A MODIFIED BIASING RIDGE ESTIMATORS FOR ADDRESSING MULTICOLLINEARITY PROBLEM IN LINEAR REGRESSION MODEL

Adedoyin, M. A.[1], [2*], Oladapo, O. J.[3], Adejumo, A. O.[1]

[1]Department of Statistics. University of Ilorin
[2]Kwara State Bureau of Statistic, Ilorin.
[3]Department of Statistics. Ladoke Akintola University of Technology, Ogbomoso

ABSTRACT

This study conduct an extensive analysis of various biasing estimators in the context of multiple explanatory variable (p) and varying degree of multicollinearity, a number of approaches have been developed for deriving biasing estimators. In this study, a new approach to obtain the ridge biasing parameter k is suggested and then evaluated by Monte Carlo simulations. A number of different models are investigated for different number of observations, the strength of correlation between the explanatory variables, and distribution of the error terms. The mean squared error (MSE) criterion is used to examine the performance of the proposed estimators when compared with other well-known estimators. Accordingly, the analysis revealed that generally, mean square error (MSE) value of the estimators decrease as the degree of correlation among explanatory variables increased with a few exceptions. In conclusion, kibria's biasing estimator proposed in 2022 exhibit effectiveness across multicollinearity level, error terms, sample sizes and correlation levels. These results provide valuable insight for researcher and practitioner seeking to choose appropriate biasing estimators in similar statistical scenarios.

Keywords: Biased, Estimator, MSE, Multicollinearity, Parameter, Correlation

1 Introduction

Regression analysis is a statistical tool used for investigating the relationships between variables. Usually, the investigator seeks to ascertain the causal effect of one variable upon another. For example, the effects of a price increase upon demand and the effect of changes in the money supply upon the inflation rate. Also, one may wish to examine whether drinks' consumption is related to various socioeconomic and demographic variables such as age, education, income, and price of drinks. The relationship is expressed in the form of an equation or a model connecting the response or dependent variable with one or more explanatory or predictor variables. The response variable is denoted by Y and the set of predictor variables by $X_1, X_{2,} ..., X_p$, where ***p*** denotes the number of predictor variables. The true relationship between Y and $X_1, X_{2,} ..., X_p$ can be approximated by the regression model given as:

ISSN NUMBER: 1116-249X

$$Y = f(X_1, X_2, ..., X_p) + \varepsilon \tag{1}$$

where ε is assumed to be a random error representing the discrepancy in the approximation, it also accounts for the failure of the model to fit the data exactly. The function $f(X_1, X_2, ..., X_p)$ describes the relationship between Y and $X_1, X_{2,} ..., X_p$. An example is the linear regression model

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_p X_p + \varepsilon \tag{2}$$

Where, $\beta_0, \beta_1, \beta_2, ..., \beta_p$**,** are called the regression parameters or coefficients, which are unknown constants to be determined (estimated) from the data (Chatterjee and Hadi 2015).

The Ordinary Least Squares (OLS) estimator stands as the preferred method for estimating the parameters of regression model, provided certain assumptions are satisfied.

However, when the assumptions of independence among explanatory variables are violated, the problem of multicollinearity occurs (Owolabi *et al*, 2022). In the presence of multicollinearity, OLS estimator performance is unsatisfactory in that the standard errors of the regression coefficients are high leading to insufficient contribution of the regressors (Oladapo *et al*, 2022;Idowu *et al*, 2023). When these happens, the OLS is no more Best Linear Unbiased Estimator (BLUE). Several estimators have been proposed to handle the problem of multicollinearity, such as Stein estimator, Ridge estimator, Liu estimator, Principal Component and the likes.

When applying the biased estimators in handling multicolinearity, the biasing parameter has been studied to have effect on the value of the Mean Square Error (MSE) (Kibra, 2003; Oladapo *et al* 2023, 2024). In 2022 Kibra proposed some biasing parameter k but didn't subject it to either simulation or empirical studies. The objective of this paper is to subject some of the biasing

ISSN NUMBER: 1116-249X

parameter k propose by Kibra and also proposed some modified biasing parameter k using the concept of Kibra 2022 estimators to simulation studies. Performance of all the proposed biasing estimators is examined over several existing estimators with OLS using the mean square error (MSE) criterion through Monte Carlo Simulation.

2. METHODOLOGY

The regression model given as:

$$Y = X\beta + \varepsilon \tag{3}$$

Consider the canonical form of the regression model in (3) as:

$$Y = X^*\alpha + \varepsilon \tag{4}$$

Where $X^* = XD$, ($\alpha = D'\beta = (\alpha_{1,}\alpha_{2,} ..., \alpha_P)'$ and D is the orthogonal matrix given that D'D = I_P and $D'X'XD = \Delta$, where Λ = diag $\Delta = diag(\lambda_1, \lambda_2, ..., \lambda_P)$ contains the eigen values of $X'X$ matrix. Then, OLS (α_{OLS}) and Ridge estimator (α_R) in canonical form are given as:

$$\hat{\alpha}_{OLS} = \left(X^{*'}X^*\right)^{-1}\left(X^{*'}y\right) \tag{5}$$

$$\hat{\alpha}_R = \left(X^{*'}X^* + k\right)^{-1}\left(X^{*'}y\right), \tag{6}$$

where k = $diag(k_1, k_2, k_3, ..., k_p)$ and $k_i > 0$, for i = 1,2,3,…,p. The MSE of OLS and ridge regression estimation are given respectively as follows:

ISSN NUMBER: 1116-249X

$$MSE\left(\hat{\alpha}_{OLS}\right) = \sigma^2 \sum_{i=1}^{p} \frac{1}{\lambda_i} \quad (7)$$

$$MSE\left(\hat{\alpha}_{R}\right) = \sigma^2 \sum_{i=1}^{p} \frac{\lambda_i}{(\lambda_i + k)^2} + \sum_{i=1}^{p} \frac{k^2\alpha^2}{(\lambda_i + k)^2} \quad (8)$$

2.1 Some Ridge Estimators

In this section, we reviewed and proposed some ridge parameters.

Hoerl and Kennard (1970) suggest to replace σ^2 and α^2 by their corresponding unbiased estimator $\hat{\sigma}^2$ and, $\hat{\alpha}^2$ respectively. That is

$$k_i = \frac{\sigma^2}{\alpha_i^2}, \text{ i= 1,2,…,p} \quad (9)$$

Where, $\hat{\sigma}^2 = \frac{\sum \hat{e}_i^2}{n-p} = \frac{(y-\hat{y})'(y-\hat{y})}{n-p}$ is the residual mean square error, which is unbiased estimator of σ^2. Hoerl and Kennard (1970) suggest k to be

$$k_{HK1} = \hat{k}_{HK} = \frac{\hat{\sigma}^2}{\hat{\alpha}_{\max}^2} \quad (10)$$

Where $\hat{\alpha}_{\max}^2$ is the maximum element of α^2. If σ^2 and α^2 are known, then $\hat{k}_{Hk}$ will give smaller MSE than the OLS.

Kibria (2003) proposed some new estimator based on generalized ridge regression approach, by using the geometric mean of equation (9), which produces the following estimator as expressed below:

ISSN NUMBER: 1116-249X

$$\hat{k}_{GM} = \frac{\hat{\sigma}^2}{\left(\prod_{i=1}^{p} \hat{\alpha}_i^{\,2}\right)^{1/p}} \tag{11}$$

By using the median of equation (9), which produces the following estimator as expressed below:

$$k_{MED} = Median\left(\frac{\hat{\sigma}^2}{\hat{\alpha}^2{}_i}\right), i = 1,2,\ldots,p \tag{12}$$

Alkamisi and Shukur (2007), proposed six biasing estimators of k:

$$k_{20} = \frac{\hat{\sigma}^2}{\beta_{\max}^2} + \frac{1}{\lambda_{\max}} \tag{13}$$

$$k_{21} = \max\left(\frac{\hat{\sigma}^2}{\hat{\beta}_i^2} + \frac{1}{\lambda_i}\right) \tag{14}$$

$$k_{22} = \frac{1}{p}\sum_{i=1}^{p}\left(\frac{\hat{\sigma}^2}{\hat{\beta}_i^{\,2}} + \frac{1}{\lambda_i}\right) \tag{15}$$

$$k_{23} = Median\left(\frac{\hat{\sigma}^2}{\hat{\beta}_i^{\,2}} + \frac{1}{\lambda_i}\right) \tag{16}$$

$$k_{24} = \frac{p\hat{\sigma}^2}{\sum_{i=1}^{p}\beta_i^{\,2}} + \frac{1}{\lambda_{\max}} \tag{17}$$

$$k_{25} = \frac{p\hat{\sigma}^2}{\sum_{i=1}^{p}\lambda_i\beta_i^{\,2}} + \frac{1}{\lambda_{\max}} \tag{18}$$

Kibria (2022) also proposed the following estimators

$$k_{104GM} = GM\left(\frac{\hat{\sigma}^2}{\hat{\beta}^2} + \frac{1}{\lambda_i}\right) \tag{19}$$

Where GM stand for geometric mean

ISSN NUMBER: 1116-249X

$$k_{105\,HM} = HM\left(\frac{\hat{\sigma}^2}{\hat{\beta}^2} + \frac{1}{\lambda_i}\right) \tag{20}$$

Where HM stand for harmonic mean

$$k_{106\,MED} = Median\left(\frac{\hat{\sigma}^2}{\hat{\beta}^2} + \frac{1}{\lambda_i}\right) \tag{21}$$

$$k_{107\,MAX} = \max\left(\frac{\hat{\sigma}^2}{\hat{\beta}^2} + \frac{1}{\lambda_i}\right) \tag{22}$$

Where max stand for maximum

$$k_{108\,MIN} = \min\left(\frac{\hat{\sigma}^2}{\hat{\beta}^2} + \frac{1}{\lambda_i}\right) \tag{23}$$

Where min stands for minimum

2.2 The proposed Biasing Estimator.

Following the works of Alkhamisi and Shukur (2007) and the works of Kibria (2022), the following estimators are proposed:

$$k_{AOM\,1} = \frac{\hat{\sigma}^2}{\hat{\beta}^2_{\min}} + \frac{1}{\lambda_{mean}} \tag{24}$$

$$k_{AOM\,2} = \left(\frac{\hat{\sigma}^2}{\hat{\beta}^2_{\min}} + \frac{1}{\lambda_i}\right) \tag{25}$$

$$k_{AOM\,3} = \frac{\hat{\sigma}^2}{\hat{\beta}^{\,2}} + \frac{1}{\lambda_{\min}} \tag{26}$$

$$k_{AOM\,4} = \left(\frac{\hat{\sigma}^2}{\hat{\beta}_i^{\,2}} + \frac{1}{\sum_{i=1}^{p} \lambda_i}\right) \tag{27}$$

ISSN NUMBER: 1116-249X

$$k_{AOM5} = \frac{p\hat{\sigma}^2}{\sum_{i=1}^{p}\hat{\beta}_i^2} + \frac{1}{\lambda_{\min}} \tag{28}$$

$$k_{AOM6} = \frac{p\hat{\sigma}^2}{\sum_{i=1}^{p}\lambda_i\hat{\beta}_i^2} + \frac{1}{\lambda_{\min}} \tag{29}$$

2.3 Monte Carlo Simulation

In this section, the procedure of simulation study that has been conducted to compare the performance of the proposed biasing estimators with the OLS and Ridge Regression estimators are discussed.

2.3.1 Model Formulation for Simulation and Data Generation Procedure

Consider the linear regression model of the form

$$Y_t = \beta_0 X_{t1} + \beta_2 X_{t2} + ... + \beta_p X_{tp} + U_t \tag{30}$$

$t = 1, 2, ..., n; p = 3, 7$

where $U_t \sim$ NID $(0, \sigma^2 I_n)$

The model will be studied with fixed regressor, X_{ti}, $t = 1, 2, ..., n$; $i = 1, 2, ..., p$

2.3.2 Procedures for Generating the Explanatory Variables

The simulation procedure used by McDonald and Galarneau (1975), Wichern and Churchill (1978), Gibbons (1981), Kibria (2003), Lukman and Ayinde (2017), Owolabi *et al* (2022), Oladapo *et al* (2022) and Iyabo *et al* (2023), will be used to generate the explanatory variables in this study: This is given as:

$$x_{ij} = \left(1-\rho^2\right)^{1/2} z_{ij} + \rho z_{i,p+1},\ i = 1, 2, ..., m,\ j = 1, 2, ..., p. \tag{31}$$

ISSN NUMBER: 1116-249X

Where z_{ij} is an independent standard normal distribution with mean zero and unit variance, ρ is the correlation between any two explanatory variables and p is the number of explanatory variables. The values of ρ will be taken as 0.8, 0.9, 0.95, 0.99 and 0.999 respectively. Thus, the correlations between the variable are the same. In this study, the number of explanatory variables (p) will be seven (7).

2.4 Criteria for evaluation of the estimators

The performance of the estimators will be compared using estimated mean square error (EMSE). For any estimator $\hat{\beta}_{ij}$ EMSE is defined as follows:

$$EMSE(\hat{\beta}) = \frac{1}{5000}\sum_{J=1}^{5000}\left(\hat{\beta}_{ij} - \beta\right)'\left(\hat{\beta}_{ij} - \beta\right) \quad (32)$$

Where $\hat{\beta}_{ij}$ is i^{th} element of the estimator β in the j^{th} replication which gives the estimate of β_i . β_i are the true value of the parameter previously mentioned. Estimators with the minimum MSE are considered best.

3 Simulation Results

The performance of the proposed estimator is examined over several existing estimators. The result of Monte Carlo experiment concerning the MSE of different estimator compared to the OLS are presented

3.1 The performance of the estimators when number of parameters, p=2

The MSE values from the simulation studies for the design used are presented for different sample size n, ρ and σ^2 in Table 1 to Table 7.

ISSN NUMBER: 1116-249X

Table 1: Best three k-estimators when numbers of parameter, p = 2 and sample size, n = 10

σ^2	Multicollinearity level (ρ)				
	0.8	0.9	0.95	0.99	0.9999
3	k_{105HM} **(0.8221)**	k_{105HM} **(0.7211)**	k_{105HM} **(0.6404)**	k_{107MAX} **(0.4948)**	k_{AOM4} **(0.4221)**
	k_{107MAX}(1.2634)	k_{107MAX}(1.0699)	k_{107MAX} (0.8022)	k_{AM} (0.5123)	k_{104GM} (0.5403)
	k_{AM} (1.4401)	k_{106MED} (1.2236)	k_{106MED}(0.8971)	k_{106MED}(0.5123)	k_{AOM2}(0.5938)
5	k_{105HM}**(1.2335)**	k_{105HM}**(1.0606)**	k_{105HM}**(0.9881)**	k_{105HM} **(0.7353)**	k_{106MED}**(0.8221)**
	k_{107MAX}(2.4971)	k_{107MAX}(2.0862)	k_{107MAX}(1.5676)	k_{24} (0.8113)	k_{AM} (1.2633)
	k_{MAX}(2.7581)	k_{MAX}(2.4497)	k_{AM}(1.8372)	k_{107MAX} (0.8172)	k_{104GM} (1.4402)
10	k_{105HM}**(3.0804)**	k_{105HM}**(2.5431)**	k_{105HM}**(2.0391)**	k_{105HM} **(1.1520)**	k_{24} **(0.8864)**
	k_{107MAX}(8.1657)	k_{107MAX} (6.7091)	k_{107MAX}(4.992)	k_{107MAX} (2.2553)	k_{25}(0.8864)
	k_{MAX} (9.0082)	k_{MAX} (7.7426)	k_{AM}(6.0934)	k_{AM} (3.0098)	k_{107MAX} (0.8874)

MSE bold is the smallest

Table 2: Best three k-estimators when numbers of parameter, p = 2 and sample size, n = 20

σ^2	Multicollinearity level (ρ)				
	0.8	0.9	0.95	0.99	0.9999
3	k_{105HM} **(0.5469)**	k_{105HM} **(0.4651)**	k_{105HM} **(0.3995)**	k_{24} **(0.2623)**	k_{AOM2} **(0.2052)**
	k_{107MAX} (0.6478)	k_{107MAX}(0.5816)	k_{107MAX} (0.4953)	K25 0.2671	k_{AOM1} (0.2305)
	k_{AM} (0.6519)	k_{MAX}(0.5866)	k_{MAX} (0.5022)	k_{104GM} (0.2745)	k_{AM} (0.2439)
5	k_{105HM} **(0.7101)**	k_{105HM} **(0.6382)**	k_{105HM} **(0.5837)**	k_{105HM} **(0.5068)**	k_{AM} **(0.4045)**
	k_{107MAX} (0.9504)	k_{107MAX} (0.8441)	k_{107MAX} (0.7389)	k_{107MAX} (0.5599)	k_{106MED} (0.4045)
	k_{MAX} (0.9546)	k_{MAX} (0.8493)	k_{MAX} (0.7447)	k_{MAX} (0.5614)	k_{107MAX} (0.4111)
10	k_{105HM} **(1.1853)**	k_{105HM}**(1.1126)**	k_{105HM}**(1.0661)**	k_{105HM} **(0.9569)**	k_{105HM} **(0.7235)**
	k_{107MAX} (2.0848)	k_{107MAX}(1.8423)	k_{107MAX}(1.6574)	k_{107MAX} (1.4268)	k_{107MAX}(0.8116)
	k_{MAX} (2.0943)	k_{MAX} (1.8514)	k_{AM} (1.6668)	k_{AM} (1.4624)	k_{MAX} (1.4232)

MSE bold is the smallest

ISSN NUMBER: 1116-249X

Table 3: Best three k-estimators when numbers of parameter, p = 2 and sample size, n = 30

σ^2	Multicollinearity level (ρ)				
	0.8	0.9	0.95	0.99	0.9999
3	k_{105HM}**(0.4215)**	k_{105HM}**(0.3398)**	k_{105HM}**(0.2737)**	k_{105HM}**(0.1755)**	k_{25}**(0.0842)**
	k_{107MAX}(0.4876)	k_{107MAX}(0.4467)	k_{107MAX}(0.3784)	k_{107MAX}(0.2044)	k_{24}(0.0870)
	k_{AM} (0.4884)	k_{MAX}(0.4479)	k_{MAX}(0.3796)	k_{MAX} (0.2058)	k_{AOM2} (0.09493)
5	k_{105HM}**(0.5499)**	k_{105HM}**(0.4705)**	k_{105HM}**(0.4113)**	k_{105HM}**(0.3351)**	k_{25} **(0.1874)**
	k_{107MAX} (0.6668)	k_{107MAX}(0.5777)	k_{107MAX} (0.4833)	k_{107MAX}(0.3355)	k_{24} (0.1899)
	k_{MAX}(0.6674)	k_{MAX}(0.5784)	k_{MAX} (0.4839)	k_{MAX}(0.3362)	k_{108MIN}(02249)
10	k_{105HM} **(0.8165)**	k_{105HM}**(0.7477)**	k_{105HM}**(0.7050)**	k_{105HM}**(0.6584)**	k_{105HM} **(0.7235)**
	k_{107MAX} (.1773)	k_{107MAX}(1.0147)	k_{107MAX}(0.8947)	k_{107MAX}(0.7606)	k_{107MAX} (0.6824)
	k_{MAX}(1.1781)	k_{MAX}(1.0155)	k_{MAX} (0.8955)	k_{MAX}(0.7612)	k_{MAX} (0.7295)

MSE bold is the smallest

Table 4: Best three k-estimators when numbers of parameter, p = 2 and sample size, n = 40

σ^2	Multicollinearity level (ρ)				
	0.8	0.9	0.95	0.99	0.9999
3	k_{24} **(0.4431)**	k_{105HM} **(0.4654)**	k_{105HM} **(0.3848)**	k_{105HM} **(0.2421)**	k_{25} **(0.0955)**
	k_{HM}(0.4446)	k_{107MAX}(0.5192)	k_{107MAX} (0.4793)	k_{107MAX} (0.2964)	k_{24} (0.0987)
	k_{21} (0.4461)	k_{MAX} (0.5201)	k_{MAX}(0.4808)	k_{MAX} (0.2979)	k_{104GM} (0.1188)
5	k_{105HM} **(0.6823)**	k_{105HM} **(0.5968)**	k_{105HM} **(0.5237)**	k_{105HM} **(0.4071)**	k_{25} **(0.2448)**
	k_{107MAX} (0.7731)	k_{107MAX} (0.7213)	k_{107MAX}(0.6413)	k_{107MAX} (0.4383)	k_{24} (0.2467)
	k_{MAX} (0.7737)	k_{MAX} (0.7213)	k_{MAX} (0.6424)	k_{MAX} (0.4393)	k_{108MIN} (0.2676)
10	k_{105HM} **(0.9556)**	k_{105HM} **(0.8807)**	k_{105HM} **(0.8201)**	k_{105HM} **(0.7423)**	k_{105HM} **(0.6563)**
	k_{107MAX} (1.4522)	k_{107MAX} (1.3221)	k_{107MAX} (1.1716)	k_{107MAX} (0.9212)	k_{107MAX} (0.8153)
	k_{MAX} (1.4534)	k_{MAX} (1.3234)	k_{MAX} (1.1730)	k_{MAX}(0.9226)	k_{MAX} (0.8335)

MSE bold is the smallest

ISSN NUMBER: 1116-249X

Table 5: Best three k-estimators when numbers of parameter, p = 2 and sample size, n = 50

σ^2	Multicollinearity level (ρ)				
	0.8	0.9	0.95	0.99	0.9999
3	k_{21} **(0.3956)**	k_{105HM} **(0.4755)**	k_{105HM} **(0.4136)**	k_{105HM} **(0.2839)**	k_{25} **(0.1154)**
	k_{24} (0.3991)	k_{107MAX} (0.5257)	k_{107MAX} (0.5068)	k_{107MAX} (0.3451)	k_{24} (0.1196)
	k_{108MIN} (0.4317)	k_{MAX} (0.5266)	k_{MAX} (0.5084)	k_{MAX} (0.3475)	k_{104GM} (0.1447)
5	k_{105HM} **(0.6541)**	k_{105HM} **(0.6058)**	k_{105HM} **(0.5545)**	k_{105HM} **(0.4541)**	k_{25} **(0.2585)**
	k_{107MAX} (0.7372)	k_{107MAX} (0.7329)	k_{107MAX} (0.6887)	k_{107MAX} (0.5085)	k_{24} (0.2696)
	k_{MAX} (0.7376)	k_{MAX} (0.7337)	k_{MAX} (0.6897)	k_{MAX} (0.5097)	k_{108MIN}(0.3245)
10	k_{105HM} **(0.9042)**	k_{105HM}**(0.8839)**	k_{105HM} **(0.8574)**	k_{105HM}**(0.8063)**	k_{105HM} **(0.7111)**
	k_{107MAX} (1.3415)	k_{107MAX} (1.3353)	k_{107MAX} (1.2657)	k_{107MAX} (0.0549)	k_{107MAX}(0.9362)
	k_{MAX} (1.3423)	k_{MAX} (1.3364)	k_{MAX} (1.2672)	k_{MAX} (1.0567)	k_{MAX} (0.9622)

MSE bold is the smallest

Table 6: Best three k-estimators when numbers of parameter, p = 2 and sample size, n = 60

σ^2	Multicollinearity level (ρ)				
	0.8	0.9	0.95	0.99	0.9999
3	k_{21} **(0.3144)**	k_{105HM}**(0.4082)**	k_{105HM} **(0.3404)**	k_{105HM} **(0.2121)**	K_{25} **(0.0829)**
	K_{24} (0.3297)	K_{107MAX} (0.4579)	K_{107MAX} (0.4336)	K_{107MAX} (0.2704)	K_{24} (0.08553)
	K_{108MIN} (0.3554)	K_{MAX}(0.4584)	K_{MAX}(0.4345)	K_{MAX} (0.7122)	K_{104GM} (0.0991)
5	K_{105HM} **(0.5956)**	K_{105HM} **(0.5257)**	K_{105HM} **(0.4631)**	K_{105HM} **(0.3609)**	K_{25} **(0.2191)**
	K_{107MAX}(0.6461)	K_{107MAX} (0.6140)	K_{107MAX} (0.5528)	K_{107MAX}(0.3771)	K_{24} (0.2225)
	K_{MAX} (0.6463)	K_{MAX} (0.6144)	K_{MAX} (0.5533)	K_{MAX} (0.3776)	K_{108MIN} (0.2302)
10	K_{105HM} **(0.8026)**	K_{105HM} **(0.7578)**	K_{105HM} **(0.7214)**	K_{105HM} **(0.6688)**	K_{105HM} **(0.6196)**
	K_{107MAX} (1.0816)	K_{107MAX} (1.0278)	K_{107MAX} (0.9467)	K_{107MAX}(0.7772)	K_{107MAX} (0.7095)
	K_{MAX} (1.0819)	K_{MAX} (1.0283)	k_{MAX} (0.9472)	k_{MAX} (0.7778)	k_{MAX} (0.7148)

MSE bold is the smallest

ISSN NUMBER: 1116-249X

Table 7: Best three k-estimators when numbers of parameter, p = 2 and sample size, n = 70

σ^2	Multicollinearity level (ρ)				
	0.8	0.9	0.95	0.99	0.9999
3	k_{21} **(0.2933)**	k_{105HM}**(0.4005)**	k_{105HM} **(0.3293)**	k_{105HM} **(0.1946)**	k_{25} **(0.0772)**
	k_{24} (0.3139)	k_{107MAX} (0.4098)	k_{107MAX} (0.4312)	k_{107MAX} (0.2704)	k_{24} (0.07888)
	k_{108MIN} (0.3145)	k_{MAX} (0.4129)	k_{MAX} (0.4350)	k_{MAX} (0.2712)	k_{104GM} (0.8741)
5	k_{105HM} **(0.5923)**	k_{105HM} **(0.5143)**	k_{105HM} **(0.4435)**	k_{105HM} **(0.3230)**	k_{25} (0.2110)
	k_{107MAX} (0.6494)	k_{107MAX} (0.6186)	k_{107MAX} (0.5526)	k_{107MAX} (0.3514)	k_{24} (0.2085)
	k_{MAX} (0.6496)	k_{MAX} (0.6190)	k_{MAX} (0.5531)	k_{MAX} (0.3518)	k_{108MIN} (0.2152)
10	k_{105HM} **(0.8084)**	k_{105HM} **(0.7465)**	k_{105HM} **(0.6909)**	k_{105HM} **(0.6107)**	k_{105HM} **(0.5556)**
	k_{107MAX} (1.1239)	k_{107MAX} (1.0449)	k_{107MAX} (0.9289)	k_{107MAX} (0.7075)	k_{107MAX} (0.6184)
	k_{MAX} (1.1242)	k_{MAX} (1.0453)	k_{MAX} (0.9293)	k_{MAX} (0.7079)	k_{MAX} (0.6198)

MSE bold is the smallest

3.2 The performance of the estimators when number of parameters, p=7

The MSE values from the simulation studies for the design used are presented for different sample size n, ρ and σ^2 in Table 8 to Table 10.

Table 8: Best three k-estimators when numbers of parameter, p = 7 and sample size, n = 10

σ^2	Multicollinearity level (ρ)				
	0.8	0.9	0.95	0.99	0.999
3	k_{105HM} **(0.7999)**	k_{107MAX} **(0.7151)**	k_{107MAX} **(0.4572)**	k_{24} **(0.4948)**	k_{AOM2} **(0.1009)**
	k_{107MAX} (0.8602)	k_{105HM} (0.7261)	k_{24} (0.5938)	k_{25} (0.1161)	k_{21} (0.1028)
	k_{AM} (1.0305)	k_{MAX} (1.2275)	k_{25}(0.6073)	k_{AM} (0.5542)	k_{106MED} (0.1571)
5	k_{105HM} **(0.8256)**	k_{105HM} **(0.7624)**	k_{107MAX} **(0.6882)**	k_{24} **(0.2078)**	k_{AOM4} **(0.1917)**
	k_{107MAX}(1.3095)	k_{107MAX}(1.0740)	k_{105HM} (0.7267)	k_{25} (0.2086)	k_{104GM} (0.2583)
	k_{MAX} (1.6057)	k_{MAX} (1.7898)	k_{MAX} (1.6064)	k_{107MAX} (0.3012)	k_{AOM2} (0.2691)
10	k_{105HM} **(0.8735)**	k_{105HM} **(0.8207)**	k_{105HM} **(0.7883)**	k_{107MAX} **(0.6421)**	k_{AM} **(0.4962)**
	k_{107MAX} (2.5321)	k_{107MAX} (1.9678)	k_{107MAX} (1.2933)	k_{25} (0.6391)	k_{104GM} (0.5094)
	k_{MAX} (3.1089)	k_{MAX} (3.0739)	k_{AM} (2.9373)	k_{24} (0.6437)	k_{MED} (0.5425)

MSE bold is the smallest

ISSN NUMBER: 1116-249X

Table 9: Best three k-estimators when numbers of parameter, p = 7 and sample size, n = 20

σ^2	Multicollinearity level (ρ)				
	0.8	0.9	0.95	0.99	0.999
3	k_{107MAX}**(0.6754)**	k_{107MAX} **(0.6172)**	k_{107MAX} **(0.5762)**	k_{24} **(0.2174)**	k_{AM} **(0.1086)**
	k_{MAX} (0.6767)	k_{MAX} (0.6233)	k_{MAX} (0.5989)	k_{25} (0.2177)	k_{108MIN} (0.1284)
	k_{AM} (0.7457)	k_{105HM} (0.8288)	k_{10HM} (0.7454)	k_{107MAX} (0.2777)	k_{AOM2} (0.1320)
5	k_{MAX} **(0.8076)**	k_{107MAX} **(0.7607)**	k_{107MAX} **(0.6948)**	k_{107MAX} **(0.4013)**	k_{104GM} **(0.1665)**
	k_{105HM} (0.9122)	k_{MAX} (0.7660)	k_{MAX} (0.7058)	k_{MAX} (0.4804)	k_{106MED} (0.1672)
	k_{MAX} (1.6057)	k_{105HM} (0.8495)	k_{105HM} (0.8313)	k_{25} (0.5761)	k_{AOM4}(0.1755)
10	k_{105HM} **(0.9254)**	k_{105HM} **(0.8788)**	k_{105HM} **(0.7883)**	k_{107MAX} **(0.7289)**	k_{24} **(0.4137)**
	k_{107MAX} (1.1336)	k_{107MAX} (1.0857)	k_{107MAX} (0.9993)	k_{105HM} (0.7657)	k_{25} (0.4143)
	k_{MAX} (1.1364)	k_{MAX}(1.0901)	k_{MAX} (1.0085)	k_{MAX} (0.7711)	k_{107MAX} (0.4665)

MSE bold is the smallest

Table 10: Best three k-estimators when numbers of parameter, p = 7 and sample size, n = 30

σ^2	Multicollinearity level (ρ)				
	0.8	0.9	0.95	0.99	0.999
3	k_{AM} **(0.6119)**	k_{107MAX} **(0.5740)**	k_{107MAX} **(0.5341)**	k_{107MAX}**(0.35810**	k_{MIN} **(0.1116)**
	k_{MEAN}(0.6131)	k_{MAX} (0.5754)	k_{MAX} (0.5397)	k_{MAX} (0.3986)	k_{107MAX}(0.1142)
	k_{MAX} (0.6503)	k_{AM} (0.6909)	k_{10HM} (0.7454)	k_{105HM}(0.5553)	k_{MAX} (0.1431)
5	k_{107MAX} **(0.7431)**	k_{107MAX} **(0.6861)**	k_{24} **(0.0709)**	k_{AOM4} **(0.0243)**	k_{21} **(0.0097)**
	k_{MAX} (0.7444)	k_{MAX} (0.6884)	k_{25} (0.0757)	k_{104GM} (0.0314)	k_{AOM2}(0.0153)
	k_{105HM} (0.9164)	k_{105HM} (0.8516)	k_{AM} (0.8516)	k_{106MED} (0.0345)	k_{AOM4} (0.0501)
10	k_{105HM} **(0.9291)**	k_{105HM} **(0.8810)**	k_{107MAX} (0.8284)	k_{107MAX} (0.6114)	k_{25}**(0.3588)**
	k_{107MAX} (0.9646)	k_{107MAX} (0.91323)	k_{105HM} (0.8285)	k_{MAX} (0.6157)	k_{24}(0.4318)
	k_{MAX} (0.9645)	k_{MAX} (0.9145)	k_{MAX}(0.8302)	k_{105HM} (0.7272)	k_{AM} (0.4318)

MSE bold is the smallest

ISSN NUMBER: 1116-249X

Table 11: Best three k-estimators when numbers of parameter, p = 7 and sample size, n = 40

σ^2	Multicollinearity level (ρ)				
	0.8	0.9	0.95	0.99	0.999
3	k_{AM} **(0.5151)**	k_{107MAX} **(0.5115)**	k_{107MAX}**(0.4761)**	k_{107MAX}**(0.3729)**	k_{24} **(0.0347)**
	k_{MEAN}(0.5157)	k_{MAX} (0.5122)	k_{MAX} (0.4794)	k_{MAX} (0.3907)	k_{25}(0.0347)
	k_{107MAX} (0.5914)	k_{AM} (0.5742)	k_{10HM} (0.7013)	k_{105HM}(0.4688)	k_{104GM} (0.0518)
5	k_{107MAX} **(0.6721)**	k_{107MAX}**(0.6192)**	k_{107MAX} **(0.5752)**	k_{107MAX} **(0.3692)**	k_{24}**(0.0868)**
	k_{MAX} (0.6723)	k_{MAX} (0.6201)	k_{MAX} (0.5722)	k_{MAX} (0.3738)	k_{25} (0.0871)
	k_{AM} (0.8167)	k_{105HM} (0.8185)	k_{105HM}(0.7279)	k_{105HM} (0.5296)	k_{104GM} (0.1221)
10	k_{107MAX} **(0.8610)**	k_{107MAX} (0.8100)	k_{107MAX}**(0.7256)**	k_{107MAX}**(0.5058)**	k_{25} **(0.3289)**
	k_{MAX} (0.8613)	k_{MAX} (0.8106)	k_{MAX}(0.7263)	k_{MAX} (0.5072)	k_{24} (0.3323)
	k_{105HM} (0.9115)	k_{105HM} (0.7892)	k_{105HM}(0.07892)	k_{105HM}(0.6622)	k_{107MAX} (0.3599)

MSE bold is the smallest

Table 12: Best three k-estimators when numbers of parameter, p = 7 and sample size, n = 50

σ^2	Multicollinearity level (ρ)				
	0.8	0.9	0.95	0.99	0.999
3	k_{104GM} **(0.5560)**	k_{AM} **(0.5734)**	k_{107MAX} **(0.5288)**	k_{107MAX} **(0.4095)**	k_{24}**(0.0476)**
	k_{AM} (0.5655)	k_{MEAN} (0.5757)	k_{MAX} (0.5304)	k_{MAX} (0.4234)	k_{25}(0.0476)
	k_{GM}(0.5591)	k_{107MAX} (0.5836)	k_{AM} (0.6959)	k_{105HM} (0.5592)	k_{104GM} (0.0729)
5	k_{107MAX} **(0.7425)**	k_{107MAX} **(0.6771)**	k_{107MAX}**(0.6248)**	k_{107MAX} **(0.4291)**	k_{25} **(0.1221)**
	k_{MAX} (0.7426)	k_{MAX} (0.6775)	k_{MAX} (0.6262)	k_{MAX} (0.4332)	k_{24}(0.1239)
	k_{AM} (0.8104)	k_{105HM}(0.8823)	k_{105HM} (0.8079)	k_{105HM} (0.6162)	k_{104GM} (0.1695)
10	k_{107MAX} **(0.9214)**	k_{107MAX} **(0.8761)**	k_{107MAX} **(0.8102)**	k_{107MAX}**(0.59302)**	k_{25} **(0.4302)**
	k_{MAX} (0.9216)	k_{MAX} (0.8765)	k_{MAX} (0.8113)	k_{MAX} (0.5946)	k_{24}(0.4375)
	k_{105HM} (0.9455)	k_{105HM} (0.9015)	k_{105HM}(0.8478)	k_{105HM}(0.7227)	k_{107MAX}(0.471)

MSE bold is the smallest

ISSN NUMBER: 1116-249X

Table 13: Best three k-estimators when numbers of parameter, p = 7 and sample size, n = 60

σ^2	Multicollinearity level (ρ)				
	0.8	0.9	0.95	0.99	0.999
3	k_{104GM}**(0.4635)**	k_{AM} **(0.4901)**	k_{107MAX}**(0.4632)**	k_{107MAX}**(0.3791)**	k_{24}**(0.0437)**
	k_{GM} (0.4653)	k_{MEAN} (0.4240)	k_{MAX} (0.4645)	k_{MAX} (0.3910)	k_{25}(0.04366)
	k_{AM} (0.4913)	k_{MAX} (0.5223)	k_{AM}(0.6603)	k_{105HM} (0.4907)	k_{104GM} (0.0802)
5	k_{107MAX}**(0.6852)**	k_{107MAX} **(0.6118)**	k_{107MAX}**(0.5638)**	k_{107MAX}**(0.3858)**	k_{25}**(0.1455)**
	k_{MAX} (0.6852)	k_{MAX} (0.6122)	k_{MAX} (0.5649)	k_{MAX} (0.3886)	k_{24}(0.1573)
	k_{AM} (0.70850)	k_{AM} (0.8032)	k_{105HM} (0.7635)	k_{105HM} (0.5415)	K104GM 0.1708
10	k_{107MAX} **(0.8468)**	k_{107MAX} **(0.7924)**	k_{107MAX} **(0.7194)**	k_{107MAX}**(0.4975)**	k_{25} **(0.3410)**
	k_{MAX} (0.8469)	k_{MAX} (0.7927)	k_{MAX} (0.7191)	k_{MAX} (0.4984)	k_{24}(0.3431)
	k_{105HM} (0.9306)	k_{105HM} (0.8759)	k_{105HM} (0.8101)	k105HM (0.6608)	k_{107MAX} (0.2558)

MSE bold is the smallest

Table 14: Number of times the k-estimators produced minimum MSE.

	P=2							P=7				
Est	**0.8**	**0.8**	0	**0.95**	**0.99**	**0.999**	**Total**	**0.9**	**0.95**	**0.99**	**0.999**	**Total**
OLS	0	0	0	0	0	0	0	0	0	0	0	0
K_{HK}	0	0	0	0	0	0	0	0	0	0	0	0
K_{GM}	0	0	0	0	0	0	0	0	0	0	0	0
K_{MED}	0	0	0	0	0	0	0	0	0	0	1	1
K_{MAX}	0	0	0	0	0	0	0	0	0	0	0	0
K_{MIN}	0	0	0	0	0	0	0	0	0	0	0	0
K_{HM}	0	0	0	0	0	0	0	0	0	0	0	0
K_{MEAN}	0	0	0	0	0	0	0	0	0	0	0	0
K_{21}	3	0	2	0	2	1	3	0	0	0	0	3
K_{AM}	0	2	0	0	0	2	6	0	0	0	1	1
K_{24}	1	0	0	0	2	3	5	0	0	0	2	3

ISSN NUMBER: 1116-249X

K_{25}	0	0	0	0	0	5	5	0	0	1	8	9
K_{104GM}	0	0	3	0	0	1	3	0	0	0	0	2
K_{105HM}	**17**	5	0	2	0	0	8	**21**	**21**	**20**	**6**	**85**
K_{106MED}	0	0	**13**	0	0	0	0	0	0	0	1	1
K_{107MAX}	0	**9**	0	**15**	**13**	**3**	**53**	0	0	0	0	0
K_{108MIN}	0	0	0	0	0	1	1	0	0	0	0	0
K_{AOM1}	0	0	0	0	1	0	1	0	0	0	0	0
K_{AOM2}	0	0	0	0	0	1	1	0	0	0	1	1
K_{AOM3}	0	0	0	0	0	0	0	0	0	0	0	0
K_{AOM4}	0	0	0	0	0	1	1	0	0	0	1	1
K_{AOM5}	0	0	0	0	0	0	0	0	0	0	0	0
K_{AOM6}	0	0	0	0	0	0	0	0	0	0	0	0

NOTE: the most frequency efficient estimator is bolded over the levels of multicollinearity.

4. Discussion

In this study, observations were made based on different scenarios involving the number of explanatory variables (p), levels of multicollinearity, error terms values and samples sizes (n). The k Harmonic mean (HM) version of the biasing estimator proposed by kibria in 2022 consistently demonstrated superiority across all levels or degree of multicollinearity compared to other estimator. The superiority was evident when error terms were set at values of 5 and 10 for all sample sizes, except at a degree of multicollinearity of 0.999. The k Maximum version of Kibria's biasing estimator outperformed all other estimators at all levels of multicollinearity, with the only exception being when the degree of multicollinearity was 0.999

5. Conclusion and Recommendation

This study evaluated the performance of a newly proposed estimator through extensive simulation experiments under varying conditions of multicollinearity levels, error structures, and

ISSN NUMBER: 1116-249X

sample sizes. The results revealed that the proposed biasing estimators did not meet expectations, often underperforming across the tested scenarios. In contrast, Kibria's (2022) biasing estimator consistently demonstrated superior performance, exhibiting the lowest mean squared error in nearly all cases. Based on these findings, the study recommends the use of Kibria's (2022) biasing estimators for sample sizes up to 100. It is particularly effective in settings with a large number of explanatory variables and remains a viable choice even when only two explanatory variables are involved, provided the sample size does not exceed 100.

References

Alkhamisi, M.A and G. Shukur (2007). A Monte Carlo Study of Recent Ridge Parameters, Communications in Statistics—Simulation and Computation, 36:3, 535-547.

Gibbons, D.G (1981). "A simulation study of some ridge estimators," Journal of the American Statistical Association, vol. 76, no. 373, p. 131–139.

Hoerl, A. E., Kennard, R. W. (1970). Ridge regression: biased estimation for nonorthogonal Problems. Technometrics 12:55–67

Idowu, J. I., Oladapo, O. J., Owolabi, A. T. and Ayinde, K.,(2022) "On the biased Two Parameter Estimator to Combat Multicollinearity in Linear Regression Model," African Scientific Reports, vol. 1, p. 188–204,

Idowu, J. I., Oladapo, O. J**.,** Owolabi, A.T., Ayinde, K. and Akinmoju, O. (2023), Combating multicollinearity: A new two-parameter approach, *Nicel Bilimler Dergisi*, 5(2), 90-116. doi: 10.51541/nicel.1084768

Idowu, J. I., Owolabi, A.T., Oladapo, O. J., Ayinde, K. and Oshuoporu, O. A., Alao, A.N. (2023). Mitigating Multicollinearity in Linear Regression Model with Two Parameter Kibria-Lukman Estimators. *WSEAS Transactions on Systems and Control*, 18: 612 -635. DOI: 10.37394/23203.2023.18.63

Kibria, B.M (2003). Performance of some New Ridge Regression estimators. Communications in Statistics part B: Simulation and Computation,32(2),419-435. https://doi.org/10.1080/SAC-120017499

Kibria, B. M. G. (2022). More than hundred (100) estimators for estimating the shrinkage parameter in a linear and generalized linear ridge regression models. *Journal of* Econometrics and Statistics 2(2):233-252.

Lukman, A. F., Ayinde, K., & Ajiboye, A. S. (2017). Monte Carlo study of some classification-based ridge parameter estimators. *Journal of Modern Applied Statistical Methods*, 16(1), 428–451. https://doi.org/10.22237/jmasm/1493598240

McDonald, G. C. and Galarneau, D. I., (1975)"A Monte Carlo evaluation of some ridge-type estimators," Journal of the American Statistical Association, vol. 70, pp. 407-416, 1975.

Oladapo, O.J., Owolabi, A.T., Idowu, J.I. and Ayinde, K. (2022), "A New Modified Liu Ridge-Type Estimator for the Linear Regression Model: Simulation and Application," Int J Clin Biostat Biom, vol. 8, no. 048.

Oladapo, O.J**.,** Idowu, J.I., Owolabi, A.T., and Ayinde, K. (2023). A New Biased Two-Parameter Estimator in Linear Regression Model. *EQUATIONS.* 3: 73-92. http://doi.org/10.37394/232021.2023.3.10.

Oladapo, O.J, Alabi O.O, Ayinde, K. (2023). Performance of Some Dawoud-Kibra Estimators for Logistic Regression Model: Application to Pena data set. *EQUATIONS*. 3:130-13., DOI:10.37394/232021.2023.3.16

Oladapo, O.J, Alabi O.O, Ayinde, K. (2024). Another New Two Parameter Estimator in Dealing with Multicollinearity in the Logistic Regression Model. *International Journal of Mathematical Sciences and Optimization: Theory and Applications*. 10(2):22–35. https://doi.org/10.5281/zenodo.10937145

Owolabi, A. T., Ayinde, K., Idowu, J. I., Oladapo, O. J. and Lukman, A. F.,(2022) "A New Two-Parameter Estimator in the Linear Regression Model with Correlated Regressors," Journal of Statistics Applications & Probability, vol. 11, no. 2, pp. 185-201.

Wichern, D. and Churchill, G.,(1978) "A Comparison of Ridge Estimators," Technometrics, vol. 20, p. 301–311.

ISSN NUMBER: 1116-249X