linear estimation which has been shown to be more efficient than the OLS estimator when data sets exhibit multicollinearity. Hoerl and Kennard (1970). The estimator of β is defined as:

$$\hat{\beta}(k) = (X'X + KI)^{-1}X'y \tag{3}$$

The constant K is known as bias or ridge parameter and it yield minimum MSE compared to the OLS estimator. This study employed the followings estimators:

1. **Generalized Ridge Estimator**

The optimum value of K had been obtained by Hoerl, Kennard and Baldwin (1975) as:

$$K_i = \frac{\sigma^2}{\alpha_i^2} \quad i = 1,2,3,\dots,p. \tag{4}$$

Since σ^2 and α_i^2 are generally unknown, the K_i needs to be estimated.

2. **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 that of Lukman and Ayinde (2017) and Sclove (1973) given as:

$$\widehat{K}_{LA} = \frac{\hat{\sigma}^2}{[Max(\hat{\alpha})]^2} \tag{5}$$

Sclove (1973) suggested an empirical K-Bayesian Ridge Parameter given as:

$$\widehat{K}_{BAY} = \frac{(SSR/n-p)}{(\sum y2\ - SSR/\ trace(X'X))} \tag{6}$$

Where SSR = Sum of Square of Regression

3. **Cochrane-Orcutt Estimator**
4. **Maximum Likelihood Estimator**

Model Formulation

Consider the linear regression model given as:

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \mu \tag{7}$$

Where, $\mu \sim N\ (0, \sigma^2)$.

The regressors are fixed and exhibit different degree of multicollinearity.

The Monte Carlo Experiments

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

Correlated Normally distributed Variables

The equations provided and used by Ayinde (2007a, b) and Ayinde and Adegboye (2010) were used to generate normally distributed random variables with specified inter-correlation. With p = 3, the equations are:

$$X_1 = \mu_1 + \sigma_1z_1 \tag{8}$$

$$X_2 = \mu_2 + \lambda_{12}\sigma_{12} + \sqrt{m_{22}z_2} \tag{9}$$

$$X_3 = \mu_3 + \lambda_{13}\sigma_3z_1 + \frac{m_{23}}{\sqrt{m_{22}}} + \sqrt{n_{33}z_3} \tag{10}$$

Where $m_{22} = \sigma^2{}_2(1-\lambda_{12}^2)$

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

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

$\lambda_{12} = \lambda_{13} = \lambda_{23} = \lambda =$ 0.2, 0.4, 0.6, 0.8, 0.95 *and* 0.99; *and* $x_i \sim N(0,1)$, $i = 1, 2, 3$

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 $x_i \sim U(0,1)$, $i = 1, 2, 3$ (Schumann, 2009)

The error terms were generated with $\mu \sim N(0, \sigma^2)$.
The model parameter values were taken as $\beta_0 = 0$, $\beta_1 = 0.8$, $\beta_2 = 0.1$ and $\beta_3 = 0.6$.

Modified Ridge Estimators:
The ridge estimators discussed in (3) require estimation of

$$\hat{K}_i = \frac{\hat{\sigma}^2}{\hat{\alpha}_i^2} \qquad (12)$$

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.

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:

Cochrane-Orcutt Modified Generalized Ridge Estimator (CORCMGRE):

The procedures are as follows:

i. CORC estimator is used instead of the OLS estimator to obtain the regression co-efficients and the MSE,

ii. Thereafter, the different $\hat{K}_i$ are obtained and used in the ridge estimator to obtain the regression co-efficient of the model.

All others proposed estimators follows the same procedure.

Cochrane Orcutt Modified Ordinary Ridge Estimator-KBAY
(CORCMORE-KBAY)

Cochrane Orcutt Modified Ordinary Ridge Estimator-KLA
(CORCMORE-KLA)

Maximum Likelihood Modified Generalized Ridge Estimator (MLMGRE)

Maximum Likelihood Modified Ordinary Ridge Estimator-KBAY
(MLMORE-KBAY)

Maximum Likelihood Modified Ordinary Ridge Estimator-KLA
(MLMORE-KLA)

The existing estimation techniques used are:
Ordinary Least Squares Estimator (OLSE)
Generalized Ridge Estimator (GRE)
Ordinary Ridge Estimator with K –Bayesian (OREKBAY)
Ordinary Ridge Estimator with K-Lukman and Ayinde (OREKLA)

Criteria for Evaluation

Mean Square Error (MSE) defined as:

$$\text{MSE}(\hat{\beta}_i) = \text{Bias}(\hat{\beta}_i)^2 + \text{Var}(\hat{\beta}_i) \qquad (13)$$

3.0 Data Analysis and Results

The table showing the Mean Square Error (MSE) at different sample sizes and levels of multicollinearity with Normally and Uniformly distributed Regressors are in tables 1 and 2

	Table1		Mean Square Error of the Estimators at different levels of multicollinearity (Normal Regressors)										
		OLSE	GRE	OREKBAY	OREKLA	CORC	CORC MGRE	CORCMO REKBAY	CORCMO REKLA	ML	MLMGRE	MLMO REKBAY	MLMORE KLA
0	10	2.0892	0.625	0.7647	1.0093	5.9414	0.803	0.9936	1.2647	2.8095	0.6712	0.8689	1.0582
	20	1.4456	0.646	0.5922	0.6446	1.8352	0.663	0.6244	0.6539	1.5987	0.6478	0.6282	0.651
	30	0.8009	0.477	0.441	0.4425	0.8751	0.473	0.4509	0.4431	0.8401	0.4759	0.4549	0.4448
	50	0.525	0.329	0.3385	0.333	0.5444	0.33	0.3414	0.3311	0.5391	0.3307	0.3433	0.3331
	100	0.2425	0.243	0.1868	0.1932	0.2487	0.242	0.1874	0.1929	0.2454	0.2423	0.1878	0.1932
	250	0.1013	0.246	0.0909	0.0921	0.103	0.245	0.091	0.0921	0.1026	0.2452	0.091	0.0921
0.2	10	2.4032	0.667	0.8	1.1015	6.8277	0.858	1.0687	1.4172	3.2391	0.7122	0.9206	1.1658
	20	1.8048	0.726	0.6203	0.7226	2.3525	0.751	0.6612	0.7461	1.988	0.7379	0.6627	0.7278
	30	0.924	0.507	0.4787	0.4795	1.0063	0.505	0.4905	0.4814	0.9718	0.5084	0.4953	0.4834
	50	0.6259	0.337	0.3753	0.3708	0.6498	0.341	0.379	0.3687	0.6441	0.3413	0.3814	0.3712
	100	0.2906	0.259	0.2126	0.222	0.2993	0.257	0.2134	0.2217	0.2946	0.258	0.2139	0.2221
	250	0.1215	0.259	0.1067	0.1074	0.1236	0.257	0.1067	0.1073	0.1231	0.2573	0.1068	0.1074
0.4	10	3.0455	0.73	0.8555	1.2872	8.6544	0.971	1.197	1.7146	4.1104	0.7914	1.0097	1.3734
	20	2.3396	0.847	0.645	0.8347	3.1829	0.9	0.6978	0.8911	2.5592	0.8674	0.6945	0.8424
	30	1.1527	0.561	0.5346	0.5436	1.2564	0.56	0.5493	0.5466	1.2155	0.5636	0.5556	0.5479
	50	0.7982	0.37	0.4269	0.4271	0.8297	0.377	0.4319	0.4252	0.8233	0.3775	0.4352	0.4282
	100	0.3719	0.289	0.2508	0.2675	0.3846	0.286	0.252	0.2673	0.3778	0.2872	0.2526	0.2677
	250	0.1557	0.272	0.1319	0.1326	0.1586	0.27	0.132	0.1325	0.1578	0.2702	0.132	0.1326
0.6	10	4.6661	0.85	0.9412	1.7528	12.4787	1.218	1.4246	2.4288	6.2723	0.9507	1.1647	1.8841
	20	3.3034	1.038	0.6596	1.0452	4.7692	1.17	0.7286	1.1702	3.5856	1.0764	0.7166	1.058
	30	1.7189	0.683	0.6213	0.6907	1.8708	0.682	0.6421	0.6956	1.8162	0.6864	0.6514	0.6976
	50	1.1395	0.448	0.4989	0.5267	1.1854	0.457	0.5059	0.5258	1.1775	0.4574	0.5107	0.5296
	100	0.5372	0.345	0.3118	0.3517	0.5579	0.341	0.3138	0.3515	0.5465	0.3423	0.3149	0.3519
	250	0.2258	0.182	0.1781	0.1786	0.2302	0.184	0.1783	0.1784	0.2288	0.1838	0.1784	0.1786
0.8	10	10.7102	1.172	1.0103	2.6782	23.6285	1.763	1.7694	3.6841	14.0662	1.352	1.3519	2.8915
	20	5.9339	1.375	0.6354	1.6632	9.4018	1.845	0.7215	1.97	6.427	1.4621	0.6969	1.691
	30	3.9683	1.12	0.7028	1.7362	4.3236	1.133	0.7342	1.7257	4.2064	1.1358	0.7494	1.7415
	50	2.1444	0.654	0.5896	0.7884	2.231	0.665	0.6002	0.786	2.2185	0.6678	0.6078	0.792
	100	1.0394	0.457	0.416	0.5664	1.0834	0.456	0.4199	0.5663	1.0585	0.4565	0.422	0.5663
	250	0.4392	0.256	0.284	0.2868	0.4486	0.257	0.2844	0.2867	0.4451	0.2572	0.2848	0.287
0.95	10	47.1169	2.637	0.8443	11.166	89.4567	3.963	1.8808	14.472	63.9459	3.0054	1.1254	11.7882
	20	20.5864	2.859	0.5102	5.4679	35.7444	4.66	0.5787	6.4649	22.5105	3.1209	0.554	5.6126
	30	21.6115	3.472	0.512	6.4217	23.6071	3.486	0.5323	6.4912	23.0155	3.5062	0.5434	6.5152
	50	8.1293	1.62	0.5447	2.3838	8.4583	1.655	0.5554	2.3698	8.416	1.664	0.5641	2.3883
	100	4.0185	0.976	0.4731	1.2112	4.1898	0.985	0.4793	1.2189	4.0984	0.9803	0.4829	1.2112
	250	1.7692	0.531	0.4813	0.652	1.8115	0.533	0.4831	0.6527	1.7928	0.5333	0.4843	0.6523
0.99	10	2.25E+02	9.001	0.5434	55.189	4.24E+02	13.91	1.3036	69.6719	3.17E+02	10.5314	0.7352	58.6597
	20	97.3444	10.55	0.3901	26.416	1.71E+02	16.54	0.4133	30.2446	1.07E+02	11.1878	0.4048	26.9762
	30	1.25E+02	14.94	0.3493	37.029	1.37E+02	14.71	0.3536	37.2221	1.33E+02	14.865	0.3565	37.4465
	50	40.2337	5.497	0.3701	11.643	41.8744	5.636	0.3737	11.5461	41.6644	5.684	0.3772	11.6398
	100	19.8406	3.052	0.3485	5.5	20.6363	3.092	0.3513	5.5123	20.2349	3.066	0.3531	5.4746
	250	9.0298	1.614	0.4128	2.5202	9.2523	1.615	0.4144	2.5347	9.1437	1.6194	0.4155	2.5276

	Table2		Mean Square Error of the Estimators at different levels of multicollinearity (Uniform Regressors)										
		OLSE	GRE	OREKBAY	OREKLA	CORC	CORC MGRE	CORCMO REKBAY	CORCMO REKLA	ML	MLMGRE	MLMO REKBAY	MLMO REKLA
0	10	0.3085	0.276	0.2751	0.2834	0.4853	0.303	0.2778	0.2841	0.4013	0.2908	0.2753	0.2821
	20	0.1643	0.26	0.1486	0.2532	0.2176	0.258	0.1488	0.2518	0.1823	0.2567	0.1488	0.2339
	30	0.129	0.176	0.1208	0.2053	0.1375	0.182	0.1205	0.198	0.1359	0.1801	0.1205	0.1964
	50	0.0627	0.062	0.0636	0.0635	0.0655	0.062	0.0634	0.0635	0.0652	0.0618	0.0634	0.0635
	100	0.0292	0.028	0.0295	0.0293	0.0299	0.028	0.0295	0.0293	0.0297	0.0277	0.0295	0.0293
	250	0.0116	0.012	0.0116	0.0115	0.0118	0.012	0.0116	0.0115	0.0117	0.0116	0.0116	0.0115
0.2	10	0.3078	0.301	0.2632	0.2537	0.4968	0.311	0.2703	0.2615	0.395	0.3087	0.267	0.2566
	20	0.183	0.27	0.1561	0.1589	0.2415	0.263	0.1574	0.1597	0.2028	0.266	0.1575	0.1593
	30	0.1306	0.173	0.118	0.1222	0.1389	0.176	0.1181	0.1219	0.1371	0.1752	0.1184	0.1221
	50	0.0627	0.062	0.0636	0.0635	0.0655	0.062	0.0634	0.0635	0.0652	0.0618	0.0634	0.0635
	100	0.0307	0.038	0.0302	0.0303	0.0314	0.039	0.0302	0.0303	0.0312	0.0388	0.0302	0.0303
	250	0.0123	0.19	0.0122	0.0123	0.0125	0.189	0.0122	0.0123	0.0125	0.189	0.0122	0.0123
0.4	10	0.3539	0.297	0.2813	0.2806	0.5814	0.311	0.2948	0.2844	0.4509	0.3045	0.29	0.2823
	20	0.2256	0.281	0.1811	0.1842	0.3004	0.275	0.1835	0.1861	0.25	0.277	0.1838	0.185
	30	0.1306	0.173	0.118	0.1222	0.1389	0.176	0.1181	0.1219	0.1371	0.1752	0.1184	0.1221
	50	0.0809	0.067	0.0761	0.0767	0.0839	0.068	0.0762	0.0767	0.0835	0.0679	0.0762	0.0767
	100	0.0364	0.063	0.0352	0.0355	0.0373	0.064	0.0352	0.0355	0.037	0.0631	0.0352	0.0355
	250	0.0148	0.167	0.0146	0.0147	0.0151	0.167	0.0146	0.0147	0.015	0.1667	0.0146	0.0147
0.6	10	0.482	0.357	0.3411	0.3335	0.802	0.358	0.367	0.3397	0.616	0.3525	0.3591	0.3365
	20	0.3173	0.304	0.2318	0.2379	0.4306	0.306	0.2364	0.2421	0.3525	0.3022	0.2368	0.2395
	30	0.2125	0.189	0.1758	0.1808	0.2256	0.19	0.1766	0.1797	0.2223	0.1893	0.1774	0.1806
	50	0.1116	0.09	0.1008	0.1025	0.1156	0.09	0.1009	0.1024	0.1151	0.0902	0.1011	0.1025
	100	0.0501	0.1	0.0474	0.0478	0.0515	0.101	0.0474	0.0478	0.051	0.1002	0.0474	0.0478
	250	0.0206	0.159	0.0202	0.0202	0.021	0.159	0.0202	0.0202	0.0209	0.1592	0.0202	0.0202
0.8	10	0.9107	0.438	0.4968	0.5367	1.5348	0.469	0.5623	0.5754	1.1859	0.4422	0.5428	0.5513
	20	0.5985	0.375	0.3466	0.3373	0.8331	0.401	0.3579	0.3347	0.667	0.3803	0.3586	0.3372
	30	0.4202	0.248	0.2955	0.313	0.4467	0.246	0.2975	0.3099	0.4401	0.2472	0.3001	0.3125
	50	0.2084	0.157	0.1705	0.1699	0.2156	0.161	0.1709	0.1692	0.2147	0.1607	0.1714	0.1698
	100	0.0941	0.159	0.0842	0.0849	0.0968	0.158	0.0843	0.0848	0.0958	0.158	0.0844	0.0849
	250	0.039	0.181	0.0374	0.0375	0.0396	0.181	0.0374	0.0375	0.0394	0.1809	0.0374	0.0375
0.95	10	3.5904	0.784	0.8026	1.124	6.1401	1.014	1.057	1.3224	4.8246	0.8817	0.961	1.1945
	20	2.2914	0.739	0.5477	0.7626	3.2511	0.887	0.5865	0.8118	2.5589	0.7745	0.5861	0.7728
	30	1.758	0.521	0.5781	0.7215	1.8731	0.524	0.5875	0.7132	1.8465	0.5272	0.5999	0.7241
	50	0.7985	0.348	0.4112	0.4114	0.8262	0.353	0.4135	0.41	0.8229	0.3545	0.417	0.4136
	100	0.3638	0.284	0.2444	0.2637	0.3738	0.285	0.245	0.2634	0.3699	0.2855	0.2457	0.2639
	250	0.1509	0.257	0.1282	0.1282	0.1535	0.255	0.1282	0.1281	0.1526	0.2552	0.1283	0.1282
0.99	10	18.0601	2.033	0.734	4.8455	31.1254	2.738	1.2071	5.9832	24.6891	2.2686	0.9643	5.2011
	20	11.303	2.067	0.4752	3.1119	16.0851	2.756	0.5191	3.401	12.6218	2.2448	0.5153	3.1593
	30	9.1048	1.649	0.5344	2.8059	9.7114	1.672	0.5463	2.7875	9.5768	1.6763	0.5625	2.8173
	50	3.9567	0.961	0.5477	1.2076	4.097	0.978	0.5532	1.1989	4.0807	0.9832	0.5623	1.2105
	100	1.8091	0.568	0.4538	0.6103	1.857	0.568	0.4559	0.6097	1.8386	0.569	0.4592	0.6108
	250	0.7494	0.448	0.3736	0.3879	0.7628	0.446	0.3739	0.388	0.7579	0.4455	0.3746	0.3884

The table below shows the frequency of the Rank of the best Five (5) estimators at different levels of multicollinearity with Normally and Uniformly Regressors.

Table 1: Ranks of the Best 5 Estimators when there is Multicollinearity Problem

		Normally Distributed Regressors		Uniformly Distributed Regressors	
Multicollinearity. Levels	N	Estimators	Frequency	Estimators	Frequency
Low (0.2 – 0.4)	10	GRE OREKBAY CORCMGRE MLMGRE MLMOREKBAY	2 2 2 2 2	OREKBAY OREKLA CORCMOREKLA MLMOREKBAY MLMOREKLA	2 2 2 2 2
	20	GRE OREKBAY OREKLA CORCMOREKBAY MLMOREKBAY MLMOREKLA	1 2 2 2 2 1	OREKBAY OREKLA CORCMOREKBAY MLMOREKBAY MLMOREKLA	2 2 2 2 2
	30	OREKBAY OREKLA CORCMOREKBAY CORCMOREKLA MLMOREKLA	2 2 2 2 2	OREKBAY CORCMOREKBAY CORCMOREKLA MLMOREKBAY MLMOREKLA	2 2 2 2 2
	50	GRE OREKBAY OREKLA CORCMGRE CORCMOREKLA MLMGRE	2 1 1 2 2 2	OLSE GRE OREKBAY CORCMGRE CORCMOREKBAY MLMGRE MLMOREKBAY	1 2 1 2 1 2 1
	100	OREKBAY OREKLA CORCMOREKBAY CORCMOREKLA MLMOREKLA	2 2 2 2 2	OREKBAY OREKLA CORCMOREKBAY CORCMOREKLA MLMOREKLA	2 2 2 2 2
	250	OREKBAY CORCMOREKBAY CORCMOREKLA MLMOREKBAY MLMOREKLA	2 2 2 2 2	OREKBAY CORCMOREKBAY CORCMOREKLA MLMOREKBAY MLMOREKLA	2 2 2 2 2
	10	GRE OREKBAY CORCMGRE MLMGRE MLMOREKBAY	2 2 2 2 2	GRE OREKBAY OREKLA CORCMGRE CORCMOREKLA MLMGRE MLMOREKLA	1 2 2 1 1 2 1
	20	GRE OREKBAY OREKLA CORCMOREKBAY MLMGRE MLMOREKBAY	2 2 1 2 1 2	OREKBAY OREKLA CORCMOREKBAY CORCMOREKLA MLMOREKBAY MLMOREKLA	2 2 2 1 1 2

Moderate (0.6 – 0.8)	30	GRE OREKBAY CORCMGRE CORCMOREKBAY MLMOREKBAY	2 2 2 2 2	GRE OREKBAY CORCMGRE CORCMOREKBAY CORCMOREKLA MLMGRE MLMOREKBAY MLMOREKLA	1 2 1 2 1 1 1 1
	50	GRE OREKBAY CORCMGRE CORCMOREKBAY MLMGRE MLMOREKBAY	2 2 2 2 1 1	GRE OREKBAY CORCMGRE CORCMOREKBAY CORCMOREKLA MLMGRE MLMOREKLA	2 1 2 1 1 2 1
	100	OREKBAY CORCMGRE CORCMOREKBAY MLMGRE MLMOREKBAY	2 2 2 2 2	OREKBAY OREKLA CORCMOREKBAY CORCMOREKLA MLMOREKBAY	2 2 2 2 2
	250	GRE OREKBAY OREKLA CORCMGRE CORCMOREKBAY CORCMOREKLA MLMGRE MLMOREKBAY	1 2 1 1 2 1 1 1	OREKBAY CORCMOREKBAY CORCMOREKLA MLMOREKBAY MLMOREKLA	2 2 2 2 2
High (0.95– 0.99)	10	GRE OREKBAY CORCMOREKBAY MLMGRE MLMOREKBAY	2 2 2 2 2	GRE OREKBAY CORCMGRE CORCMOREKBAY MLMGRE MLMOREKBAY	2 2 1 1 2 2
	20	GRE OREKBAY CORCMOREKBAY MLMGRE MLMOREKBAY	2 2 2 2 2	GRE OREKBAY CORCMGRE CORCMOREKBAY MLMGRE MLMOREKBAY	2 2 1 1 2 2
	30	GRE OREKBAY CORCMGRE CORCMOREKBAY MLMGRE MLMOREKBAY	1 2 2 2 1 2	GRE OREKBAY CORCMGRE CORCMOREKBAY MLMGRE MLMOREKBAY	2 2 2 2 1 1
	50	GRE OREKBAY CORCMGRE CORCMOREKBAY MLMOREKBAY	2 2 2 2 2	GRE OREKBAY CORCMGRE CORCMOREKBAY CORCMOREKLA MLMGRE MLMOREKBAY	2 2 2 1 1 1 1

	100	GRE OREKBAY CORCMOREKBAY MLMGRE MLMOREKBAY	2 2 2 2 2	GRE OREKBAY OREKLA CORCMGRE CORCMOREKBAY CORCMOREKLA MLMOREKBAY	1 2 1 1 2 1 2
	250	GRE OREKBAY CORCMGRE CORCMOREKBAY MLMOREKBAY	2 2 2 2 2	OREKBAY OREKLA CORCMOREKBAY CORCMOREKLA MLMOREKBAY MLMOREKLA	2 2 2 2 1 1

3.1 Results of the analysis

Low Level of Multicollinearity

With normally distributed regressors, the best estimator with highest frequency at different sample sizes are OREKBAY, MLMOREKBAY and CORCMOREKBAY.
With uniformly distributed regressors, the results of the analysis does not deviate much with the best estimators in normally distributed regressors except with addition of CORCMOREKLA which also compete favorably well. This is an indication that the above-named estimators can be used to address the problem of multicollinearity.

Moderate Level of Multicollinearity

With normally distributed regressors OREKBAY outperformed all other estimators. However the proposed modified ridge estimators like MLMOREKBAY, CORCMOREKBAY and MLMGRE also compete favorably well at different sample sizes.

With uniformly distributed regressors, while OREKBAY is still the best estimator CORCMOREKBAY and CORCMOREKLA compete favorably well among different sample sizes.

High Level of Multicollinearity

With normally distributed regressors the proposed modified ridge estimators of MLMOREKBAY and CORCMOREKBAY along with the existing estimator perform excellently well at all sample sizes.
With uniformly distributed regressors OREKBAY, MLMOREKBAY and CORCOREKBAY are the best estimators at different sample sizes. CORCMOREKLA also compete favorably well at sample size of n=50,100 and 250.

This study will best be illustrated with the aid of line graphas shown below. The levels of multicollinearity are categorize as Low (0.2 – 0.4), Moderate (0.6 – 0.8) and High (0.95 – 0.99) under Normal and Uniform Regressors.

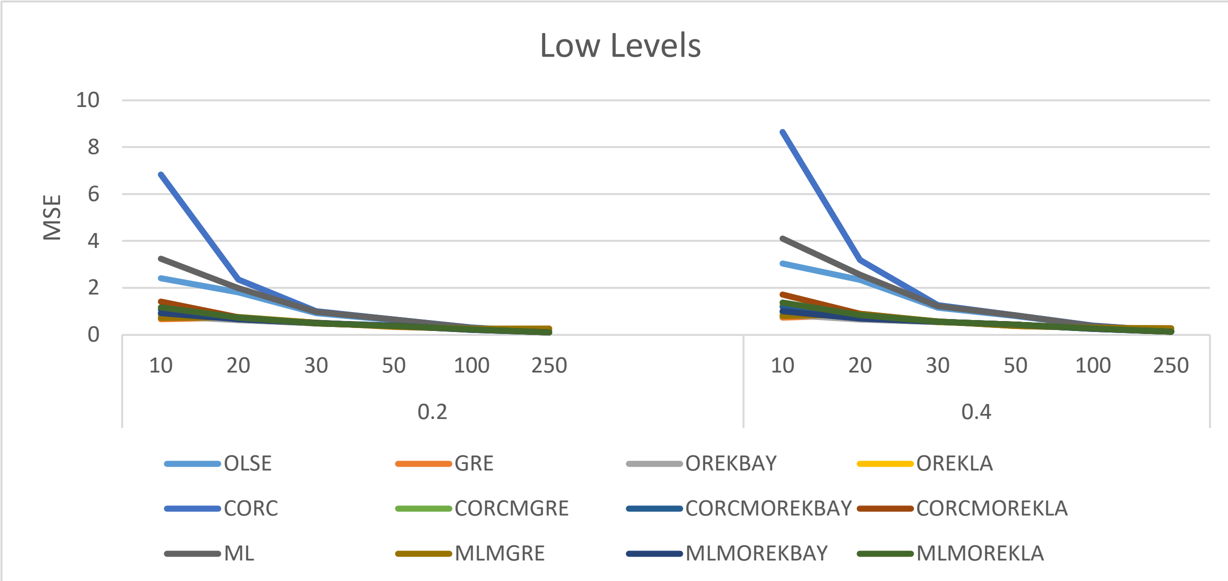

Figure 1: MSE at Low Levels of Multicollinearity with Normal Regressors

From Fig.1 above, the best estimator is OREKBAY follows by the proposed modified estimators of MLMOREKBAY and CORCOREKLA.

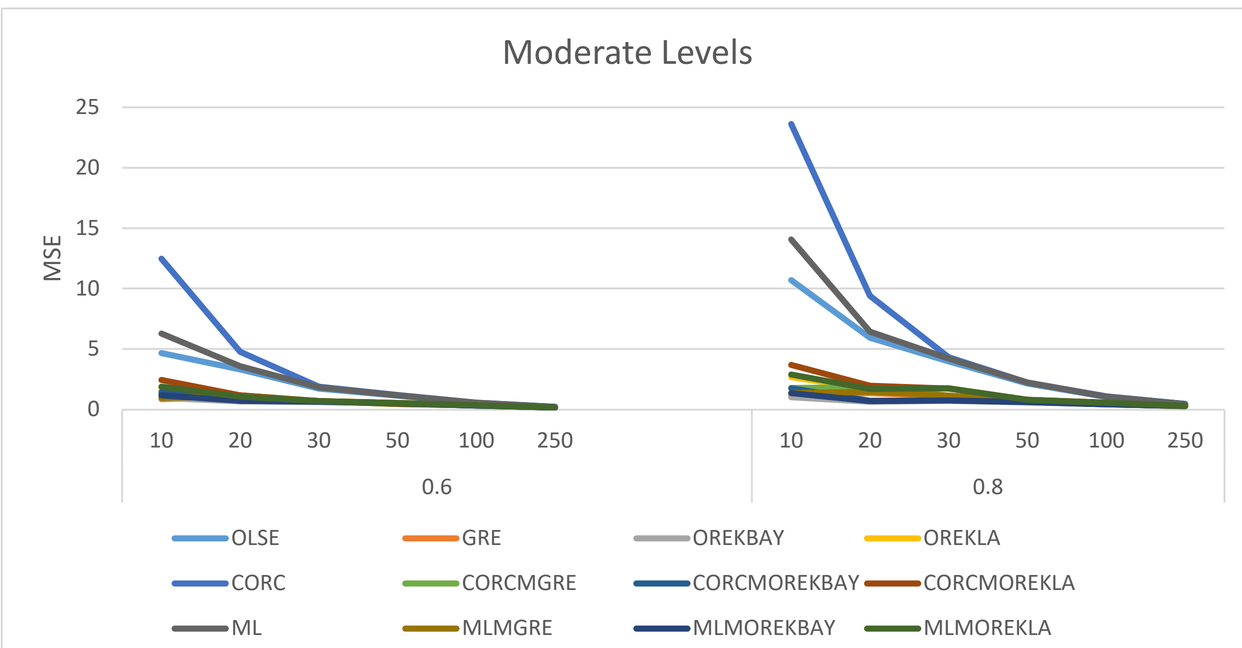

Figure 2: MSE at Moderate Levels of Multicollinearity with Normal Regressors

From Fig.2 above, the best estimator is the proposed modified ridge estimators of MLMOREKBAY and CORCMOREKBAY. However, the existing estimator OREKBAY also compete favorably well.

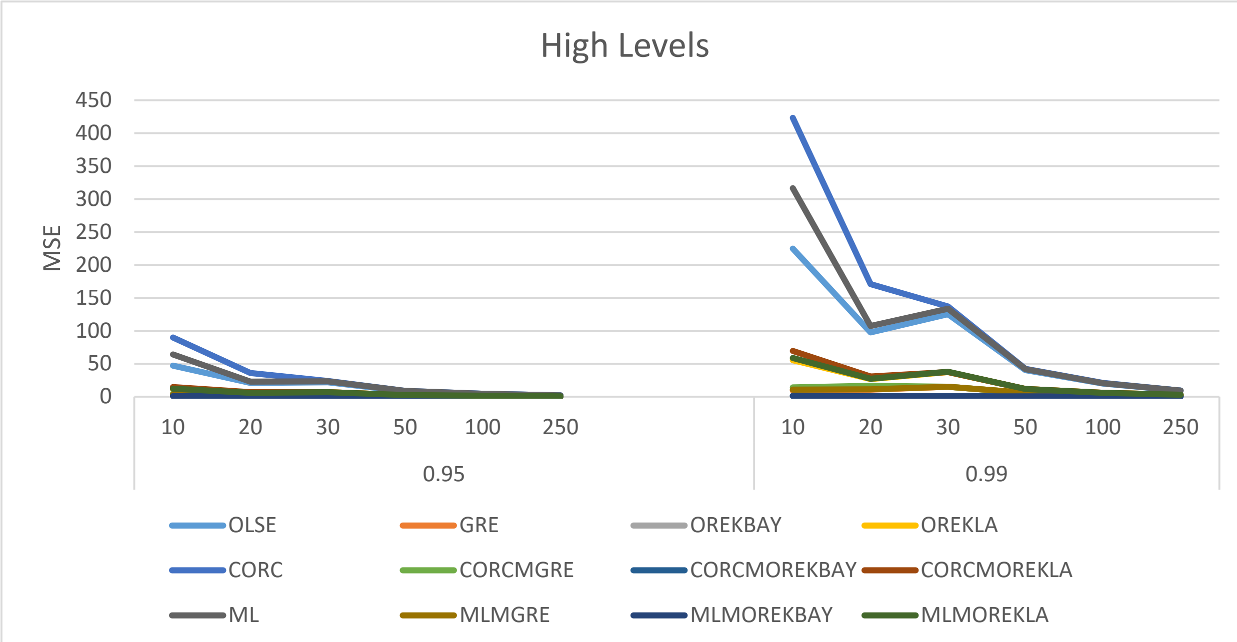

Figure 3: MSE at High Levels of Multicollinearity with Normal Regressors

From Fig.3 above, the best estimators are MLMOREKBAY, CORCMOREKBAY and OREKBAY.

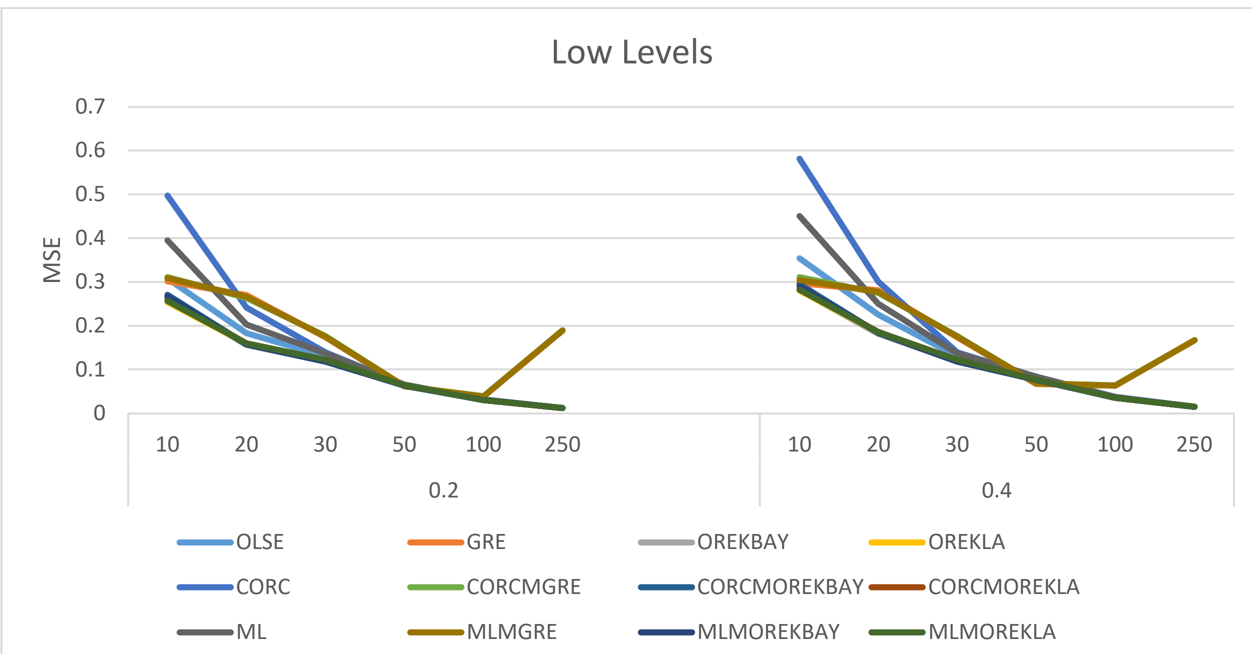

Figure 4: MSE at Low Levels of Multicollinearity with Uniform Regressors

The figure 4 above shows that the proposed estimators MLMOREKBAY perform excellently well follows by CORCMOREKBAY and CORCMOREKLA

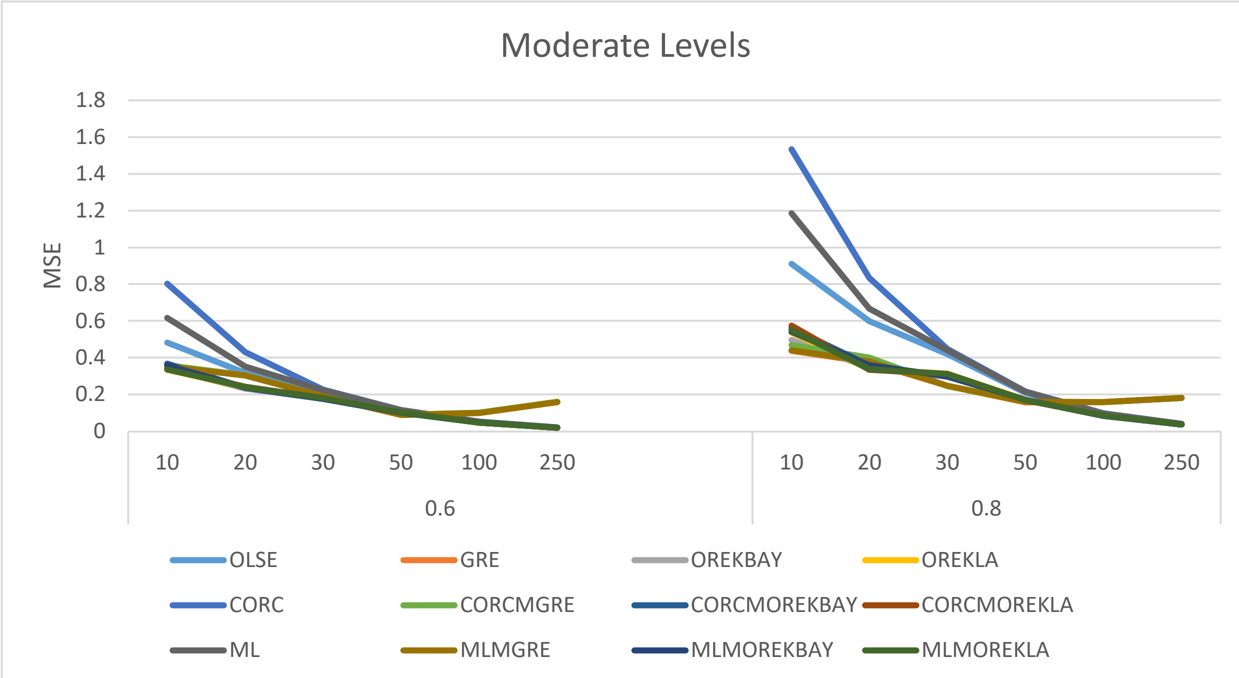

Figure 5: MSE at Moderate Levels of Multicollinearity with Uniform Regressors

From Fig.5 above, the best estimator is the proposed modified ridge estimators of CORCMOREKBAY. However, MLMOREKLA and CORCMOREKLA are best at sample size of n= 10,20,30,50, and 250.

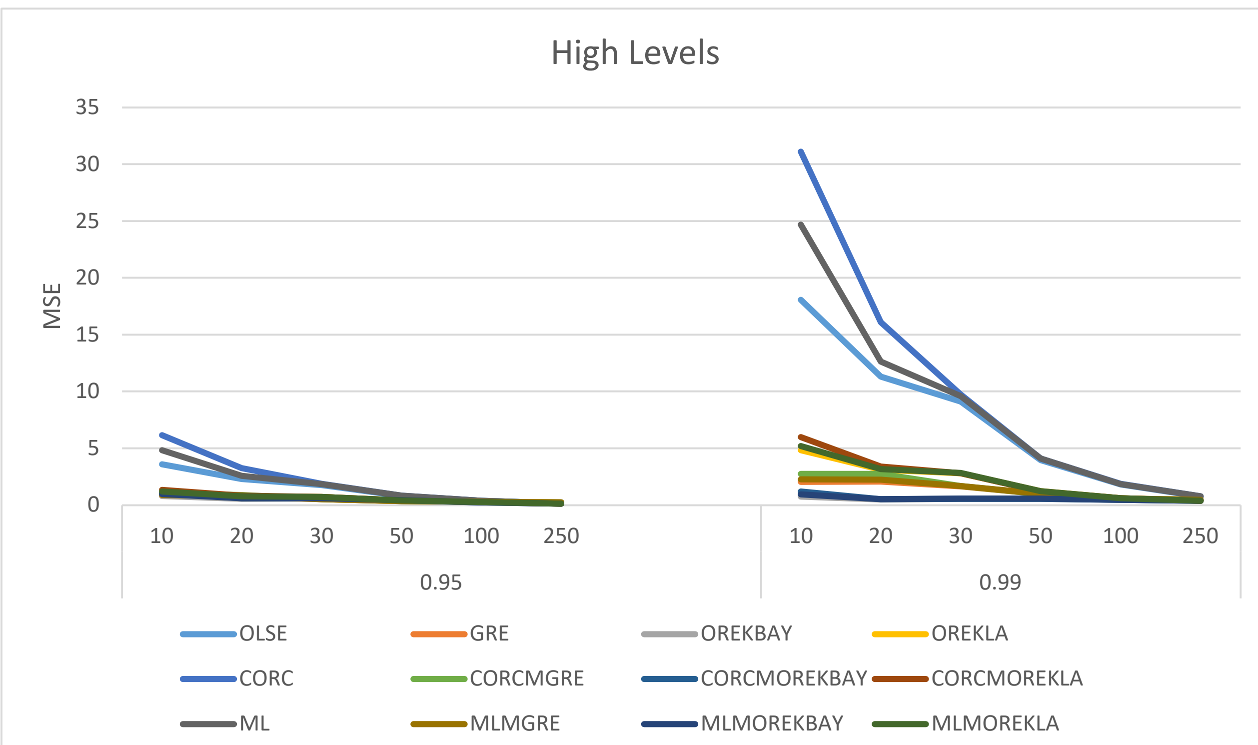

Figure 6: MSE at High Levels of Multicollinearity with Uniform Regressors

From Fig.6 above, the best estimator are GRE, OREKBAY, MLMOREKBAY and CORCMOREKBAY. However, at sample size of n=50,100 and 250 CORCMOREKLA do compete as the best.

Conclusions

A study on modified ridge estimators for solving multicollinearity in Linear Regression Model shows from the table that the best estimators among the ranks of the best five (5) estimators are categorize in to Low, Moderate and High Levels Multicollinearity. The study also revealed that among the proposed modified ridge estimator of MLMOREKBAY, CORCMOREKBAY, CORCMOREKLA and MLMGRE perform excellently well at different sample sizes and can therefore be used to solve problem of multicollinearity. Also, the existing estimator OREKBAY compete favorably well in addressing the problem of multicollinearity.

References

[1]. Ayinde, K. (2007a). Equations to generate normal variates with desiredinter - correlation matrix. International Journal of Statistics and System, 2 (2), 99 –111.

[2]. Ayinde, K. (2007b). A Comparative Study of the Performances of the OLS and Some GLS Estimators When Stochastic Regressors Are both Collinear and Correlated with Error Terms. Journal of Mathematics and Statistics, 3(4), 196 –200.

[3]. Bowerman, B. L. and O" Connell, R. T. (2006). Linear Statistical Models and Applied Approach, Boston. PWS-KENT Publishing. 4th Edition.

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

[5]. Cochrane, D. and Orcutt, G. H. (1949). Application of Least Square to relationship containing

[6]. Gujarati, D.N. and Porter, D.C. (2009). Basic Econometrics. Mc Graw-Hill, New York. 5th Edition.

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

[8]. Hoerl, A. E., Kennard, R. W. and Baldwin, K. F. (1975). Ridge Regression: Some Simulation. Journal of Communication in Statistics.

[9]. James, S. (1956). Inadmissibility of the usual estimator for the mean of a Multivariate Distribution. Proc. Third Berkelysymp.math.statist.prob.197-206

[10]. Kmenta, J. (1986). Elements of econometrics. Macmillan Publishing Co. New York.

[11]. Lukman, A. F. and Ayinde, K. (2017). Review and Classifications of the Ridge Parameter Estimation Techniques. Haccetteppe Journal of Mathematics and Statistics, 46(5),953-967

[12]. Markov, A. A. (1900). Wahrscheinlichkeitsrechnug.Leipzig:Tuebner

[13]. Massey, W. F. (1965). Principal Component Regression in exploratory statistical research. Journal of the American Statistical Association, 60, 234– 246.

[14]. Schumann, E. (2009). Generating Correlated Uniform Variate. http:// comisef.wikidot.com / tutorial:correlateduniformvariate.

[15]. 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

[16]. Sclove, S. (1973). Improved estimators for coefficient in Linear regression. Journal of American Statistical Association 63, 596-606

[17]. TSP (2005). User Guide and Reference Manual, Time series processor, New York.

[18]. 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.

10/11/2022