IRAQI STATISTICIANS JOURNAL

https://isj.edu.iq/index.php/isj

ISSN: 3007-1658 (Online)

New Stochastic Restricted Estimator: A Step Towards Improved Modeling in Linear Regression

Dounia J. Al-Obaidi[1] and Mustafa I. Alheety[2]

[1,2]Department of Mathematics, College of Education for Pure Sciences, University Of Anbar, Ramadi, Iraq

ARTICLE INFO

Article history:
Received 14 March 2025
Revised 16 March 2025,
Accepted 27 April 2025
Available online 28 April 2025

Keywords:
Biased estimation;
Linear convex combination;
Liu estimator,
Mixed estimator,
Mean squared error matrix.

ABSTRACT

Here, we propose a new convex linear combination estimator, called the New Mixed Estimator (NME), for multiple linear regression models with stochastic linear constraints on unknown parameters and multicollinearity between explanatory variables. The Ordinary Mixed Estimator (OME) and the Biased Stochastic Restricted Liu-type Estimator (SRLIT), two well-known estimators, are integrated to create the NME. Using Mean Square Error (MSE) as the main performance parameter, we examine the statistical characteristics of the NME and theoretically compare it to both OME and SRLIT to show its superiority. According to our research, the NME routinely performs better than the OME and SRLIT in terms of overall efficacy and statistical characteristics. Additionally, we provide a numerical example to clearly support these findings.

1. Introduction

The sample correlation matrix in multiple linear regression frequently deviates significantly from the identity matrix due to a lack of orthogonality caused by linked explanatory factors. These variations can lead to unstable ordinary least squares estimations (OLSE) of regression coefficients, which are often manifested by inflated standard errors and unreasonably high absolute values for a large number of estimated coefficients. Considerable research have been done on the difficulties multicollinearity presents and its statistical ramifications in linear regression models. There have been several solutions put forth to address these problems, the main one being the employment of biased estimators in place of the OLS estimator. Under conditions of multicollinearity, it has been demonstrated that these biased estimators improve the accuracy of parameter estimation.

Principal component regression estimators [9], ordinary ridge regression estimators (ORE) [1], r-k estimators [2], Liu estimators (LE) [3], r-d estimators [4], and Liu-type estimators [5,6] are among the various forms of biased estimators that have been presented. While Akdeniz and Kakaryanlar [7] presented the approximately unbiased generalized Liu estimator (AUGLE), Singh et al. offered the approximately unbiased generalized ridge estimator (AUGRE) using a steepening approach. Akdeniz and Erol[8] investigated the approximately unbiased ridge estimator and the approximately unbiased LE estimator as bias-corrected substitutes for ORE and LE,

Corresponding author E-mail address: dou23u2006@uoanbar.edu.iq
https://doi.org/10.62933/isj.v2i1.13
This work is an open-access article distributed under a CC BY License (Creative Commons Attribution 4.0 International) under https://creativecommons.org/licenses/by-nc-sa/4.0/

respectively, after looking more closely at specific cases of AUGRE and AUGLE.

Convex linear combination estimators are an effective statistical technique for combining several estimators, improving accuracy, robustness, and efficiency in a range of applications such as covariance matrix estimation, density estimation, and regression. These estimators, which provide advantages including increased accuracy, resilience, and optimality, are essential to statistical modeling. They also exhibit flexibility and computational efficiency, which makes them appropriate for a range of applications. Convex linear combination estimators are more dependable and efficient in real-world situations since they also offer assurances regarding small sample sizes and control over false discovery rates. These approaches provide estimators that reproduce the behavior of individual estimators while attaining better overall performance by optimizing the weights of the combinations.

The need for techniques that improve regression models' performance has grown in recent years. The recently suggested Convex Combination Estimator is one of these approaches that shows promise. By combining multiple conventional estimators into a convex mixture, this estimator capitalizes on the benefits of each estimator. A better balance between bias and variance can be achieved by weighting each estimator based on how well it performs in various scenarios. This study presents a new convex estimator that combines a biased SRLTE with an unbiased OME. The MSE was used to evaluate this novel estimator's performance. The paper's structure is set out as follows: The model specifications and details of the proposed estimator are described in Section 2;the efficacy of the new estimator is evaluated in Section 3; and a Monte Carlo simulation study and numerical example are presented in Sections 4 ,5 and 6 to demonstrate the behavior of the estimator.

2. Model Specification and the Proposed Estimator

We consider a multiple linear regression model expressed as follows:

$$Y = X\beta + \epsilon \qquad (1)$$

where Y is an $n \times 1$ vector representing the response variable, , X is an $n\times p$ known design matrix of rank p , β is a $p \times 1$ vector of unknown parameters and ϵ is an $n \times 1$ vector of disturbances. It is assumed that the expected value of the disturbances satisfies $E(\epsilon) = 0$, and the variance-covariance matrix is given by $Var(\epsilon) = \sigma^2 I_n$. where I_n is the $n \times n$ identity matrix. The OLS for model (1)is given by:

$$\hat{\beta}_{Ols} = S^{-1}X'Y \qquad (2)$$

where $S = X'X$. The OLSE may become unreliable in the presence of multicollinearity, which can result in inaccuracies in parameter estimates. According to Netter [10], multicollinearity occurs when one independent variable is approximately a linear combination of other independent variables. This condition can significantly compromise the accuracy of parameter estimates and lead to substantial challenges in the analysis.

The covariance matrix for any estimator β^* for β is defined as follows:

$$Cov\left(\hat{\beta}\right) = E(\beta^* - \beta)(\beta^* - \beta)' \ .$$

In the context of biased estimation, the MSE matrix serves as the most effective criterion for evaluating an estimator's performance. This is due to its ability to provide a holistic assessment by integrating both the variance-covariance matrix and the bias vector into a single formulation.

$$MSE\left(\hat{\beta}\right) = Cov\left(\hat{\beta}\right) + Bais\left(\hat{\beta}\right) Bais\left(\hat{\beta}\right)'$$

In the context of model (1), the analysis is based exclusively on sample data. However, if there exists prior knowledge regarding the parameters, a set of J independent stochastic linear limitations is introduced.

$$r = R\beta + e\,, \qquad (3)$$

where r is a $j \times 1$ stochastic known vector with$E(r) = R\beta$, R is a $j \times p$ random vector of disturbances with $E(e) = 0$ and $Var(e) = \sigma^2 W$,and W is assumed to be a known and positive definite (pd). Further, it is assumed that ϵ is stochastically independent of e.

The models presented in (1) and (3) serve as an alternative approach to addressing multicollinearity. Thiel and Goldberger [14,15] proposed a mixed estimation methodology that integrates sample data with prior information into a cohesive model. Within this framework, they introduced the OME, which demonstrates superior variance characteristics compared to the OLSE variance and exhibits enhanced statistical properties.

$$\hat{\beta}_{OME} = (S + R'W^{-1}R)^{-1}(X'Y + R'W^{-1}r) \quad (4)$$

The Expectancy, Variance and MSE Matrix of the OME can be derived as follows:

$$E(\hat{\beta}_{OME}) = \beta \quad (5)$$

$$MSE\ (\hat{\beta}_{OME}) = Cov(\hat{\beta}_{OME}) = \sigma^2\ (\ S + R'W^{-1}\ R)^{-1} \quad (6)$$

Alheety (2020) [16] introduced a novel estimator known as the Stochastic Restricted Liu-Type Estimator SRLTE.

$$\hat{\beta}_m = (S + I + R'W^{-1}R)^{-1}\ (X'Y + d\hat{\beta} + R'W^{-1}\ r) \quad (7)$$

The Expectancy, Variance and MSE Matrix of estimator the SRLTE can be derived as follows:

$$E(\ \hat{\beta}_m) = (S + I + R'W^{-1}\ R)^{-1}(S + d + R'W^{-1}R)\beta \quad (8)$$

$$Bias(\hat{\beta}_m) = (S + I + R'W^{-1}\ R)^{-1}(S + d + R'W^{-1}R)\beta - \beta \quad (9)$$

$$Cov(\hat{\beta}_m) = \ \sigma^2(S + I + R'W^{-1}\ R)^{-1}\ (S + d^2S^{-1} + R'W^{-1}\ R)(S + I + R'W^{-1}\ R)^{-1} \quad (10)$$

$$MSE(\hat{\beta}_m) = \sigma^2(S + I + R'W^{-1}\ R)^{-1}\ (S + d^2S^{-1} + R'W^{-1}\ R)(S + I + R'W^{-1}\ R)^{-1} + [(S + I + R'W^{-1}\ R)^{-1}(S + d + R'W^{-1}R)\beta - \beta]'[(S + I + R'W^{-1}\ R)^{-1}(S + d + R'W^{-1}R)\beta - \beta] \quad (11)$$

The OME and STRLE estimators are convexly combined to generate the NME estimator. When both estimators work well in specific situations, it may be useful to use a convex combination of two estimators. In order to integrate the OME and STRLE estimators, we investigate a linear convex combination estimator. Equations (4) and (7) give the linear convex combination for the NME estimator's mathematical representation:

$$\hat{\beta}_{NME} = A\ \hat{\beta}_{OME} + (I - A)\ \hat{\beta}_m \quad , \quad (12)$$

where the eq(12) is a convex matrix estimator model and A is Square matrix $p \times p$.

2.1 The properties of the proposed estimator.

We can find the MSE of the new estimator by finding the variance and bias of the estimator

$$MSE(\hat{\beta}_{NME}) = Cov(\hat{\beta}_{NME}) + B^2(\hat{\beta}_{NME}) = Cov(\hat{\beta}_{NME}) + [E(\hat{\beta}_{NME}) - \beta]^2 \quad (13)$$

The expectation of the new estimator is found by expecting OME in Equation (5) and expecting SRLITE in Equation (8).

$$E(\hat{\beta}_{NME}) = E[A\ (\hat{\beta}_{OME}) + (I - A)\hat{\beta}_m] \quad (14)$$

$$= A\ E(\hat{\beta}_{OME}) + (I - A)E(\hat{\beta}_m)$$

$$= A\beta + (I - A)(S + I + R'W^{-1}R)^{-1}(S\beta + d\beta + R'W^{-1}R\beta)$$

$$= A\beta + (I - A)(S + I + R'W^{-1}\ R)^{-1}(S + d + R'W^{-1}\ R)\beta$$

$$= A\beta + (I - A)C^*G_d\beta \quad (15)$$

where $C^* = (S + I + R'\ W^{-1}\ R\ \ )^{-1}$, $G_d = S + d + R'W^{-1}\ R$

As for bias, we find it using the following law:

$$[E(\hat{\beta}_{NME}) - \beta]^2 = [A\beta + (I - A)C^*G_d\beta - \beta]^2 \quad (16)$$

We find the variance of the NME estimator from the variance of OME in Equation (6) and the variance of SRLTE in Equation (10)

$$Cov(\hat{\beta}_{NME}) = A\ Cov(\hat{\beta}_{OME})A' + (I - A)Cov(\hat{\beta}_m(I - A)' \quad (17)$$

$$= A\sigma^2(S + R'W^{-1}\ R)^{-1}A' + (I - A)\sigma^2C^*(S + d^2S^{-1} + R'W^{-1}\ R)C^*(I - A)'$$

$$= A\sigma^2QA' + (I - A)\sigma^2C^*Q_d\ C^*(I - A\ )' \quad (18)$$

where $= (S + R'W^{-1}\ R)^{-1}$, $Q_d = S + d^2S^{-1} + R'W^{-1}\ R$.

By replacing equation (18) and equation (16) with equation (13), we find the mean square error of the new estimator.

$$MSE(\hat{\beta}_{NME}) = A\sigma^2QA' + (I-A)\sigma^2C^*Q_dC^*(I-A)' + [A\beta + (I-A)C^*G_d\beta - \beta]^2 \qquad (19)$$

By deriving mean square error in equation (19) with respect to A, we find the value of parameter A.

$$\frac{\partial\, MSE}{\partial A} = 2A\sigma^2Q - 2(I-A)\sigma^2C^*Q_dC^* + 2[A\beta + (I-A)\,C^*\,G_d\beta - \beta](\beta - C^*\,G_d\,\beta) \qquad (20)$$

$$\frac{\partial\, MSE}{\partial A} = 0 \Longrightarrow$$

$$2A\sigma^2Q - 2(I-A)\sigma^2C^*Q_dC^* + 2[A\beta + (I-A)\,C^*\,G_d\,\beta - \beta](\beta' - \beta'C^{*\prime}\,G_d{}') = 0$$

$$2A\sigma^2Q + 2A\sigma^2C^*Q_dC^* + 2A\beta(\beta' - \beta'C^{*\prime}\,G_d{}') - 2A\,C^*\,G_d\beta(\beta' - \beta'C^{*\prime}\,G_d{}')$$

$$= 2\sigma^2C^*Q_dC^* - 2C^*\,G_d\,\beta(\beta' - \beta'^{C^*}{}'G_d') + 2\beta(\beta' - \beta'C^{*\prime}\,G_d{}')$$

$$A_{opt} = (\,\sigma^2Q + \sigma^2C^*Q_dC^* + \beta\beta' - C^*G_d\beta\beta' - C^*\,G_d\beta\beta' + C^*C^*G_dG_d'\beta\beta'\,)^{-1}\;(\sigma^2C^*Q_dC^* + \beta\beta' - C^*G_d\,\beta\beta' + C^*C^{*'}\,G_d\,G_d'\,\beta\beta' - C^*G_d\beta\beta') \qquad (21)$$

When $\beta$and σ^2 are not known, we use $(\hat{\beta}_{OLS}, \hat{\sigma}^2)$ The ordinary least squares estimator

3. Evaluation of Estimator Performance

This section compares NME to both the performance of the SRLTE and the OME using the MSE criteria .We introduce some basic concepts and terms that are useful in our current research.[17]

Lemma1. For a positive definite matrix M (denoted by $M > 0$) and a vector α, the inequality $M - \alpha\alpha' \geq 0$ holds if and only if $\alpha' M\,\alpha \leq 1$.

Lemma 2. For two competing linear estimators $\hat{\beta}_j = A_j y$, $j = 1{,}2$ if the difference between the covariance matrices of these estimators, represented by $cov(\hat{\beta}_j) > 0$, $j = 1{,}2$ then the condition $MSE(\hat{\beta}_1) - MSE(\hat{\beta}_2) \geq 0$ is satisfied if and only if the inequality $d_2'(\sigma^2 D + d_1 d_1')d_2 \leq 1$ holds, where $MSE(\hat{\beta}_j)$, d_j represent the mean squared error matrix and bias vector of $\hat{\beta}_j$, respectively.

Lemma 3. For n x n matrices $M > 0\; and\;\; N \geq 0\; the\; inequality\; M > N$ holds if and only if the largest eigenvalue of the matrix NM^{-1}, denoted by $\lambda_1(NM^{-1}) < 1$.

Lemma 4. For any two n x n matrices $A\; and\; B\;, where\; A\; >\; 0\; (or\; A\; \geq\; 0)\; and\; B\; >\; 0, the\; inequality\; B\; -\; A\; >\; 0$ holds if and only if the eigenvalues of the matrix $\lambda_i^B(A) < 1$.

Lemma 5. (Hu Yang et al. 2009). Supposed that M is a real symmetric matrix and P is a matrix then $M \geq 0 \Leftrightarrow \forall P.\, P'MP \geq 0 \Leftrightarrow$ each eigenvalues of M is non negative.

Definition. defines that for two competing estimators $\hat{\beta}_1 and \hat{\beta}_2\; the\; estimator\; \hat{\beta}_2$ is considered superior to $\hat{\beta}_1$according to the Mean Squared Error (MSE) criterion if and only if $MSE(\hat{\beta}_1) - MSE(\hat{\beta}_2) \geq 0$.

3.1: Comparison between NME and OME using the MSE performance of the estimators.

We will make a more detailed Using Equation (6) and Equation (19) we find

$$\Delta_1 = MSE(\hat{\beta}_{OME}) - MSE(\hat{\beta}_{NME}) \qquad (22)$$

$$= \sigma^2Q - (A\sigma^2QA' + (I-A)\sigma^2C^*Q_dC^*(I-A)' + [A\beta + (I-A)C^*G_d\beta - \beta]^2\,)$$

$$= \sigma^2Q - (A\sigma^2QA' + (I-A)\sigma^2C^*Q_dC^*(I-A)') - \;[A + (I-A)C^*G_d - I]'\beta'\beta[A(I-A)C^*G_d - I]$$

$$=\; \sigma^2\,D_1 - b_1'b_1 \qquad (23)$$

where $b_1 = [A + (I-A)C^*G_d - I]$, $D_1 = Q - AQA' - (I-A)C^*Q_d\,C^*(I-A\,)'$

Theorem 1. The NME under the namely stochastic linear regression model, the superiority of the NME over the OME in terms of the MSE matrix is determined $\Delta_1 \geq 0$ if and only if $b_1' D_1^{-1} b_1 \leq \sigma^2$.

Proof: The MSE difference between the NME and OME is given in eq.(23) is

$\Delta_1 = \sigma^2 D_1 - b_1' b_1$. To apply lemma1 to eq.(23) we need to prove that D_1 is a positive definite matrix (pd)

Not that $D_1 = Q - (AQA' + (I - A)C^* Q_d C^*(I - A)')$

$= Q[Q^{-1} - (AQ^{-1}A' + Q^{-1}(I - A)C^* Q_d C^*(I - A)'Q^{-1})]Q$

For $Q = (S + R'W^{-1}R)^{-1}$is a positive definite (pd) (namely > 0) . There exists an orthogonal matrix P such that $P'P = PP' = I$ and $P[Q^{-1} - (AQ^{-1}A' + Q^{-1}(I - A)C^* Q_d C^*(I - A)'Q^{-1})]P' = diag\{\xi_1, \dots, \xi_P\}$ There for D_1 is (p.d) if and only if $\xi_i > 0$, $i = 1, \dots, p$

$$\xi_i = (\lambda_i + v_i) - a_i^2(\lambda_i + v_i) - \frac{(\lambda_i + v_i)^2 (1 - a_i)^2 (\lambda_i + d^2\lambda_i^{-1} + v_i)}{(\lambda_i + 1 + v_i)^2}$$

$$\xi_i = (\lambda_i + v_i)\left(1 - a_i^2 - \frac{(\lambda_i + v_i) (1 - a_i)^2 (\lambda_i + d^2\lambda_i^{-1} + v_i)}{(\lambda_i + 1 + v_i)^2}\right)$$

Therefore, $D_1\ is\ (p.d)\ if\ \xi_i > 0, \forall i$. That means D_1 is (p.d) if and only if

$$\frac{(\lambda_i + v_i) (1 - a_i)^2 (\lambda_i + d^2\lambda_i^{-1} + v_i)}{(\lambda_i + 1 + v_i)^2} < 1$$

This implies that D_1, is clearly a positive definite matrix. Hence according to lemma 1, the NME is superior to OME if and only if $b_1' D_1^{-1} b_1 \leq \sigma^2$ This completes the proof.

3.2: Comparison between NME and SRLTE using the MSE performance of the estimators.

We will make a more detailed Using Equation (11) and Equation (19) we find

$$\Delta_2 = MSE(\hat{\beta}_m) - MSE(\hat{\beta}_{NME}) \quad (24)$$

$$= \sigma^2 C^* Q_d C^* + [C^* G_d - I]'\beta'\beta[C^* G_d - I] - (A\sigma^2 QA' + (I - A)\sigma^2 C^* Q_d C^*(I - A)' + [A\beta + (I - A)C^* G_d \beta - \beta]^2)$$

$$= \sigma^2 C^* Q_d C^* + [C^* G_d - I]'\beta'\beta[C^* G_d - I] - (A\sigma^2 QA' + (I - A)\sigma^2 C^* Q_d C^*(I - A)') - [A + (I - A)C^* G_d - I]'\beta'\beta[A + (I - A)C^* G_d - I]$$

$$= \sigma^2 C^* Q_d C^* - (A\sigma^2 QA' + (I - A)\sigma^2 C^* Q_d C^*(I - A)') + b_2' b_2 - b_1' b_1$$

$$= \sigma^2 D_2 + b_2' b_2 - b_1' b_1' \quad (25)$$

where $b_2 = [C^* G_d - I]$, $D_2 = C^* Q_d C^* - AQA' - (I - A)C^* Q_d C^*(I - A)'$

Theorem 2. When the maximum Eigen value of $(AQA' + (I - A)C^* Q_d C^*(I - A)')(C^* Q_d C^*)^{-1} < 1$. then the NME is superior to the SRLTE in the MSE if and only if $b_1'(\sigma^2 D_2 + b'_2 b_2)b_1 \leq 1$

Proof: The MSE difference between the NME and SRLTE given in eq.(25) is

$\Delta_2 = \sigma^2 D_2 + b_2' b_2 - b_1' b_1'$

To show that $\Delta_2 \geq 0$, lemma 2 can be used. A requirement to apply lemma 2 is that D_2 to be positive definite matrix. It is clear that $C^* Q_d C^* > 0 \quad and \quad AQA' + (I - A)C^* Q_d C^*(I - A)' \geq 0$.

According to lemma 3 , $C^* Q_d C^* > AQA' + (I - A)C^* Q_d C^*(I - A)'$ if and only if $\lambda_1 < 1$,Where $\lambda_1$is the maximum eigen value of $(AQA' + (I - A)C^* Q_d C^*(I - A)')(C^* Q_d C^*)^{-1}$

There fore D_2 is a positive definite matrix. Then according to lemma 2 ,Δ_2 is a non negative definite matrix if and only if $b_1'(\sigma^2 D_2 + b'_2 b_2)b_1 \leq 1$ .This completes the proof .

4. Simulation Study

The purpose of this section is to compare the many biased estimators in order to determine which one is the best. Thus, we use the MATLAB application to perform a simulation investigation. This simulation is based on variables that influence the estimator's duo characteristics and the degree of collinearity between multiple explanatory variables. The explanatory variables were generated using the equation in accordance with Kibria (2003).

$$x_{ij} = (1 - \Upsilon^2)^{\frac{1}{2}} z_{ij} + \Upsilon z_{ip} ,$$
$$i = 1 \dots \dots . . n \quad , j = 1 \dots \dots . . p \qquad (26)$$

where the z_{ij} independent standard normal pseudo-random numbers and Υ represents the correlation between any two variables. These variables are standardized so that $X'X$ is being in correlation form. The response variable y is considered by:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots \dots \dots \dots + \beta_p x_{ip} + e_i, \qquad i = 1,2, \dots \dots , n \qquad (27)$$

where e_i is i.i.d. $N(0, \sigma 2)$. Therefore, zero intercept for (27) will be assumed. Also the number of explanatory variables $p = 4$, while the values of σ are choose as (.1, 1, 5). The correlation Υ will choose as (0.85, 0.95,0.99) and sample size n as (50, 100, 150,).The coefficients $\beta_1, \beta_2, \dots, \beta_p$ are selected as the eigenvectors corresponding to the largest eigenvalue of the matrix $X'X$ subject to constraint $\beta'\beta = 1$. Thus, for all $n , \sigma , \lambda , p , \beta$ and Υ , sets of Xs are created. The experiment was replicated 5000 times by creating new error terms. Estimated mean square error (EMSE) is calculated as follows:

$$EMSE(\beta^*) = \frac{1}{10000} \sum_{i=1}^{5000} (\beta^* - \beta)'(\beta^* - \beta),$$

where β^* would be any estimator (OME,SRLTE or NME).

5. The Simulation Results

From Tables 1, 2 and 3 when the $(n = 50), (\gamma = 0.85)$ for all standard divisionσ, the SRLTE estimator is better than of the NME estimator because the SRLTE has minimum mean square error. While, when ($\gamma = 0.95$) and $(\sigma = 0.1 , 1)$ the SRLTE estimator is better than of the NME, while, when $(\sigma = 5)$ the performance of the NME estimator is better than of SRLTE and OME. In Table 3, when $(\gamma = 0.99, \sigma = 0.1 , 1,5)$ the NME has lowest mean square error. that mean , when σ increasing the MNE is better than of any estimator in this case.

Tables 3, 5, and 6 show that the SRLTE estimator performs better than the NME and OME estimators when (n=100) and (γ=0.85, 0.95,σ=0.1,1) because it has good qualities in this instance and has a minimum mean square error. Since the NME has the lowest mean square error when compared to the SRLTE and OME estimators, it performs better than the SRLTE estimator when (σ=5). The NME estimator performs better than the SRLTE and OME estimators in Table 3 (γ=0.99 σ=0.1, 1, 5).

Tables 7 through 9 show that the SRLTE estimator outperforms the NME and OME estimators when (n=150) (γ=0.85,0.95,σ=0.1) due to its good characteristics and minimum mean square error. However, the MNE outperforms all estimators when σ increases and becomes (σ=1,5). The NME estimator performs better than the SRLTE and OME estimators in Table 3 (γ=0.99 σ=0.1,1,5).

The simulation's conclusion. It can be observed that the suggested estimator performs better than the other estimators the larger the sample size and the higher the correlation coefficient.

Table 1: Estimated Mean Square Error with $n = 50,\ \gamma = 0.85$

$\sigma = 0.1$				$\sigma = 1$				$\sigma = 5$		
OME	$SRLTE$	NME	d	OME	$SRLTE$	NME	d	OME	$SRLTE$	NME
2.4081	0.2332	.9174	.10	2.0116	0.3340	١.4365	.10	0.6230	0.3937	0.4921
2.4081	0.2345	2.1410	.30	2.0116	0.3349	2.8400	.30	0.6230	0.3948	0.4935
2.4081	0.2357	1.4407	.50	2.0116	0.3358	1.9130	.50	0.6230	0.3958	0.4944
2.4081	0.2369	1.1489	.75	2.0116	0.3368	1.4723	.75	0.6230	0.3969	0.4952
2.4081	0.2375	1.0615	.80	2.0116	0.3372	1.3332	.80	0.6230	0.3975	0.4955
2.4081	0.2382	0.9954	.85	2.0116	0.3377	1.2259	.85	0.6230	0.3980	0.4958
2.4081	0.2388	0.9439	.88	2.0116	0.3381	1.1410	.88	0.6230	0.3986	0.4960
2.4081	0.2515	0.6727	.90	2.0116	0.3476	0.6784	.90	0.6230	0.4100	0.4984
2.4081	0.2647	0.6263	.95	2.0116	0.3574	0.5907	.95	0.6230	0.4222	0.4986
2.4081	0.2928	0.5968	.99	2.0116	0.3784	0.5047	.99	0.6230	0.4489	0.4965

Table 2: Estimated Mean Square Error with $n = 50,\ \gamma = 0.95$

$\sigma = 0.1$				$\sigma = 1$				$\sigma = 5$		
OME	$SRLTE$	NME	d	OME	$SRLTE$	NME	d	OME	$SRLTE$	NME
1.5651	0.3382	0.8979	.10	1.6841	0.3164	0.8399	.10	3.956095	1.416983	0.385184
1.5651	0.3395	0.6492	.30	1.6841	0.3184	0.7503	.30	3.956095	1.41585	0.431775
1.5651	0.3407	0.5927	.50	1.6841	0.3207	0.6518	.50	3.956095	1.414894	0.521684
1.5651	0.3420	0.5686	.75	1.6841	0.3231	0.5646	.75	3.956095	1.414113	0.669478
1.5651	0.3426	0.5610	.80	1.6841	0.3245	0.5310	.80	3.956095	1.413789	0.757036
1.5651	0.3433	0.5549	.85	1.6841	0.3259	0.5051	.85	3.956095	1.413509	0.839193
1.5651	0.3439	0.5500	.88	1.6841	0.3273	0.4870	.88	3.956095	1.413273	0.90208
1.5651	0.3569	0.5154	.90	1.6841	0.3684	0.6892	.90	3.956095	1.417791	0.493303
1.5651	0.3702	0.5301	.95	1.6841	0.4325	0.7766	.95	3.956095	1.439914	0.422773
1.5651	0.3980	0.6727	.99	1.6841	0.6303	0.8187	.99	3.956095	1.536971	0.395247

Table 3: Estimated Mean Square Error with $n = 50,\ \gamma = 0.99$

$\sigma = 0.1$				$\sigma = 1$				$\sigma = 5$		
OME	$SRLTE$	NME	d	OME	$SRLTE$	NME	d	OME	$SRLTE$	NME
1.6719	0.4264	0.4932	.10	1.9208	0.4843	0.7165	.10	5.5593	2.9540	0.5802
1.6719	0.4277	0.4839	.30	1.9208	0.4872	0.4551	.30	5.5593	2.9435	0.4579
1.6719	0.4290	0.4876	.50	1.9208	0.4908	0.3322	.50	5.5593	2.9330	0.4876
1.6719	0.4304	0.5177	.75	1.9208	0.4951	0.5137	.75	5.5593	2.9227	0.5832
1.6719	0.4311	0.5459	.80	1.9208	0.4976	0.6242	.80	5.5593	2.9175	0.6448
1.6719	0.4317	0.5843	.85	1.9208	0.5002	0.7274	.85	5.5593	2.9123	0.7111
1.6719	0.4324	0.6337	.88	1.9208	0.5031	0.8195	.88	5.5593	2.9072	0.7797
1.6719	0.4462	2.8066	.90	1.9208	0.5984	1.4454	.90	5.5593	2.8081	1.6019
1.6719	0.4605	3.0930	.95	1.9208	0.7679	1.5543	.95	5.5593	2.7155	1.7465
1.6719	0.4905	2.2239	.99	1.9208	1.3293	1.6199	.99	5.5593	2.5501	1.7510

Table 4: Estimated Mean Square Error with $n = 100,\ \gamma = 0.85$

$\sigma = 0.1$				$\sigma = 1$				$\sigma = 5$		
OME	$SRLTE$	NME	d	OME	$SRLTE$	NME	d	OME	$SRLTE$	NME
1.9769	0.2944	0.9963	.10	1.3152	0.2573	0.8026	.10	4.1534	2.8732	0.5023
1.9769	0.3072	0.5888	.30	1.3152	0.2711	0.5249	.30	4.1534	2.8720	0.5008
1.9769	0.2913	1.6277	.50	1.3152	0.2540	1.3091	.50	4.1534	2.8740	0.5045
1.9769	0.3203	0.4976	.75	1.3152	0.2855	0.4706	.75	4.1534	2.8739	0.5030
1.9769	0.3339	0.4391	.80	1.3152	0.3004	0.4343	.80	4.1534	2.8791	0.5080
1.9769	0.3409	0.4144	.85	1.3152	0.3081	0.4178	.85	4.1534	2.8829	0.5116
1.9769	0.3444	0.4027	.88	1.3152	0.3119	0.4098	.88	4.1534	2.8851	0.5137
1.9769	0.3480	0.3916	.90	1.3152	0.3159	0.4020	.90	4.1534	2.8874	0.5160
1.9769	0.3515	0.3808	.95	1.3152	0.3198	0.3943	.95	4.1534	2.8900	0.5185
1.9769	0.3544	0.3726	.99	1.3152	0.3230	0.3882	.99	4.1534	2.8922	0.5207

Table 5: Estimated Mean Square Error with $n = 100,\ \gamma = 0.95$

$\sigma = 0.1$				$\sigma = 1$				$\sigma = 5$		
OME	$SRLTE$	NME	d	OME	$SRLTE$	NME	d	OME	$SRLTE$	NME
1.6328	0.3478	0.5227	.10	1.209878	1.243662	0.941944	.10	2.342715	0.679662	0.306241
1.6328	0.3589	0.4014	.30	1.209878	1.156086	0.919943	.30	2.342715	0.634382	0.314621
1.6328	0.3451	0.6233	.50	1.209878	1.176757	0.929837	.50	2.342715	0.679726	0.306236
1.6328	0.3702	0.3105	.75	1.209878	1.151407	0.917122	.75	2.342715	0.715338	0.380152
1.6328	0.3818	0.2413	.80	1.209878	1.333343	0.942293	.80	2.342715	0.613471	0.334392
1.6328	0.3878	0.2190	.85	1.209878	1.074725	0.878287	.85	2.342715	0.635226	0.340796
1.6328	0.3907	0.2114	.88	1.209878	1.041954	0.867982	.88	2.342715	0.68536	0.305833
1.6328	0.3937	0.2064	.90	1.209878	1.055692	0.869392	.90	2.342715	1.141849	0.761083
1.6328	0.3967	0.2038	.95	1.209878	1.015017	0.913187	.95	2.342715	1.129965	0.781093
1.6328	0.3992	0.2036	.99	1.209878	1.02578	0.88024	.99	2.342715	1.143662	0.758309

Table 6: Estimated Mean Square Error with $n = 100,\ \gamma = 0.99$

$\sigma = 0.1$				$\sigma = 1$				$\sigma = 5$		
OME	$SRLTE$	NME	d	OME	$SRLTE$	NME	d	OME	$SRLTE$	NME

1.6216	0.4135	0.4857	.10	2.0365	0.6269	0.4790	.10	1.9819	0.6221	0.2161
1.6216	0.4145	0.4570	.30	2.0365	0.6289	0.4364	.30	1.9819	0.6212	0.2817
1.6216	0.4155	0.4240	.50	2.0365	0.6310	0.4329	.50	1.9819	0.6213	0.1258
1.6216	0.4166	0.3905	.75	2.0365	0.6331	0.5383	.75	1.9819	0.6225	0.1749
1.6216	0.4171	0.3746	.80	2.0365	0.6342	0.6٢35	.80	1.9819	0.6234	0.2233
1.6216	0.4176	0.3597	.85	2.0365	0.6352	0.٦489	.85	1.9819	0.6246	0.2741
1.6216	0.4181	0.3461	.88	2.0365	0.6363	٠.4231	.88	1.9819	0.6261	0.3235
1.6216	0.4285	0.4455	.90	2.0365	0.6609	٠.٦289	.90	1.9819	0.7094	0.8284
1.6216	0.4390	0.6928	.95	2.0365	0.6905	٠.٦131	.95	1.9819	0.8951	1.0498
1.6216	0.4609	0.7305	.99	2.0365	0.7646	٠.٦101	.99	1.9819	1.5742	1.4277

Table 7: Estimated Mean Square Error with $n = 150,\ \gamma = 0.85$

$\sigma = 0.1$				$\sigma = 1$				$\sigma = 5$		
OME	$SRLTE$	NME	d	OME	$SRLTE$	NME	d	OME	$SRLTE$	NME
1.8000	0.3096	٠.2317	.10	1.4835	0.2127	0.6760	.10	4.5696	3.2870	0.4943
1.8000	0.3235	0.7574	.30	1.4835	0.2210	0.5218	.30	4.5696	3.2755	0.4929
1.8000	0.3062	6.7363	.50	1.4835	0.2107	0.8515	.50	4.5696	3.2900	0.5085
1.8000	0.3378	0.5867	.75	1.4835	0.2297	0.4740	.75	4.5696	3.2652	0.4976
1.8000	0.3526	0.5058	.80	1.4835	0.2389	0.4380	.80	4.5696	3.2560	0.5029
1.8000	0.3601	0.4749	.85	1.4835	0.2437	0.4213	.85	4.5696	3.2518	0.5058
1.8000	0.3639	0.4607	.88	1.4835	0.2461	0.4131	.88	4.5696	3.2498	0.5074
1.8000	0.3678	0.4473	.90	1.4835	0.2486	0.4050	.90	4.5696	3.2479	0.5090
1.8000	0.3716	0.4345	.95	1.4835	0.2510	0.3970	.95	4.5696	3.2460	0.5106
1.8000	0.3747	0.4246	.99	1.4835	0.2531	0.3906	.99	4.5696	3.2446	0.5119

Table 8: Estimated Mean Square Error with $n = 150,\ \gamma = 0.95$

$\sigma = 0.1$				$\sigma = 1$				$\sigma = 5$		
OME	$SRLTE$	NME	d	OME	$SRLTE$	NME	d	OME	$SRLTE$	NME
1.4742	0.3780	0.5877	.10	4.0300	1.8499	0.5100	.10	4.9948	3.5653	0.4934
1.4742	0.3896	0.4637	.30	4.0300	1.8630	0.4694	.30	4.9948	3.5741	0.4952
1.4742	0.3752	0.7929	.50	4.0300	1.8470	0.5655	.50	4.9948	3.5640	0.4980
1.4742	0.4017	0.3892	.75	4.0300	1.8782	0.4589	.75	4.9948	3.5888	0.5188
1.4742	0.4143	0.3196	.80	4.0300	1.8956	0.4833	.80	4.9948	3.6095	0.6029
1.4742	0.4207	0.2884	.85	4.0300	1.9050	0.5153	.85	4.9948	3.6221	0.7112
1.4742	0.4240	0.2741	.88	4.0300	1.9100	0.5378	.88	4.9948	3.6289	0.8049
1.4742	0.4273	0.2608	.90	4.0300	1.9150	0.5653	.90	4.9948	3.6361	0.9473
1.4742	0.4306	0.2487	.95	4.0300	1.9202	0.5984	.95	4.9948	3.6437	1.1743
1.4742	0.4333	0.2399	.99	4.0300	1.9245	0.6293	.99	4.9948	3.6501	1.4643

Table 9: Estimated Mean Square Error with $n = 150,\ \gamma = 0.99$

$\sigma = 0.1$				$\sigma = 1$				$\sigma = 5$		
OME	$SRLTE$	NME	d	OME	$SRLTE$	NME	d	OME	$SRLTE$	NME
1.6216	0.4135	0.0164	.10	1.4194	0.7185	0.5276	.10	1.8816	1.4173	0.7963
1.6216	0.4145	0.0626	.30	1.4194	0.7275	43.6884	.30	1.8816	1.4250	0.6760
1.6216	0.4155	0.0609	.50	1.4194	0.7169	0.4858	.50	1.8816	1.4175	0.5767
1.6216	0.4166	0.1041	.75	1.4194	0.7410	٠.2699	.75	1.8816	1.4459	0.6202
1.6216	0.4171	0.1554	.80	1.4194	0.7590	٠.4932	.80	1.8816	1.4801	0.6007
1.6216	0.4176	0.1872	.85	1.4194	0.7698	٠.4831	.85	1.8816	1.5021	0.5946
1.6216	0.4181	0.2050	.88	1.4194	0.7756	٠.5145	.88	1.8816	1.5144	0.5920
1.6216	0.4285	0.2242	.90	1.4194	0.7816	٠.5633	.90	1.8816	1.5275	0.5896
1.6216	0.4390	0.2449	.95	1.4194	0.7880	٠.6262	.95	1.8816	1.5414	0.5874
1.6216	0.4609	0.2625	.99	1.4194	0.7933	٠.6852	.99	1.8816	1.5532	0.5858

6.Numerical Example

A numerical example is provided to show how well the NME estimator performs. This paper uses the acetylene data set that was employed extensively by Bashtain (2011)[18]. The scalar mean square error for the OME, SRLTE, and NME estimators are compared. Using scalar mean square error, Table 10 compares the

NME estimator's performance to that of other estimators. Using Lemma 5, we obtain Δ_i , $i = 1,2,3$ which, if and only if every eigenvalue of Δ_i is nonnegative definite, is nonnegative definite. Comparing the scalar mean square error (SRMTE, NME, and OME) is the aim. Examine the stochastic linear constraints that are employed in equation (3):

$$r = R\beta + e \ , R = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & 1 & 1 \end{pmatrix},$$

$$r = \begin{pmatrix} 63.9498 \\ 2.5648 \end{pmatrix}, \quad e \sim (0, \sigma^2 I)$$

The mse of OME estimator is given by:

$$mse\left(\hat{\beta}_{OME}\right) = \sigma^2 \sum_{i=1}^{p} \frac{1}{\lambda_i + v_i} \qquad (28)$$

The mse of SRLTE estimator is given as follows:

$$mse\left(\hat{\beta}_m\right) = \sigma^2 \sum_{i=1}^{p} \frac{\lambda_i + d^2\lambda_i^{-1} + v_i}{(\lambda_i + 1 + v_i)^2} + \sum_{i=1}^{p} \frac{((\lambda_i + d + v_i) - (\lambda_i + 1 + v_i)^{-1})^2 \alpha_i^2}{(\lambda_i + 1 + v_i)^2} \qquad (29)$$

The mse of NME estimator is given as follows :

$$mse\left(\hat{\beta}_{NME}\right) = \sigma^2 \sum_{i=1}^{p} \frac{a_i^2}{\lambda_i + v_i} + \sigma^2 \sum_{i=1}^{p} \frac{(1-a_i)^2(\lambda_i + d^2\lambda_i^{-1} + v_i)}{(\lambda_i + 1 + v_i)^2} + \sum_{i=1}^{p} \left(a_i + \frac{(1-a_i)(\lambda_i + d + v_i)}{(\lambda_i + 1 + v_i)} - 1\right)^2 \alpha_i^2 \qquad (30)$$

We obtain $\hat{\sigma}^2 = \frac{\left|(Y - X\hat{\beta})^2\right|}{n-p} = 0.0061$ and the three eigen values of $X'X$ are 2.0647, 0.8936 and 0.0417 .The $X'X$ matrix will be as follows :

$$X'X = \begin{pmatrix} 1.0000 & 0.2303 & -0.9582 \\ 0.2303 & 1.0000 & -0.2459 \\ -0.9582 & -0.2459 & 1.0000 \end{pmatrix}$$

The variables in the $X'X$ matrix struggle to have strong relationships with one another, as we can see. One benefit of standardizing the X matrix is that it makes it easier to identify the variables that have a high degree of correlation. The Condition Number $10 < C.N = \sqrt{\frac{2.0647}{0.04171}} = 7.03 < 30$ which indicates that there is a moderate multicollinearity and may be corrected.

comparing with the MURR, RRE and RLS estimators.

Table 10 : The scalar mean square error for different estimators and different estimated ridge parameter

d	OME	$SRLTE$	NME
0.30	132.4815	522.0347	647.8892
0.20	132.4815	260.2956	281.1237
0.10	132.4815	103.8709	98.7015
0.090	132.4815	94.0207	88.2613
0.080	132.4815	85.2237	79.0499
0.070	132.4815	77.4798	71.0305
0.050	132.4815	65.415	58.4392
0.040	132.4815	60.5669	53.8131
0.030	132.4815	57.0356	50.2704
0.010	132.4815	53.1324	46.3720

Table 10 illustrates how well the NME estimator performs when compared to other estimators for every different parameter d.

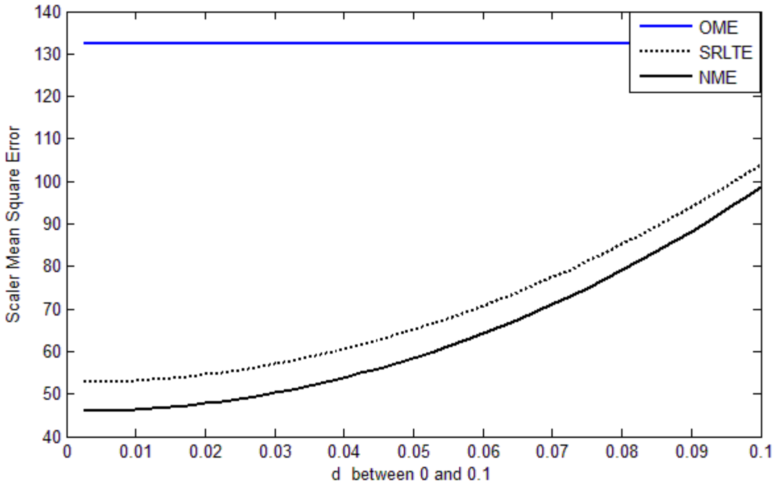

Figure 1: **The scalar mean square error for different estimated ridge parameter for OME, SRLTE and NME**

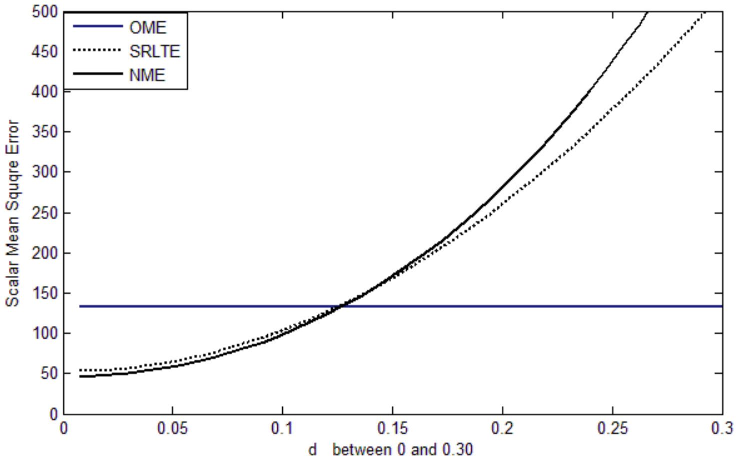

Figure 2: The scalar mean square error for different estimated ridge parameter for OME, SRLTE and NME

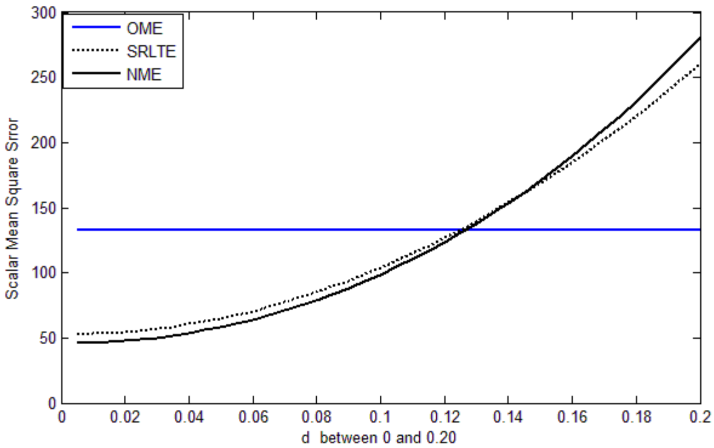

Figure 3: The scalar mean square error for different estimated ridge parameter for OME, SRLTE and NME

7. Conclusion

For the stochastic restricted linear regression model, we introduced a new stochastic restricted estimator in this work. We examined this estimator's characteristics and showed that, according to certain standards, it performs better than other biased estimators. By combining several estimators into a convex mixture framework, the suggested estimator—known as the NME—addresses current issues and produces a mean square error that is lower than that of OME and SRLTE. In addition, we conducted a simulation study to assess how multicollinearity affects the suggested estimator's performance.

References

[1] Hoerl, A. E., Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1), 55-67.

[2] Baye, M. R., Parker, D. F. (1984). Combining ridge and principal component regression: A money demand illustration. *Communications in Statistics-Theory and Methods*, 13(2), 197-205.

[3] Liu, K. (1993). A new class of biased estimate in linear regression. *Communications in Statistics-Theory and Methods*, 22(2), 393-402.

[4] Kaçıranlar, S., Sakallıoğlu, S. (2001). Combining the Liu estimator and the principal component regression estimator. *Communications in Statistics-Theory and Methods*, 30(10), 2699-2705.

[5] Liu, K. (2003). Using Liu-type estimator to combat collinearity. *Communications in Statistics-Theory and Methods*, 32(5), 1009-1020.

[6] Liu, K. (2004). More on Liu-type estimator in linear regression. *Communications in Statistics-Theory and Methods*, 33(11), 2723-2733.

[7] Akdeniz, F., Kaçıranlar, S. (1995). On the almost unbiased generalized Liu estimator and unbiased estimation of the bias and MSE. *Communications in Statistics-Theory and Methods*, 24(7), 1789-1797.

[8] Akdeniz, F., Erol, H. (2003). Mean squared error matrix comparisons of some biased estimators in linear regression. *Communications in Statistics-Theory and Methods*, 32(12), 2389-2413.

[9] Massy, W. F. (1965). Principal components regression in exploratory statistical research. *Journal of the American Statistical Association*, 60(309), 234-256.

[10] Neter, J., Wasserman, W., Kutner, M. H. (1989). Applied Linear Regression Models (Vol. 2). Irwin Press.

[11] Özbey, F., Kaçıranlar, S. (2015). Evaluation of the predictive performance of the Liu estimator. *Communications in Statistics-Theory and Methods*, 44(10), 1981-1993.

[12] Arumairajan, S., Wijekoon, P. (2017). Modified almost unbiased Liu estimator in linear regression model. *Communications in Mathematics and Statistics*, 5, 261-276.

[13] Theil, H., Goldberger, A. S. (1992). On pure and mixed statistical estimation in economics. In H. Theil's Contributions to Economics and Econometrics: Econometric Theory and Methodology (pp. 317-332). Dordrecht: Springer Netherlands.

[14] Theil, H. (1963). On the use of incomplete prior information in regression analysis. *Journal of the American Statistical Association*, 58(302), 401-414.

[15] Alheety, M. I. N. (2020). New versions of Liu-type estimator in weighted and non-weighted mixed regression model. *Baghdad Science Journal*, 17(1 Suppl.), 0361-0361.

[16] Trenkler, G., Toutenburg, H. (1990). Mean squared error matrix comparisons between biased estimators—An overview of recent results. *Statistical Papers*, 31(1), 165-179.

[17] Rao, H., Toutenburg, H., Heumann, S.C. (2008). Linear Models and Generalizations: Least Squares and Alternatives. Springer-Verlag.

[18] Bashtian, M., Hassanzadeh, M., Arashi, M., Tabatabaey, S.M.M. (2011). Using improved estimation strategies to combat multicollinearity. *Journal of Statistical Computation and Simulation*, 81(12), 1773-1797.

[19] Dawoud, I., Eledum, H. (2024). New Stochastic Restricted Biased Regression Estimators. *Mathematics*, 13(1), 15.

[20] Alheety, M.I., Nayem, H.M., Kibria, B.M.G. (2025). An Unbiased Convex Estimator Depending on Prior Information for the Classical Linear Regression Model. *Stats*, 8(1), 16.

[21] Li, Y., Yang, H. (2010). A new stochastic mixed ridge estimator in linear regression model. *Statistical Papers*, 51, 315-323.