Journal of NAMP 70, (2025) 125-142

The Nigerian Association of Mathematical Physics

Journal homepage: https://nampjournals.org.ng

ON THE PERFORMANCE OF NEW WEIGHTED RIDGE ESTIMATORS FOR MITIGATING HETEROSCEDASTICITY PROBLEM ON A LINEAR REGRESSION MODEL WITH APPLICATION

***Amalare, A. A., Ayinde, K., Ekum, M. I., Ogunsanya, A. S. and Agbona, A. A.**

[1,3]*Department of Statistics, Lagos State University of Science and Technology, Ikorodu, Lagos State, Nigeria.*

[2]*Department of Mathematics and Statistics, Northwest Missouri State University, Maryville, Missouri, USA.*

[4]*Department of Statistics, University of Ilorin, Kwara State, Nigeria.*

[5]*Department of Statistics, Federal Polytechnic, Ede, Osun state Nigeria..*

ARTICLE INFO

Article history:
Received 22/5/2025
Revised 8/7/2025
Accepted 10/7/2025
Available online 17/7/2025

Keywords:
Error variances, Heteroscedasticity, Linear regression model, Monte Carlo simulation, New weighted ridge estimators.

ABSTRACT

Heteroscedasticity, a common violation of the homoscedasticity assumption in classical linear regression models, adversely affects parameter estimation and predictive accuracy, is a critical issue to be addressed. This study developed and evaluated new weighted ridge estimators for mitigating the effects of heteroscedasticity in linear regression models, particularly where there is little or no multicollinearity. The proposed estimators were derived by combining ridge regression and weighted least squares techniques. Monte Carlo simulations at varying levels of heteroscedasticity, error variances, and sample sizes, were used to assess their performance with the mean square error (MSE) showing that the proposed ORERW estimator performed best when the nature of heteroscedasticity is known, while OREW2 was superior when heteroscedasticity is unknown. Real-life data on passenger car mileage was used to validate the estimators with GREW3 and OREW3 outperforming the traditional methods.

1. INTRODUCTION

The classical linear regression model is widely recognized as a foundational tool for predictive analysis. Its appeal lies in its simplicity, interpretability, and the extensive range of computational tools available for validating theoretical assumptions and analyzing diverse data types. In essence, linear regression predicts the behaviour of a response variable based on its linear relationship with one or more predictor variables [1]. This capability has made it a staple in data analysis across numerous disciplines, including but not limited to the natural sciences, engineering, medicine, and social sciences by examining the relationship between a dependent variable and explanatory variables. Regression analysis provides valuable insights into the underlying dynamics of various phenomena.

*Corresponding author: AMALARE, A. A.
E-mail address: amalareasm@gmail.com

https://doi.org/10.60787/jnamp.vol69no2.531

Its adaptability and robust theoretical underpinnings ensure its continued relevance in both academic research and practical applications, from modeling scientific processes to informing decision-making in complex societal and industrial contexts [1, 2].

The simple linear regression model is given by:

$$y_i = \beta_0 + \sum_{j=1}^{k} \beta_{ij} x_{ij} + \varepsilon_i, \qquad i = 1,2,\dots,n \tag{1}$$

where y_i is the ith dependent (response) variable, $x_1, \dots, x_k$ are the k independent (explanatory) variables, $\beta_j,\ j = 0, 1, 2, \dots k$ are the $k + 1$ unknown regression parameters, which values are to be estimated. ε_i is the ith stochastic (disturbance) term and it is assumed to follow the distribution of y, which is normally distributed with mean zero and a constant variance σ^2.

Given a sample of n independent observations:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \varepsilon_i, \qquad i = 1,\ 2, \dots, n \tag{2}$$

Thus,

$$y_1 = \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \cdots + \beta_k x_{1k} + \varepsilon_1$$
$$y_2 = \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \cdots + \beta_k x_{2k} + \varepsilon_2$$
$$. \quad . \quad . \quad . \quad \dots\dots\dots \ . \quad .$$
$$. \quad . \quad . \quad . \quad \dots\dots\dots \ . \quad .$$
$$y_n = \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \cdots + \beta_k x_{nk} + \varepsilon_n.$$

In vector form, equation (3) is obtained as:

$$Y = [y_1 y_2 \dots y_n] = [1x_{11} \dots x_{1k} 1x_{21} \dots x_{2k} \dots 1 \dots x_{n1} \quad \dots \dots \dots \ ..x_{nk}][\beta_0 \beta_1 \dots \beta_k] + [\varepsilon_1 \varepsilon_2 \dots \varepsilon_n] \tag{3}$$

The general form of equation (3) is given by:

$$Y = X\beta + \varepsilon \tag{4}$$

where y is an (n x 1) vector of observations on a response variable, X matrix is an n×(k+1) full rank of independent variables, β is a (k+1) ×1 vector of unknown parameters to be estimated, ε is (n×1) vector of random error [1].The parameter β in a linear regression model are commonly estimated using the Ordinary Least squares Estimator (OLSE).
The OLSE of β is given as:

$$\hat{\beta}_{OLS} = (X'X)^{-1} X'Y \tag{5}$$

Where $\hat{\beta}_{OLS}$ is an unbiased estimator β. The unbiased variance-covariance matrices of the OLS estimator is defined as:

$$Var(\hat{\beta}_{OLS}) = \hat{\sigma}^2 (X'X)^{-1} \tag{6}$$

Where $\hat{\sigma}^2 = \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n-p}$, proposed by Maddalla [2] and Kibra [3]. (7)

The estimator in equation (5) is widely preferred as it is the best linear unbiased estimator (BLUE), provided that the assumptions underpinning the linear regression model remain intact [4, 5]. However, in practical applications, many of these assumptions are seldom fully met. Authors such as Neter and Wasserman [6], Fomby *et al*. [7], and Ayinde *et al*. [8] have highlighted various scenarios where these assumptions may be violated. For instance, a breach in the independence of regressors in multiple linear regression models results in multicollinearity, a significant issue reported by several researchers. Despite being unbiased, the Ordinary Least Squares (OLS) Estimator becomes inefficient under multicollinearity, with parameter estimates highly sensitive to minor data variations [9, 10, 11].

This sensitivity inflates the predicted value's variance [2, 7, 12], diminishing the performance of OLS estimator under multicollinearity. Consequently, numerous authors have proposed diagnostic methods and alternative estimators, such as the ridge regression estimator developed by Hoerl and Kennard [13], alongside contributions from McDonald and Galarneau [14], Lawless and Wang [15], Wichern and Churchill [16], Gibbon [17], Kibra [3], Durogade and Kashid [18], Kibra and Shipra [19], Lukman *et al*. [20], Aslam and Ahmad [21], and Zubair and Adenomon [22].

Principal Component Regression, introduced by Hotelling [23], offers another approach, reducing correlated variables to a smaller number of uncorrelated ones with maximum variances. Additionally, partial least squares (PLS) regression, as proposed by [24, 25], integrates elements from principal component analysis and multiple linear regression. Ridge regression, first conceptualized by Hoerl and Kennard [13], is particularly noteworthy by incorporating a biasing parameter k into the diagonal of the
$X'X$ correlation matrix. Ridge regression provides a biased but more efficient estimator with a smaller mean square error (MSE) than OLS.

The biased ridge estimator with constant $k > 0$ is defined as:

$$\hat{\beta}_R = (X'X + KI)^{-1}X'Y \quad (8)$$

where k is called the ridge (biasing) parameter. If $k = 0$, then equation (8) reduces to equation (5), OLS estimator. This implies that OLS estimator is a special case of ridge estimator. The optimal value of k in ridge regression have been proposed by various authors using different estimation techniques, among these authors are: Hoerl and Kennard [13], McDonald and Galarneau [14], Hoerl *et al*. [26], Liu [27], Khalaf and Shukur [28], Durogade and Kashid [18], Ayinde*et al*. [11], Lukman*et al* [29], recently Zubair and Adenomon [22] and Kibra and Lukman [30]. When different k-values are used, the estimator is referred to as the Generalized Ridge Estimator (GRE), whereas a constant k-value results in the Ordinary Ridge Estimator (ORE). To analyze the characteristics of the ridge estimator in equation (8), its mean (expected value) is derived by calculating the expectation of equation (8).:

$$E(\hat{\beta}_R) = F\beta \quad (9)$$

where $F = (X'X + KI)^{-1}X'X$.

The bias of the estimator is defined as:

$$Bias(\beta_R) = E(\hat{\beta}_R) - \beta \quad (10)$$

Then, substituting (9) into (10) gives

$$Bias(\hat{\beta}_R) = -(X'X + KI)^{-1}K\beta. \quad (11)$$

Squaring both sides gives

$$Bias^2(\hat{\beta}_R) = K^2(X'X + KI)^{-1}\beta\beta'(X'X + KI)^{-1} \quad (12)$$

The variance of the ridge estimator is derived by taking the variance of both sides:

$$Var(\hat{\beta}_R) = Var(F\beta)$$
$$= F^2 Var(\beta)$$
$$= \lfloor (X'X + KI)^{-1} \rfloor^2 (X'X)^2 \sigma^2 (X'X)^{-1}$$

$$Var(\hat{\beta}_R) = \sigma^2 \lfloor (X'X + KI)^{-1} \rfloor^2 X'X \,. \tag{13}$$

The mean square error (MSE) of the ridge estimator is defined by:

$$MSE(\hat{\beta}_R) = \sigma^2 \lfloor (X'X + KI)^{-1} \rfloor^2 X'X + \lfloor (X'X + KI)^{-1} \rfloor^2 K^2 \beta^2 \tag{14}$$

Since $X'X$ is a positive definite matrix, there exists an orthogonal matrix $\emptyset$ such that $X'\emptyset X = E$, where $E = diag(e_i, \; e_2, \; \ldots \ldots, \; e_p)$ and $e_i, \; e_2, \; \ldots \ldots, \; e_p$ are the eigenvalues of $X'X$. Now let $\propto = \emptyset'\beta$.
The scalar Mean Square Error (MSE) is obtained as:

$$MSE(\hat{\beta}_R) = \sigma^2 \sum_{i=1}^{p} \frac{e_i}{(e_i+k)^2} + K^2 \sum_{i=1}^{p} \frac{\alpha_i^2}{(e_i+k)^2} \tag{15}$$

where $\propto_i$ is the i^{th} element of the vector $\propto = \emptyset'\beta$.

Considering the above, the trace of the variance-covariance matrix is the MSE of the OLS estimator given by:

$$MSE\,(\hat{\beta}_{OLS}) = \sigma^2 trace(X'X)^{-1} = \sigma^2 \sum_{i=1}^{p} \frac{1}{e_i} \tag{16}$$

Different biasing parameter *k*, for GRE are found in previous studies such as [13], which produced *k* optimum as:

$$K_i = \frac{\sigma^2}{\alpha_i^2}, \quad i = 1, 2, \ldots \ldots .., p \tag{17}$$

Since σ^2 and α_i^2 are unknown generally, it is suggested that they should be replaced by their corresponding unbiased estimates $\hat{\sigma}^2$ and $\hat{\alpha}_i^2$ (Hoerl and Kennard [13]. Hence,

$\hat{K}_i = \frac{\hat{\sigma}^2}{\hat{\alpha}_i^2}$, where $\hat{\sigma}^2$ is defined in equation (7).

Numerous biasing ridge parameters k for the ORE are documented in previous studies, including the parameter proposed by [13], expressed as:

$$\hat{K}_{HK} = \frac{\hat{\sigma}^2}{Max(\hat{\alpha}_i^2)} \tag{18}$$

However, the violation of homoscedasticity assumption leads to heteroscedasticity, which occurs in cross sectional data. The penalties of using OLS estimator to estimate the parameters when heteroscedasticity exists are biased and inefficient estimates, which makes the usual hypothesis to be incorrect. In heteroscedasticity problem, the variances of the disturbance terms are no longer equal, that is, $E(\varepsilon'\varepsilon) \neq \sigma^2\Omega$. The resulting model is the Generalized least square (GLS) with Ω [31]. Given a non-singular symmetric matrix P : $\Omega = P'P$ is positive definite.

Then applying P^{-1} to equation (4) gives:

$$P^{-1}Y = P^{-1}X\beta + P^{-1}\varepsilon \tag{19}$$

and transforming it gives:

$$Y^* = X^*\beta + \varepsilon^* \quad (20)$$

The variance of the disturbance term (ε_i^*) in equation (20) is homoscedastic (equal variance). From equation (20), the OLS estimates of the transformed model have all the optimal properties of the OLS and the usual inferences are valid. By Gauss-markov theorem [32], the best linear unbiased estimator of β via the transformed model (20) is defined as:

$$\hat{\beta}_{GLS} = \left(X^{*'}X^*\right)^{-1}X^{*'}Y^* \quad (21)$$

Which is equivalent to:

$$\widehat{\beta}_{GLS} = [X'\Omega^{-1}X]^{-1}X'\Omega^{-1}Y \quad (22)$$

Where $\Omega^{-1} = P^{-1'}P^{-1}$, the Aitken has shown that GLS estimator $\hat{\beta}$ of β as defined in (21) is efficient among the class of linear unbiased estimators of β with variance-covariance of β given by:

$$Var(\hat{\beta}_{GLS}) = \sigma^2(X'\Omega^{-1}X)^{-1}. \quad (23)$$

When there is heteroscedasticity problem in a data set, the OLS estimates and predicted values are unbiased and also inefficient since the estimates do not obey the minimum variance property [7] and biased estimates of the standard error making valid inferences to be unreliable [6, 9]. In order to make amends for an imbalance or lost of efficiency, different methods have been developed, these include the estimators provided by Cochran *et al.* [33], Park [34], Hartley *et al.* [35], Rao [36], Hartley and Jayatillake [37], Horn *et al*. [38], Magnus [39], White [40], Cragg [41], Shin [42], Balasiddamuni *et al.* [43 and Shin and Kim [44].

In certain instances, both multicollinearity and heteroscedasticity can coexist within a dataset. Under such conditions, the ridge regression or weighted least squares estimator alone proves insufficient to effectively address both issues simultaneously. As a result, this study was undertaken to assess the performance of the proposed new estimators in managing heteroscedasticity in scenarios with zero multicollinearity, as well as in cases where both problems are present in the dataset.

2. MATERIALS AND METHODS

2.1 Derivation of the Proposed New Estimator

Recall equation (8), $\hat{\beta}_R = (X'X + KI)^{-1}X'Y$, and applying OLS estimator to equation (20) gives GLS estimator ($\hat{\beta}_{GLS} = \left(X^{*'}X^*\right)^{-1}X^{*'}Y^*$ as in equation (21). Therefore, the proposed ridge estimator is the mathematical combination of equations (8) and (21), which gives

$$\hat{\beta}_{Proposed} = \left(X^{*'}X^* + KI\right)^{-1}X^{*'}Y^*. \quad (24)$$

Then, solving equation (24) further gives

$$\hat{\beta}_{WRE} = [X'\Omega^{-1}X + KI]^{-1}(X'\Omega^{-1}Y) \quad (25)$$

As developed by Amalare *et al.* [45], assumed Ω^{-1} is known, though it is always unknown in real situation. It is often estimated but to have heteroscedasticity corrected measure for the model, weight variables are needed as suggested by Fuller and Rao [46] and Carroll and Ruppert [47].

This present study used the following proposed and existing estimators. Thus, applying OLS estimator into the transformed model (20), led to the formulation of weighted least square estimator, which includes ordinary least square estimator with real weight (OLSRW) and three (3) others: OLSW1, OLSW2, OLSW3. Furthermore, if GRE and ORE are applying into equation (20), it resulted into the proposed new estimators: Generalized ridge estimator with real weight (GRERW), ORERW, GREW1, OREW1, GREW2, OREW2, OREW3 and GREW3.

2.2 Model Formulation for Simulation Study

The multiple linear regression in equation (26) is considered in order to examine the proposed and existing estimators:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \varepsilon_i \qquad (26)$$

where $\varepsilon_i \sim N(0, \sigma_i^2)$. The X_i variables are fixed independent predictors with no multicollinearity, while y_i represents the response variable, and the β coefficients are predefined values.

2.3 Procedure for Generating the error term with varying of Heteroscedasticity

The error term was generated based on the normal variate distribution outlined in equation (27) to reflect different heteroscedasticity patterns. The study considered various heteroscedasticity structures, including: $\sigma_i^2 = \sigma^2 ABS(X_{i1})$ [48], $\sigma^2 X_{i1}^2$ [34], $\sigma^2(1 + X_{i1})^2$ [49], $\sigma^2 exp(X_{i1})$ [50] and $\sigma^2 \lfloor E(Y_i) \rfloor^2$. Following the distribution of the standard normal variate, $\varepsilon_i \sim N(0, \ \sigma_i^2)$

$\varepsilon_i = Z\sigma_i$, where $Z \sim N(0, 1)$

$$= Z\sigma\sqrt{\Omega} \ . \qquad (27)$$

2.4 Procedure for Generating Explanatory Variables

The simulation procedure provided by [14] and used by Wichern and Churchill [16], Gibbons [17], Kibra [3], Mansson *et al.* [51], Kibra and Lukman [30] and Idowu *et al.* [52], was adopted to generate explanatory variables in the study.

This is given as:

$$X_{ti} = (1 - \rho^2)^{\frac{1}{2}} Z_{ti} + \rho Z_{tp} \qquad (28)$$

t = 1, 2, 3,…, n, i = 1, 2, …, p, where $Z_{ti} \sim N(0, \ 1)$, rho (ρ) represents the correlation between any two explanatory variables. In this study, a multicollinearity level of zero is assumed (ρ = 0), and p denotes the number of explanatory variables.

2.5 Procedure for Generating the Response Variable

The true values of the model parameters in equation (26) were set as $\beta_0 = 4.0, \beta_1 = 3.4, \beta_2 = 4.5$ and $\beta_3 = 6.0$. After generating X_i with zero multicollinearity and error terms reflecting various heteroscedasticity patterns, the dependent variable values were computed using equation (26). A Monte Carlo simulation was performed 1,000 times, varying parameters across five error variances ($\sigma_i^2 = 0.01, 1.0, 25, 100, 625$), five known heteroscedasticity structures, and seven sample sizes (n = 15, 20, 30, 50, 100, 250, 500).

2.6 Criterion for Evaluation and Performance of the Proposed new Estimator

The performance of the proposed estimator is assessed through an examination, evaluation, and comparison based on its finite sampling properties, particularly the mean square error (MSE), which consists of the variance and the squared bias of the estimator. The MSE is expressed as:

$$MSE(\hat{\beta}_i) = \frac{1}{1000}\sum_{j=1}^{1000}(\hat{\beta}_{ij} - \hat{\beta}_i)^2 \qquad (29)$$

The estimated MSE for each estimator $(\hat{\beta})$ is calculated for each replicate. The best estimator is the one with the lowest estimated MSE.

RESULTS AND DISCUSSION

3.1 Simulation Result

Figures 1 - 6 visualize the complete summary of the simulated results under the mean square error criterion at the five (5) known heteroscedasticity natures with multicollinearity level of zero, as well as various error variances and sample sizes. Figures 7 - 12 illustrate the known but presumed unknown heteroscedasticity structures.

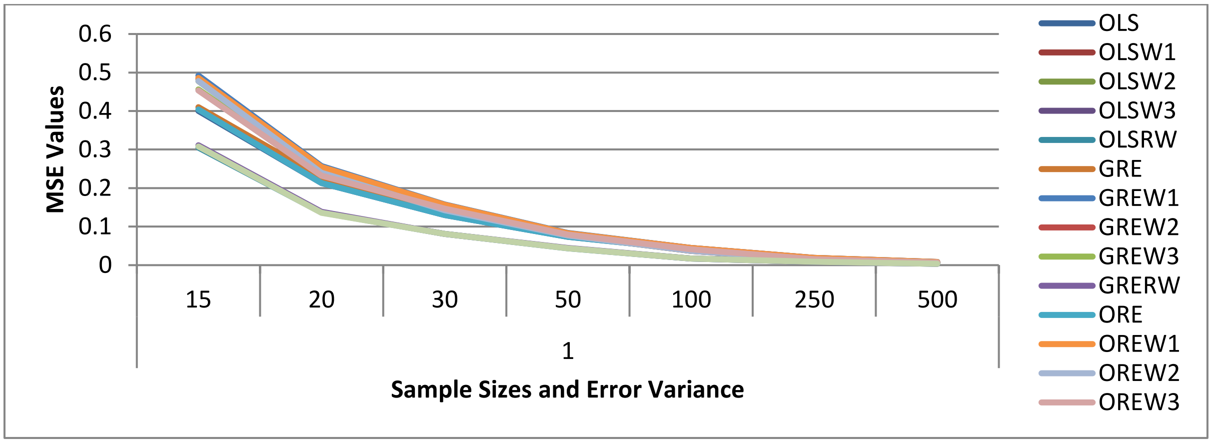

Figure 1: Graphical Representation of the Mean Squares Error of the Estimators at different Sample Sizes with known Nature of Heteroscedasticity of the form ABS(X) and Error Variance of 1.0.

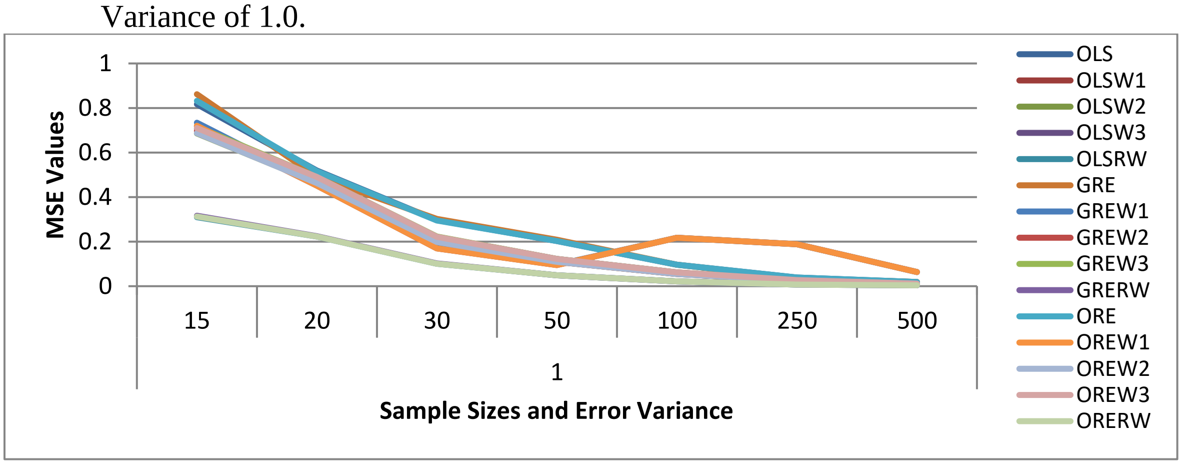

Figure 2: Graphical Representation of the Mean Squares Error of the Estimators at different Sample Sizes with known Nature of Heteroscedasticity of the form $(1 + X)^2$ and Error Variance of 1.0.

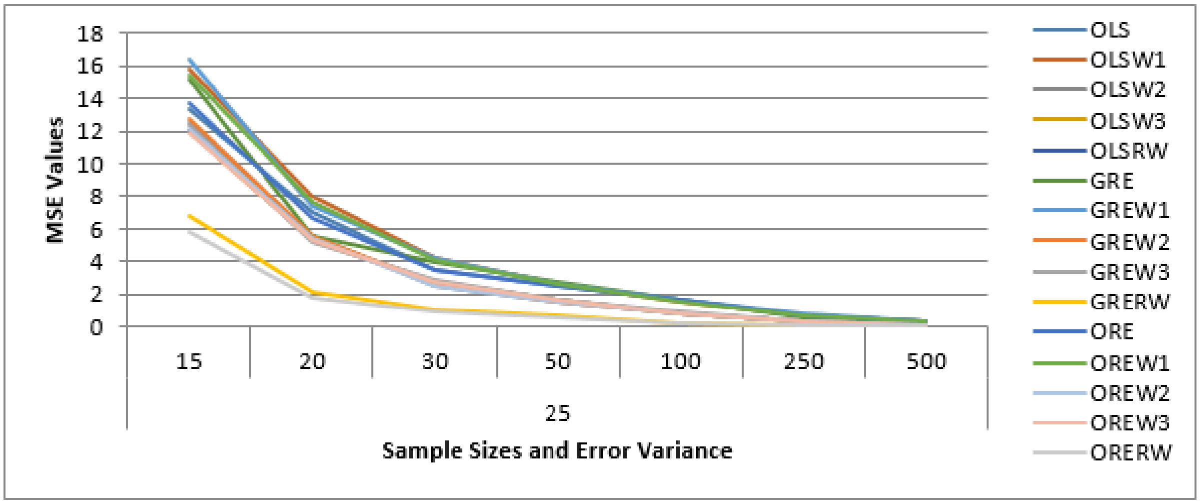

Figure 3: Graphical Representation of the Mean Squares Error of the Estimators at different Sample Sizes with known Nature of Heteroscedasticity of the form $(X)^2$and Error Variance of 25.

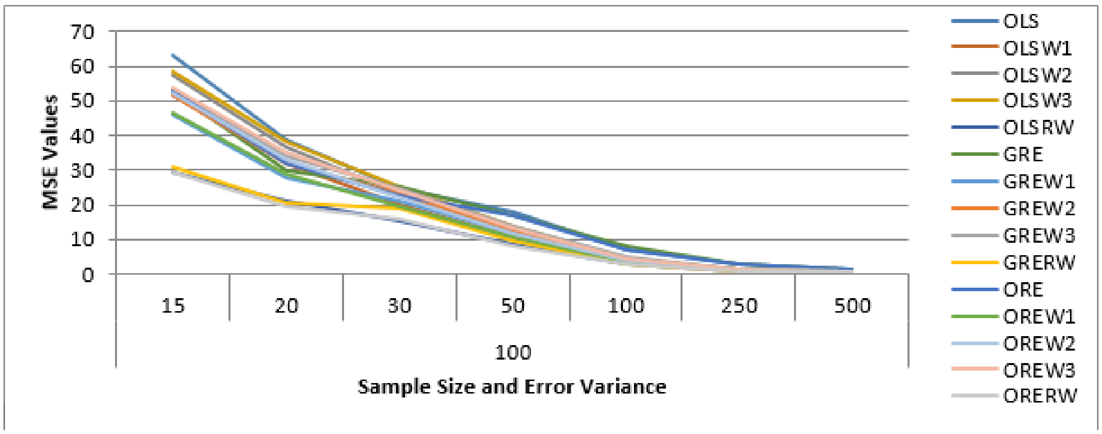

Figure 4: Graphical Representation of the Mean Square Errors of the Estimators at different Sample Sizes with known Nature of Heteroscedasticity of the form Exp(X) and Error Variance of 100.

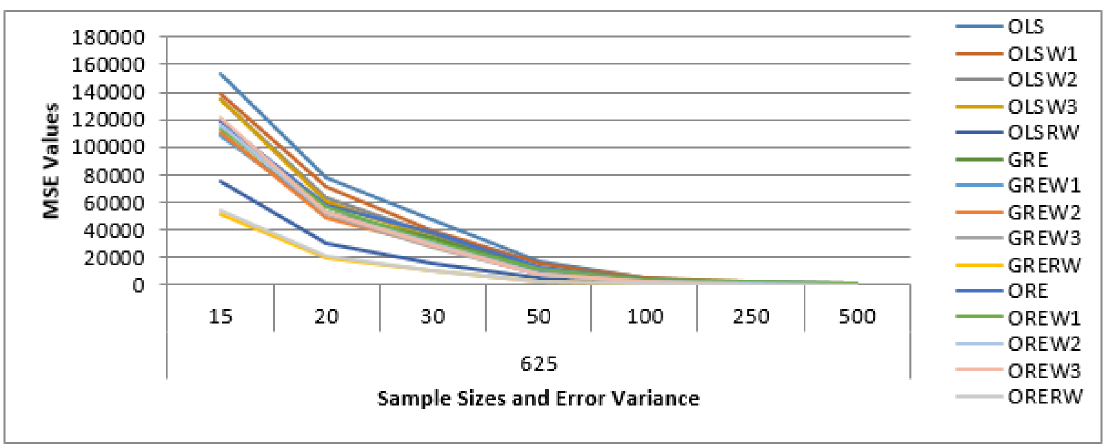

Figure 5: Graphical Representation of the Mean Square Errors of the Estimators at different Sample Sizes with known Nature of Heteroscedasticity of the form $[E(Y)]^2$ and Error Variance of 625.

Table 1: Number of Times Each Estimator Produced Minimum mean square errors when counted over the known natures of Heteroscedasticity structure and Error Variance

Sample Size	Estimators														
	OLS	OLSW1	OLSW2	OLSW3	OLSRW	GRE	GREW1	GREW2	GREW3	GRERW	ORE	OREW1	OREW2	OREW3	ORERW
15	0	0	0	0	**12**	0	0	0	0	7	0	0	0	0	6
20	0	0	0	0	4	0	0	0	0	7	0	0	0	0	**14**
30	0	0	0	0	8	0	0	0	0	5	0	0	0	0	**12**
50	0	0	0	0	2	0	0	0	0	7	0	0	0	0	**16**
100	0	0	0	0	6	0	0	0	0	3	0	0	0	0	**16**
250	0	0	0	0	6	0	0	0	0	6	0	0	0	0	**13**
500	0	0	0	0	5	0	0	0	0	6	0	0	0	0	**14**
TOTAL	0	0	0	0	43	0	0	0	0	41	0	0	0	0	91
RANK	9.5	9.5	9.5	9.5	2	9.5	9.5	9.5	9.5	3	9.5	9.5	9.5	9.5	1

NOTE: Estimator with highest frequency is bolded.

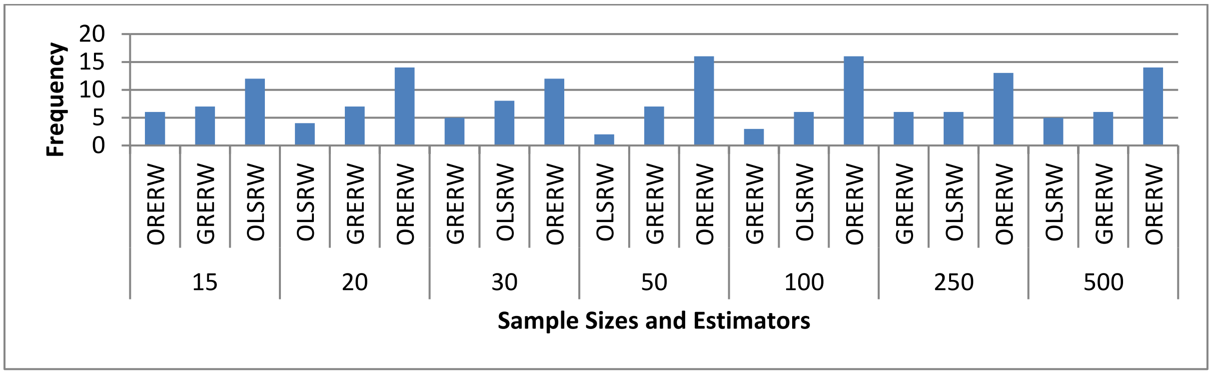

Figure 6: Graphical Representation of the Frequency of the Best Estimators under Mean Squares error Criterion at different Sample Sizes when there is known Natures of Heteroscedasticity in the Model.

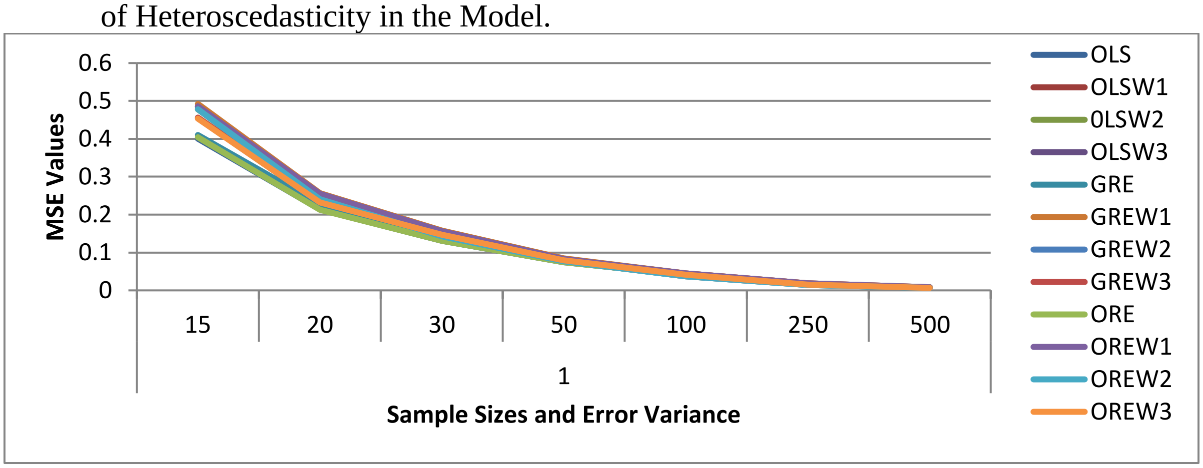

Figure 7: Graphical Representation of the Mean Square Error of the Estimators at different Sample Sizes with known but Assumed to be Unknown Nature of Heteroscedasticity of the form ABS(X) and Error Variance of 1.

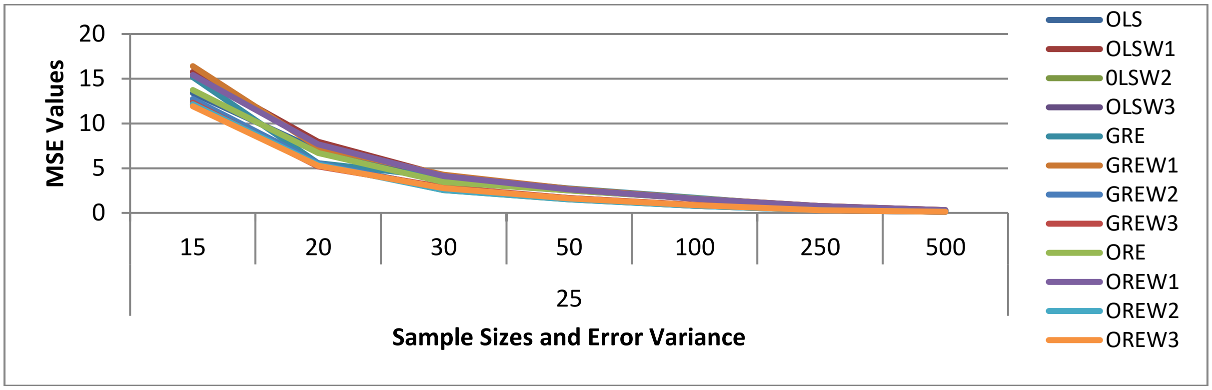

Figure 8: Graphical Representation of the Mean Square Error of the Estimators at different Sample Sizes with known but Assumed to be Unknown Nature of Heteroscedasticity of the form $(X)^2$ and Error Variance of 25.

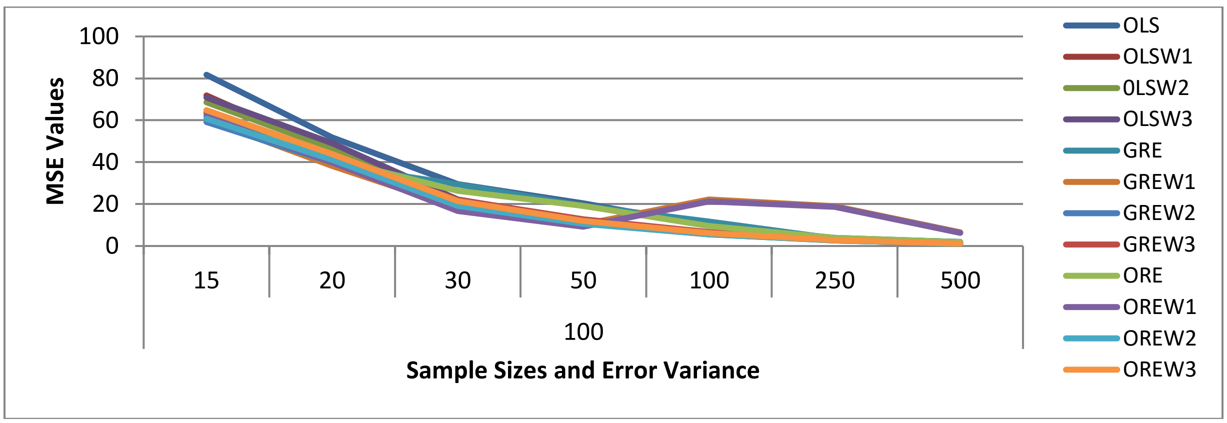

Figure 9: Graphical Representation of the Mean Square Error of the Estimators at different Sample Sizes with known but Assumed to be Unknown Nature of Heteroscedasticity of the form $(1 + X)^2$ and Error Variance of 100.

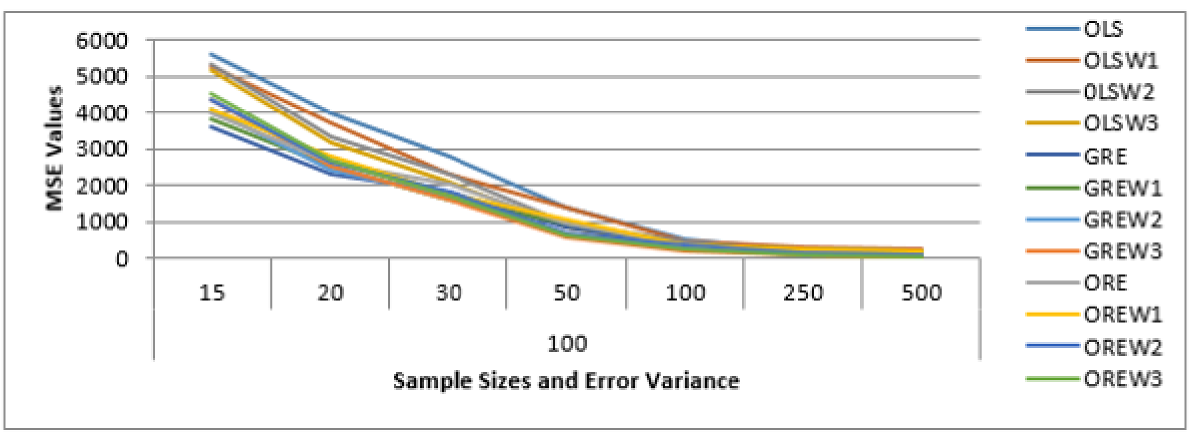

Figure 10: Graphical Representation of the Mean Square Error of the Estimators at different Sample Sizes with known but Assumed to be Unknown Nature of Heteroscedasticity of the form $[E(Y)]^2$ and Error Variance of 100.

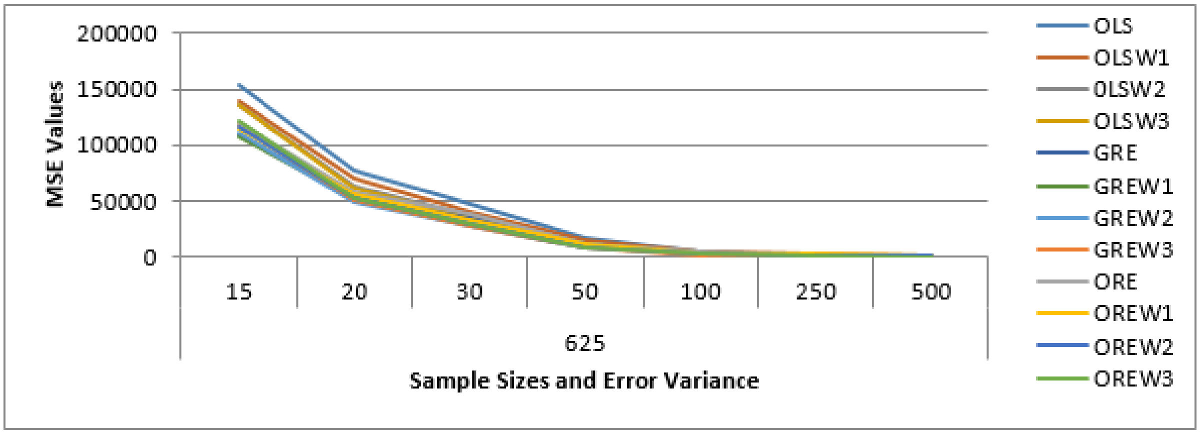

Figure 11: Graphical Representation of the Mean Square Error of the Estimators at different Sample Sizes with known but Assumed to be Unknown Nature of Heteroscedasticity of the form Exp(X) and Error Variance of 625.

Table 2: Number of Times Each Estimator Produced Minimum Mean Square Error when counted over the known but assumed to be unknown natures of Heteroscedasticity and Error Variances.

Sample Size	Estimators OLS	OLS W1	OLS W2	OLS W3	GRE	GRE W1	GRE W2	GRE W3	ORE	ORE W1	ORE W2	ORE W3
15	3	3	2	4	**5**	3	1	1	1	0	1	1
20	0	1	0	0	**7**	3	2	2	2	4	0	4
30	0	**5**	1	1	2	**5**	1	1	3	3	3	0
50	0	0	0	1	1	2	2	2	4	**8**	4	1
100	0	1	3	3	0	0	1	3	0	5	**9**	0
250	0	3	3	2	0	0	1	4	0	1	**11**	0
500	0	1	1	1	0	0	1	4	0	5	**11**	1
TOTAL	3	13	10	12	16	13	9	17	10	26	39	7
RANK	12	5.5	8.5	7	4	5.5	10	3	8.5	2	1	11

NOTE: Estimator with highest frequency is bolded.

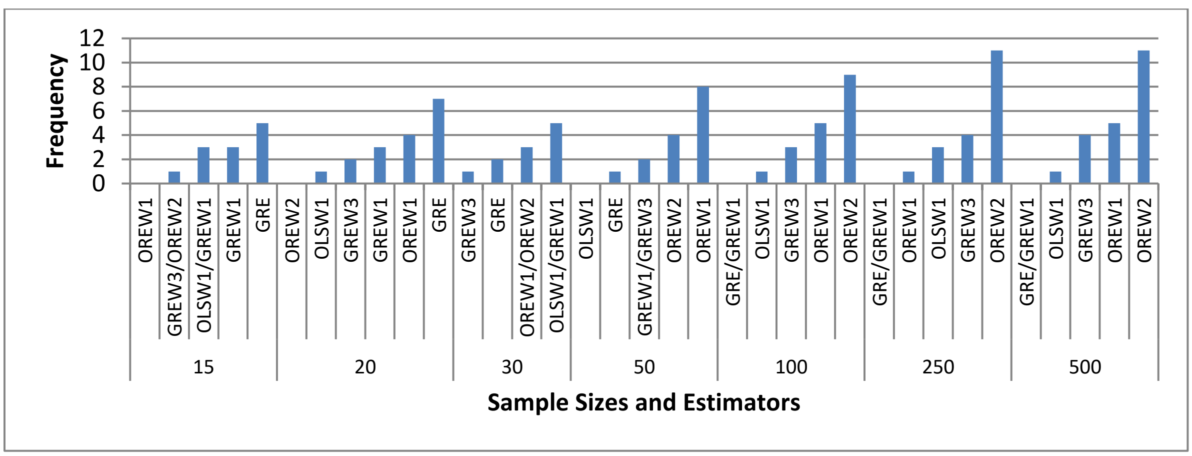

Figure 12: Graphical Representation of the Frequency of the Best Estimators under Mean Square Error Criterion at different Sample Sizes with known but Assumed to be Unknown Natures of Heteroscedasticity in the Model.

3.2 Application to Real Life Dataset

3.2.1 Passenger Car Mileage Data

This study used the dataset originally adopted by Seber [53] and later used in the subsequent analysis by Wooldridge [54] and Seber and Lee [55]. The dataset adopted by [53], comprises 81 Cars or observations about the Passenger Car Mileage with four (4) predictors in the model.

Table 3: Variable Description

Variable Name	Description
x_{i1}	VOL (Cubic feet of Cab Space)
x_{i2}	SP (Top speed, miles per hour)
x_{i3}	HP (Engine horsepower)
x_{i4}	WT (Vehincle weight, hundreds of pounds)
Y	MPG (Average miles per gallon).

Table 3 provides a comprehensively described each variable used in the models. The independent variables were standardized so that their mean would be zero and their variance of 1. The Variance inflation factor (VIF) proposed by authors [56, 57] was employed to diagnose multicollinearity in the model and some other statistics, amongst which are LM test and Durbin-Watson statistics proposed by White [40] Durbin-Watson [58], to diagnose for heteroscedasticity and auto-correlation respectively.

Table 4: Summary of Results obtained from Passenger Car Mileage Data.

Estimator	**Statistics**	**Parameter estimates and Some other Statistics**									**MSE**	**RANK**
		β_1	β_2	β_3	β_4	F Value	VIF_{Max}	LM test	DW/ Adj.R^2	Swz.BIC		
OLS	Est. Coef	0.762	0.034	0.053	0.016	255.94	6.778	14.41 (0.00)	2.3699 (0.957) 0.8586	103.402	0.1869	9
	Std. Error	0.445	0.075	0.034	0.059							
	p-value	0.093	0.000	0.191	0.788							
OLSW1	Est. Coef	0.888	0.029	0.664	0.033	212.57	3.466	1.118 (0.29)	2.3462 (0.949) 0.7923	115.507	0.1697	8
	Std. Error	0.410	0.077	0.036	0.043							
	p-value	0.035	0.001	0.071	0.448							

OLSW 2	Est. Coef	0.850	0.027	0.073	0.012	265.09	6.756	0.458 (0.50)	2.559 (0.991) 0.8501	112.873	0.1639	7
	Std. Error	0.402	0.071	0.029	0.044							
	p-value	0.04	0.000	0.016	0.785							
OLSW 3	Est. Coef	0.956	0.029	0.046	0.013	1050.6	24.79	0.436 (0.51)	2.319 (0.938) 0.9784	102.106	0.0805	5
	Std. Error	0.283	0.046	0.019	.0016							
	p-value	0.002	0.000	0.019	0.437							
GRE	Est. Coef	0.578	0.036	0.007	-0.004						0.1403	6
	Std. Error	0.326	0.006	0.003	0.015							
GRE W1	Est. Coef	0.732	0.031	0.008	0.014						0.0253	4
	Std. Error	0.338	0.007	0.003	0.024							
GRE W2	Est. Coef	0.694	0.029	0.008	-0.004						0.0249	3
	Std. Error	0.328	0.006	0.003	0.015							
GRE W3	Est. Coef	0.879	0.031	0.005	0.005						0.0060	1
	Std. Error	0.260	0.004	0.002	0.005							
ORE	Est. Coef	0.761	0.034	0.005	0.016						0.5744	10
	Std. Error	0.429	0.007	0.004	0.015							
ORE W1	Est. Coef	0.887	0.029	0.007	0.033						0.7679	11
	Std. Error	0.409	0.008	0.004	0.024							
ORE W2	Est. Coef	0.849	0.027	0.007	0.012						0.7143	12
	Std. Error	0.402	0.007	0.003	0.015							
ORE W3	Est. Coef	0.930	0.030	0.005	0.013						0.0062	2

	Std. Error	0.276	0.004	0.002	0.005							

Source: Computer Output.

Table 4 shows the summary of results obtained from passenger car mileage data regression analysis using the proposed and existing regression estimators.

i. The use of OLS estimation method shows that the data set has no multicollinearity problem since $VIF = 6.778 < 10$ but there is heteroscedasticity problem as white test equal to 14.41 ($p < 0.05$).

ii. Correcting for heteroscedasticity alone with OLSW1, OLSW2 and OLSW3. The OLSW3performs better than OLS estimator and the rest. The MSE (Mean square error) of OLSW3 (0.0805) is less than MSE of OLS (0.1868), OLSW1 (0.1697) and OLSW2 (0.1639). Although, OLSW3 introduced multicollinearity problem in the data set with VIF value of 24.79 which is greater than 10.

iii. Correcting for heteroscedasticity and multicollinearity problems using existing and proposed estimators. The most efficient estimator is GREW3 (MSE = 0.0060) followed by OREW3 (MSE = 0.0062) which are better than OLS (MSE = 0.1869). All these are presented in Table 4 and graphically presented in Figure 13.

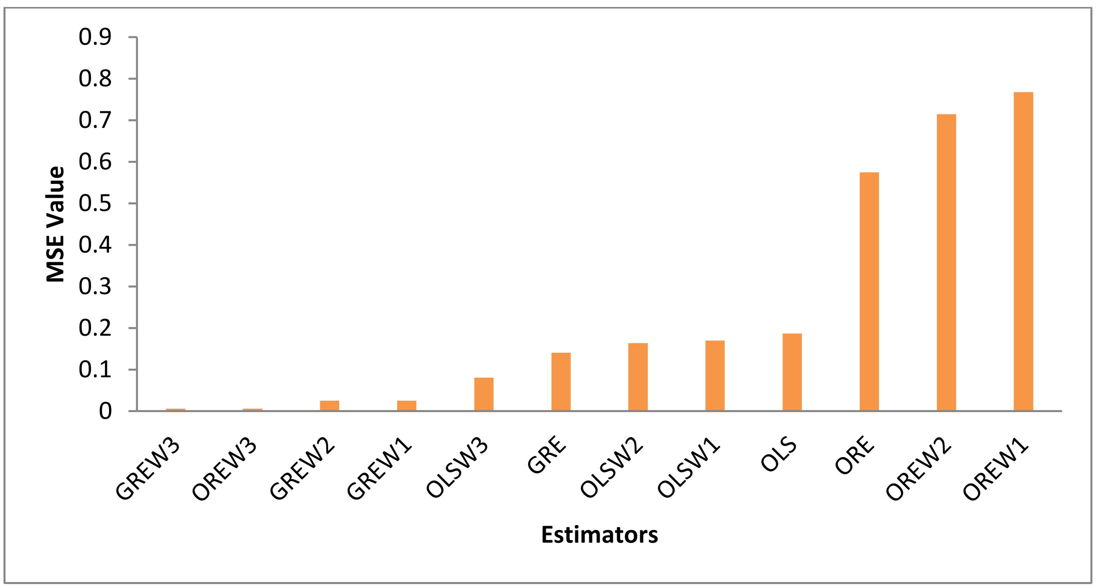

Figure 13: Mean Square Error of the estimators with Passenger Car Mileage Data

DISCUSSION OF FINDINGS

The results presented in Figures 1 through 6 reveal that as the sample size increases, the mean square error (MSE) of the estimators generally decreases. However, some estimators' MSEs converge to zero, as shown in Figures 1, 3, 4, and 5, except for a specific case of heteroscedasticity in the form of $(1 + X)^2$, illustrated in Figure 2. As the error variance increases, the MSE of the estimators also rises, which positively influences the performance of the estimators. The graphical patterns of the MSEs of the estimators are similar across different error variances. Counting the number of times each estimator exhibits the minimum MSE across five (5) types of heteroscedasticity and five (5) error variance levels, Table 1 was compiled. The maximum frequency is 25, so the estimator with a frequency closer to 25 is considered better.

From Table 1, ORERW is the best estimator because it has the highest frequency across all sample sizes, except when n = 15, where OLSRW performs better. Additionally, GRERW is close to the best estimator when the sample size is between 15 and 20, or between 250 and 500. Overall, the top three estimators based on MSE are ORERW, OLSRW, and GRERW. Figure 6 shows their frequency counts at different sample sizes. Simulated results, shown in Figures 7 through 12, also present MSE under various sample sizes and error variances, assuming different but unknown heteroscedasticity structures. As the sample size increases, the MSE of the estimators generally decreases. However, some estimators' MSEs converge to the same value, as seen in Figures 7, 8, 10, and 11, except for the specific heteroscedasticity structure of $(1 + X)^2$.

As error variance increases, so does the MSE of the estimators, which impacts their performance. Table 2 summarizes the frequency with which each estimator produces the minimum MSE across five (5) unknown heteroscedasticity structures and five (5) levels of error variance. With a maximum frequency of 25, the closer an estimator's frequency is to 25, the better it is. From Table 2, GRE (Generalized Ridge Estimator) is the best when $n \leq 20$. For $n = 30$, GREW1 and OLSW1 perform equally better than the OLS estimator. OREW1 is the best for $n = 50$, and OREW2 performs best when the sample size is between 100 and 500. In general, the top six estimators in terms of MSE are OREW2, OREW1, GREW3, GRE, OLSW1, and GREW1. Their frequency counts at different sample sizes are shown in Figure 12.

CONCLUSION

The pursuit of understanding the true nature of the heteroscedasticity problem is essential for obtaining the most efficient parameter estimates. This study tackles the challenges that heteroscedasticity presents in linear regression models, especially when multicollinearity is present, by introducing new weighted ridge estimators. The study demonstrates their efficacy in addressing these issues under various conditions of heteroscedasticity and sample sizes. Simulation results show that the ORERW estimator is optimal when the type of heteroscedasticity is known, effectively correcting for specific heteroscedasticity patterns, while the OREW2 estimator is the most efficient when the nature of heteroscedasticity is unknown. The GREW3 and OREW3 estimators also outperform existing methods in practical applications. These results highlight the robustness of the proposed estimators in delivering reliable parameter estimates and minimizing mean square error.

Thus, it is therefore recommended that in adoption of these new estimators, researchers and practitioners are encouraged to adopt the proposed ORERW and OREW2 estimators in studies involving linear regression models with heteroscedasticity and multicollinearity. Further research on the nature of heteroscedasticity should be done by making efforts to accurately identify the true nature of heteroscedasticity in datasets, as this enhances estimator performance. Software implementation by incorporating these estimators into statistical software packages would promote their accessibility and practical application across diverse disciplines. Extension to other regression models through future studies should explore the adaptation of these estimators to nonlinear and mixed-effects regression models to expand their utility.

Policy implications of this study suggest that academic and professional training programs should include modules on advanced regression techniques, focusing on the limitations of traditional OLS methods and the advantages of robust alternatives like weighted ridge estimators. Policy frameworks in research and data-driven decision-making should emphasize the adoption of robust statistical methods to ensure the accuracy and reliability of findings, particularly in contexts where data exhibits heteroscedasticity. Investments in statistical research and the development of tools like the proposed estimators should be prioritized to enhance data analysis capabilities in critical sectors, including economics, healthcare, and environmental studies. Regulatory bodies like Chattered Institute of Statisticians of Nigeria (CISON), National Bureau of Statistics (NBS) and research institutions should establish guidelines for the application of advanced regression techniques to ensure methodological consistency and improve the quality of analytical outcomes.

REFERENCES

[1] Ekum, M. I.,Akinmoladun, O. M., Aderele, O. R. and Esan, O. A. (2015). Application of Multivariate Analysis on the effects of World Development Indicators on GDP per capita of Nigeria (1981-2013). *International Journal of Science and Technology (IJST),* 4(2):254-534.

[2] Johnston, J. (1972). Econometric Methods, 2nd Ed. McGraw-Hill Book Co., Inc., New York.

[3] Gujarati, N. D., Porter, C.D. and Gunasekar, S. (2012)' Basic Econometrics.' (Fifth Edition). New Delhi: Tata McGraw-Hill.

[4] Neter, J. and Wasserman, W. (1974) 'Applied Linear Model.' Richard D. Irwin, Inc.

[5] Fomby, T. B., Hill, R. C. and Johnson, S. R. (1984). 'Advanced econometric methods.' Springer- Verlag, New York, Berlin, Heidelberg, London, Paris, Tokyo.

[6] Ayinde, K., Lukman, A. F. and Arowolo, O.T. (2015). Robust regression diagnostics of influential observations in linear regression model. Open Journal of Statistics, 5, 1-11.

[7] Chatterjee, S and Hadi, A. S. (2006). Regression Analysis by Example (Four Edition).

[8] Lukman, A. F., Ayinde, K., Okunola, A. O., Akanbi, O. B. and Onate, C. A. (2018). Classification-Based Ridge Estimation Techniques of Alkhamisi Methods. Journal of Probability and Statistical Sciences. 16 (2), 165-181.

[9] Ayinde, K., Lukman, F., Samuel, O. O. and Ajiboye, S. A. (2018) 'Some new adjusted ridge estimators of linear regression model'. International Journal of Civil Engineering and Technology, 9 (11), 2838- 2852.

[10] Maddala, G. S. (1988). Introduction to Econometrics, Macmillan, New York.

[11] Lukman, A. F., Ayinde, K. and Ajiboye, S. A. (2017) 'Monte- Carlo study of some classification-based ridge parameter estimators'. Journal of Modern Applied Statistical Methods, 16 (1), 428-451.

[12] Lukman, A. F., Ayinde, K., Binuomote, S. and Clement, O. A. (2019a). Modified ridge-type estimator to combat multicollinearity: Application to chemical data. Journal of Chemometrics, 33 (5), e3125.

[13] Hoerl, A. E and Kennard, R.W. (1970) 'Ridge regression: Biased estimation for non-orthogonal problems.' Technometrics, 12, 55-67.

[14] McDonald, G. C. and Galarneau, D. I. (1975) 'A Monte Carlo evaluation of some ridge-type estimators'. Journal of the American Statistical Association, 70 (350), 407-412.

[15] Lawless, J. F. and Wang, P. (1976) 'A simulation study of Ridge and other Regression Estimators'. Communications and Statistics, A5, 307-323.

[16] Wichern, D. and Churchill, G. (1978) 'A comparison of ridge Estimators'. Technometrics, 20, 301-311.

[17] Gibbons, D. G. (1981) 'A Simulation Study of Some Ridge Estimators'. Journal of the American Statistical Association, 76 (373), 131-139.

[18] Kibria, B. M. (2003) 'Performance of Some New Ridge Regression Estimators'. Communications in Statistics- Simulation and Computation, 32 (2), 419-435.

[19] Dorugade, A. V. and Kashid, D. N. (2010). Alternative method for choosing ridge parameter for regression.s International Journal of Applied Mathematical Sciences, 4 (9), 447- 456.

[20] Kibra, B. M. G. and Shipra, B. (2016) 'Some ridge regression estimators and their performances'. Journal of Modern Applied Statistical Methods, 15 (1), 206- 231.

[21] Lukman, A. F., Ayinde, K., Sek, S. K., and Adewuyi, E. (2019b). A modified new two-parameter in a linear regression model. Modelling and Simulation in Engineering, 2019:6342702. Doi:doi:10.1155/2019/6342702

[22] Aslam, M. and Ahmad, S. (2020). The modified Liu-ridge-type estimator. A new class of biased estimators to address multicollinearity. Communications in Statistics-Simulation and Computation. Doi:10.1080/03610918.2020.1806324.

[23] Zubair, M. A. and Adenomon, M. O. (2021). Comparison of Estimators Efficiency for Linear Regression with Joint Presence of Autocorrelation and Multicollinearity. Science World Journal, Vol. 16 (No.2), 103- 109, ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print).

[24] Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. Journal of Educational Psychology, 24, 417-441, and 498-520.

[25] Wold, Herman (1985). 'Partial least square' In Kotz, Samuel; Johnson, Norman L. (eds.). Encyclopedia of Statistical sciences. 6. New York: Willey. 581-591.

[26] Tabakan, G. (2013) 'Performance of the difference-based estimators in partially linear models.' Statistics, 47, (2), 329-347.

[27] Hoerl, A. E., Kennard, R. W. and Baldwin, K. F. (1975). 'Ridge Regression: Some Simulations'. Communications in Statistics,4, 104-123.

[28] Liu, K. (1993) 'A new class of biased estimate in linear regression.' Communications in Statistics- Theory and Methods, 22, 393-402, MR1212418.

[29] Khalaf, G. and Shukur, G. (2005). Choosing ridge parameters for regression problems. Communications in Statistics-Theory and Methods, 34, 1177-1182.

[30] Kibra, B. M. G. and Lukman, A. F. (2020). A New Ridge-Type Estimator for the Linear Regression Model: Simulations and Applications. Scientifica. https//doi.org10.1155/2020/9758378.

[31] Aitken, A. C. (1935). 'On Least Squares and Linear Combination of Observations'. Proceedings of the Royal Statistical Society of Edinburgh, 55, 42-48.

[32] Markov, A. A. (1900). 'Wahrscheinlichkeitsrechnug'. Leipzig; Tuebner.

[33] Cochran, W. G. and Carroll, S. P. (1953) 'A Sampling Investigation of the Efficiency of Weighting Inversely as the Estimated Variance.' Biometrics, 9, (4), 447-459.

[34] Park, R. E. (1966). 'Estimation with Heteroscedastic Error Terms'. Econometrica, 34, 888-892.

[35] Hartley, H. O., Rao, J.N.K. and Kiefer, G. (1969) 'Variance Estimation with One Unit per Stratum'. Journal of the American Statistical Association, 64, 841-851.

[36] Rao, C.R. (1970) 'Estimation of Heteroscedasticity Variances in Linear Models'. Journal of the American Statistical Association 65, 161-172.

[37] Hartley, H. O. and Jayatillake, K. S.E. (1973) 'Estimation for Linear Models with Unequal Variances'. Journal of the American Statistical Association. 68, 189-192.

[38] Horn, S. D., Horn, R.A. and Duncan, D.B. (1975). 'Estimating Heteroscedastic Variances in linear model'. Journal of the American Statistical Association, 70, 380-385.

[39] Magnus, J. R. (1978) 'Maximum Likelihood Estimation of the GLS Model with unknown Parameter in the Disturbance Covariance Matrix.'' Journal of Econometrics, 7, 281-312.

[40] White, H. (1980) 'A Heteroskedastic-Consistent Covariance Matrix Estimator and A Direct Test for Heteroskedasticity.' Econometrica , 48 , 817-838.

[41] Cragg, J. G. (1983) 'More Efficient Estimation in the Presence of Heteroskedasticity of Unknown Form.' Econometrica, 51, 751-763.

[42] Shin,H. C.(2013) 'Weighted Least Squares estimation with Sampling Weights.' Proceedings of the Survey Research Methods Section, the American Statistical Association, 1523-1530.
[43] Balasiddamuni, P., Prakash, K., Rao, K., Prasad, A., Abbaiah, R. and Rayalu, G. (2013) 'A Minimum Quadratic Unbiased Estimation (MINQUE) of Parameters In A Linear Regression Model With Spherical Disturbances.' International Journal of Scientific and Technology Research, 2, (5), 135-138.
[44] Shin, H. C. and Kim, J. (2014) 'Weighted Least Squares Estimation with Simultaneous Consideration of Variances and Sampling weights.' Proceedings of the Survey Research Methods Section, the American Statistical Association, 2972- 2978.
[45] Amalare, A. A., Ayinde, K. and Onanuga, K. (2023). Parameter Estimation of Linear Regression Model with Multicollinearity and Heteroscedasticity Problems. Journal of the Nigerian Association of Mathematical Physics. Vol. 65, 207- 216.
[46] Fuller, W. A. and Rao, J. N. K. (1978) 'Estimation for a Linear Regression Model with Unknown Diagonal Covariance Matrix'. Annals of Statistics, 6 (5), 1149-1158.
[47] Carroll, R. J. and Ruppert, D. (1988) 'Transformation and Weighting in Regression'. New York, NY: Chapman & Hall.
[48] Glesjer, H. (1969) 'A New Test for Heteroscedasticity'. Journal of the American Statistical Association, 64, 316- 323.
[49] Bruesch, T. S. and Pagan, A. R.(1979)' A Simple test of heteroscedasticity and random coefficient variation'. Econometrica, 47, 1287 – 1294.
[50] Box, G. E. P. and Hill, W. J.(1974) 'Correction Inhomogeneity of Variance with Power Transformation Weighting'. Technometrics, 16 (3), 385-389.
[51] Mansson, K., Shukur, G. and Kibra, B. M. G. (2010). On some Ridge Regression Estimators: A Monte Carlo Simulation study under different error variances. Journal of Statistics, 17(1), 1-22.
[52] 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 1:188-204. https://doi.org/10.46481/asr.2022.1.3.57
[53] Seber, F. A. F. (1977). Linear Regression Analysis, John Wiley & Sons, New York, p.64.
[54] Wooldridge, J. M. (2000). Introductory Econometrics: A Modern Approach, South-Western Publishing.
[55] Seber, G. A. and Lee, A. J. (2012). Linear Regression Analysis, Vol. 329: John Wiley & Sons.
[56] Marquardt, D. W. (1970). 'General inverse, Ridge Regression, Biased Linear Estimation and Non-Linear Estimation'. Technometrics, 12, 591-612.
[57] Kleinbaum D. G., Kupper, L. L. and Muller, K. E. (1988) 'Applied Regression Analysis and Other Multivariate Methods'. Second Edition, PWS-Kent, Boston, Mass, Pp. 210.
[58] Durbin, J. and Watson, G. S. (1951). 'Testing for Serial Correlation in Least Squares Regression'. Biometrika, Vol. 38, 159- 171.