Istanbul Journal of Mathematics
ijmath 2023, 1 (2), 74–85
DOI: 10.26650/ijmath.2023.00009

RESEARCH ARTICLE

A New biased estimator and variations based on the Kibria Lukman Estimator

K. U. Akay[1*], E. Ertan[1] and A. Erkoç[2]

[1]Istanbul University, Faculty of Science, Department of Mathematics, Vezneciler, 34134, Istanbul, Türkiye
[2]Mimar Sinan Fine Arts University, Faculty of Arts and Sciences, Department of Statistics, 34380, Şişli, Istanbul, Türkiye

ABSTRACT
One of the problems encountered in linear regression models is called multicollinearity problem which is an approximately linear relationship between the explanatory variables. This problem causes the estimated parameter values to be highly sensitive to small changes in the data. In order to reduce the impact of this problem on the model parameters, alternative biased estimators to the ordinary least squares estimator have been proposed in the literature. In this study, we propose a new biased estimator that can be an alternative to existing estimators. The superiority of this estimator over other biased estimators is analyzed in terms of matrix mean squared error. In addition, two different Monte Carlo simulation experiments are carried out to examine the performance of the biased estimators under consideration. A numerical example is given to evaluate the performance of the proposed estimator against other biased estimators.

Mathematics Subject Classification (2020): Primary 62J07, Secondary 62F10

Keywords: Liu estimator, Liu-type estimator, multicollinearity, Ridge estimator

1. INTRODUCTION

Regression analysis is one of the most widely used statistical techniques to explain the statistical relationship between explanatory and response variables using a model. Let us consider the following linear regression model:

$$Y = X\beta + \varepsilon \tag{1}$$

where Y is an $n \times 1$ vector of dependent variables, X is an $n \times p$ full column rank matrix of n observations on p independent explanatory variables, β is a $p \times 1$ vector of unknown parameters and ε is an $n \times 1$ vector of random errors which are distributed as Normal with mean vector 0 and covariance matrix $\sigma^2 I$. The Ordinary Least Squares (OLS) estimator of β is given by

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

where the covariance matrix of $\hat{\beta}_{OLS}$ is obtained as $cov\left(\hat{\beta}_{OLS}\right) = \sigma^2 (X'X)^{-1}$. According to the Gauss-Markov Theorem, the OLS estimator of the parameter vector β is the best linear unbiased estimator. In other words, we mean that $\hat{\beta}_{OLS}$ has the smallest variance among the class of all unbiased estimators that are linear combinations of the data. However, if there is an approximate relationship between the explanatory variables close to linear dependence, a biased estimator with a smaller variance may be found. This situation, i.e. a relationship close to linear dependence between explanatory variables, is called the multicollinearity problem in regression analysis. In the case of multicollinearity in the model, a very small change in the matrix X results in a very large change in matrix $(X'X)^{-1}$. Therefore, some values in the parameter vector of the OLS estimator will have a large variance. If there is multicollinearity in the linear regression model, then the OLS estimator given by (2) is again the best-unbiased estimator. However, since the variance of the OLS estimator will be very large, it will tend to produce unstable results. Although there are methods to overcome this situation by reducing the variables, alternative approaches can be used to solve the multicollinearity problem by keeping all explanatory variables in the model. Another method for solving this problem is to use biased estimators that can minimize parameter variances. For more detailed information about these proposed biased estimators in linear regression models, researchers can review the articles Hoerl and Kennard (1970),Liu (1993),Liu (2003),Kibria (2003),Özkale and Kaçıranlar (2007),Sakallıoğlu and Kaçıranlar (2008),Yang and Chang (2010),Kurnaz and Akay (2015),Kurnaz and Akay (2018),Qasim et al.

Corresponding Author: K. U. Akay **E-mail:** kulas@istanbul.edu.tr
Submitted: 17.11.2023 • **Last Revision Received:** 11.12.2023 • **Accepted:** 12.12.2023

This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

(2020),Lukman et al. (2019),Lukman et al. (2020), Ahmad and Aslam (2022), Zeinal and Azmoun (2023), Üstündağ et al. (2021), Aslam and Ahmad (2022),Babar and Chand (2022),Dawoud (2022), Qasim et al. (2022), Shewa and Ugwuowo (2023).

There are many biased estimators proposed in the literature to minimize the problems arising from collinearity. Among these estimators, the Ridge Estimator (RE) proposed by Hoerl and Kennard (1970) and the Liu Estimator (LE) proposed by Liu (1993) are widely preferred. The RE is defined by

$$\hat{\beta}_{RE} = (X'X + kI)^{-1} X'Y, \qquad k > 0 \tag{3}$$

where k is a biasing parameter. On the other hand, LE, which combines the advantages of the RE and Stein (1956) estimators, is defined as follows:

$$\hat{\beta}_{LE} = (X'X + I)^{-1} \left(X'Y + d\hat{\beta}_{OLS}\right), \quad 0 < d < 1 \tag{4}$$

where d is a biasing parameter. Stein (1956) defined the Stein estimator as follows: $\hat{\beta}_S = c\hat{\beta}_{OLS}$ where $0 < c < 1$.

However, although RE and LE are the first-choice estimators due to collinearity in the linear regression model, these estimators have several disadvantages. To utilize the advantageous features of both RE and LE, the researchers created estimators with two biasing parameters k and d. For example, Liu (2003) introduced an estimator that is dependent on k and d as follows:

$$\hat{\beta}_{LTE} = (X'X + kI)^{-1} \left(X'Y - d\hat{\beta}^*\right), \quad k > 0, \quad -\infty < d < \infty \tag{5}$$

where $\hat{\beta}^*$ can be any estimator of β. The estimator in (5) is known as the Liu-type estimator. The OLS method is used to produce this estimator after adding $\left(-d \middle/ k^{1/2}\right) \beta^* = k^{1/2}\beta + \varepsilon'$ to the model (1). As an alternative, Özkale and Kaçıranlar (2007) developed the following Two-parameter Estimator (TPE):

$$\hat{\beta}_{TPE} = (X'X + kI)^{-1} \left(X'Y + kd\hat{\beta}_{OLS}\right), \quad k > 0, \quad 0 < d < 1, \tag{6}$$

where k and d are two biasing parameters. The TPE is a general estimator which includes the OLS, RE, and LE as special cases.

Kurnaz and Akay (2015) presented a general Liu-type estimator as an alternative to the estimators previously introduced. This estimator includes estimators (2), (3), (4), (5), and (6) as special cases as follows:

$$\hat{\beta}_{NLTE} = (X'X + kI)^{-1} \left(X'Y + f(k)\hat{\beta}^*\right), \quad k > 0 \tag{7}$$

where $\hat{\beta}^*$ is any estimator of β and $f(k)$ is a continuous function of the biasing parameter k. Similarly, NLTE is obtained by augmenting $\frac{f(k)}{k^{1/2}}\hat{\beta}^* = k^{1/2}\beta + \varepsilon'$ to (1) and then using OLS method. For example, if $f(k) = -k$ and $\hat{\beta}^* = \hat{\beta}_{OLS}$, the KL estimator given by Kibria and Lukman (2020) is obtained. The KL estimator, which is a special case of the estimator (7), is defined as follows:

$$\hat{\beta}_{KL} = (X'X + kI)^{-1} (X'X - kI)\hat{\beta}_{OLS}, \quad k > 0 \tag{8}$$

where k is a biasing parameter. On the other hand, Qasim et al. (2022) proposed the Two-step shrinkage (TSS) estimator in the presence of multicollinearity as follows:

$$\hat{\beta}_{TSS} = (X'X + kI)^{-1} (X'X - kdI)\hat{\beta}_{OLS}, \qquad k > 0, \ 0 \le d < 1 \tag{9}$$

where k and d are two biasing parameters. Note that this estimator given in (9) can be obtained by taking $f(k) = -kd$ and $\hat{\beta}^* = \hat{\beta}_{OLS}$ in (7). On the other hand, when we take $f(k) = \frac{k}{d}$ where $d \in R - \{0\}$ and $\hat{\beta}^* = \hat{\beta}_{LE}$ in (7), a new two-parameter estimator proposed by Üstündağ et al. (2021) is obtained as follows:

$$\hat{\beta}_{STO} = (X'X + kI)^{-1} \left(X'Y + \frac{k}{d}\hat{\beta}_{LE}\right), \quad k > 0, \ d > 1 \tag{10}$$

where k and d are two biasing parameters. Furthermore, Sakallıoğlu and Kaçıranlar (2008) proposed another biased estimator based on RE which is given by

$$\hat{\beta}_{SK} = (X'X + I)^{-1} \left(X'Y + d\hat{\beta}_{RE}\right), \qquad k > 0, \quad -\infty < d < \infty \tag{11}$$

where k and d are two biasing parameters. This estimator given in (11) is a general estimator that includes the OLS, RE, and LEs as special cases. Also, this estimator is obtained by augmenting the equation $d\hat{\beta}_{RE} = \beta + \varepsilon'$ to (1) and using the OLS method. Also, Yang and Chang (2010) proposed a new biased estimator based on RE as follows:

$$\hat{\beta}_{YC} = (X'X + I)^{-1} (X'X + dI)\hat{\beta}_{RE}, \qquad k > 0, \quad 0 < d < 1 \tag{12}$$

where k and d are two biasing parameters. The estimator given in (12) is obtained by augmenting $(d-k)\hat{\beta}_{RE} = \beta + \varepsilon'$ to (1) and using the OLS method. Also, the YC estimator is a general estimator that includes OLS, RE, and LE as special cases.

On the other hand, Idowu et al. (2023) modified the LE provided by (4). They used the KL estimator provided by (8) in place of the OLS estimator in LE. The estimator is called LKL by Idowu et al. (2023) is given as follows:

$$\hat{\beta}_{LKL} = (X'X + I)^{-1} (X'X + dI)\hat{\beta}_{KL}, \quad k > 0, \ 0 < d < 1 \tag{13}$$

where k and d are two biasing parameters.

One of the common features of the estimators we consider is that they are defined based on Ridge, Liu, or Liu-type estimators with a modification on these estimators. Another important point here is that all estimators we have considered depend on the OLS estimator. Therefore, to reduce the problems that may arise due to collinearity, a new estimator is obtained by replacing the OLS estimator with a more powerful estimator. The estimators obtained in this case usually depend on the biasing parameters k and d.

In the literature, there are many estimators for linear regression models based on the biasing parameters k and d. Some of these estimators are as follows: LTE, SK, YC, TSS, TPE, STO, and LKL estimators. However, one of the major problems for these estimators is that it is also difficult to find optimal estimates of these biasing parameters (Liu (2003)), (Özkale and Kaçıranlar (2007)), (Sakallıoğlu and Kaçıranlar (2008)), (Yang and Chang (2010), (Ahmad and Aslam (2022)), (Aslam and Ahmad (2022)), (Qasim et al. (2022)), (Shewa and Ugwuowo (2023)). Therefore, our first objective in this study is to achieve a new estimator with a single biasing parameter by modifying the existing estimators. Another objective is to investigate the performance of this estimator with other estimators through different simulation studies.

The article is organized as follows. In Section 2, the proposed biased estimator is introduced. In Section 3, the proposed estimator is compared with the NLTE under the MMSE sense. Two Monte Carlo simulation studies are designed to evaluate the performances of the considered estimators in Section 4. In Section 5, the performance evaluation of all considered estimators is given in the Portland cement data. Finally, some conclusions are given in Section 6.

2. A NEW BIASED ESTIMATOR

In recent years, researchers have focused especially on the KL estimator proposed by Kibria and Lukman (2020). In the literature, they have proposed new estimators based on the KL estimator Dawoud (2022), Idowu et al. (2023), Shewa and Ugwuowo (2023). In this study, in order to take the performance of the KL estimator one step further, the RE estimator will be used instead of the OLS estimator in the KL estimator. In other words, the KL estimator is obtained by augmenting $-\sqrt{k}\hat{\beta}_{OLS} = \sqrt{k}\beta + \varepsilon'$ to (1) and then using the OLS method. As an alternative to this constraint, let us consider the constraint as follows: $-2\sqrt{k}\hat{\beta}_{RE} = \sqrt{k}\beta + \varepsilon'$. In this case, the estimator is obtained as follows:

$$\hat{\beta}_{KLR} = (X'X + kI)^{-1}(X'X - kI)(X'X + kI)^{-1} X'Y, \quad k > 0 \tag{14}$$

where k is a biasing parameter. This estimator given in (14) is called KLR. Let us consider the following objective function:

$$L(\beta) = (y - X\beta)'(y - X\beta) + \left((\beta - \hat{\beta}_{KLR})'(\beta - \hat{\beta}_{KLR}) - c\right) \tag{15}$$

where $\hat{\beta}_{KLR}$ is the KLR estimator given in (14). When Equation (15) is differentiated with respect to β, the following equation is obtained:

$$(X'X + I)\beta = X'Y + \hat{\beta}_{KLR}. \tag{16}$$

Solving the system given in (16) with respect to the parameter β, yields the following estimator:

$$\begin{aligned} \hat{\beta}_{LKLR} &= (X'X + I)^{-1}\left(X'Y + \hat{\beta}_{KLR}\right), \qquad k > 0 \\ \hat{\beta}_{LKLR} &= (X'X + I)^{-1}\left((X'X) + (X'X + kI)^{-1}(X'X - kI)(X'X + kI)^{-1}(X'X)\right)(X'X)^{-1}X'Y \end{aligned} \tag{17}$$

where k is a biasing parameter. We can obtain the estimator given in (17) estimator by augmenting $\hat{\beta}_{KLR} = \beta + \varepsilon'$ to model (1) and using the OLS method.

We rewrite the model (1) in canonical form

$$Y = Z\alpha + \varepsilon \tag{18}$$

where $Z = XQ$, $\alpha = Q'\beta$ and Q is the orthogonal matrix. The columns of the orthogonal matrix Q are the eigenvectors of $X'X$. Then $Z'Z = Q'X'XQ = \Lambda = diag\left(\lambda_1, \lambda_2, ..., \lambda_p\right)$ where $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_p \geq 0$ are the ordered eigenvalues of $X'X$. For model (18), we can rewrite the proposed estimators in canonical form as follows:

$$\hat{\alpha}_{LKLR} = (\Lambda + I)^{-1}\left(\Lambda + (\Lambda + kI)^{-1}(\Lambda - kI)(\Lambda + kI)^{-1}\Lambda\right)\hat{\alpha}_{OLS} \tag{19}$$

where $\hat{\alpha}_{OLS} = \Lambda^{-1}Z'y$.

We compute the biasing vector and variance-covariance matrix of the estimator $\hat{\alpha}_{LKLR}$:

$$var\left(\hat{\alpha}_{LKLR}\right) = cov\left(\left(\Lambda+I\right)^{-1}\left(\Lambda+\left(\Lambda+kI\right)^{-1}\left(\Lambda-kI\right)\left(\Lambda+kI\right)^{-1}\Lambda\right)\hat{\alpha}_{OLS}\right)$$
$$= \sigma^2\left(\Lambda+I\right)^{-1}\left(\Lambda+\left(\Lambda+kI\right)^{-1}\left(\Lambda-kI\right)\left(\Lambda+kI\right)^{-1}\Lambda\right)\Lambda^{-1}\left(\Lambda+\left(\Lambda+kI\right)^{-1}\left(\Lambda-kI\right)\left(\Lambda+kI\right)^{-1}\Lambda\right)\left(\Lambda+I\right)^{-1}$$
$$bias\left(\hat{\alpha}_{LKLR}\right) = E\left(\hat{\alpha}_{LKLR}\right)-\alpha = E\left[\left(\Lambda+I\right)^{-1}\left(\Lambda+\left(\Lambda+kI\right)^{-1}\left(\Lambda-kI\right)\left(\Lambda+kI\right)^{-1}\Lambda\right)\hat{\alpha}_{OLS}\right]-\alpha$$
$$= \left(\left(\Lambda+I\right)^{-1}\left(\Lambda+\left(\Lambda+kI\right)^{-1}\left(\Lambda-kI\right)\left(\Lambda+kI\right)^{-1}\Lambda\right)-I\right)\alpha$$

The MMSE and SMSE of an estimator $\tilde{\beta}$ are defined as:

$$\begin{aligned} MMSE\left(\tilde{\beta}\right) &= var\left(\tilde{\beta}\right)+\left[bias\left(\tilde{\beta}\right)\right]\left[bias\left(\tilde{\beta}\right)\right]' \\ SMSE\left(\tilde{\beta}\right) &= tr\left(MMSE\left(\tilde{\beta}\right)\right) = tr\left(var\left(\tilde{\beta}\right)\right)+bias\left(\tilde{\beta}\right)' bias\left(\tilde{\beta}\right). \end{aligned} \tag{20}$$

where $var\left(\tilde{\beta}\right)$ is the variance-covariance matrix and $bias\left(\tilde{\beta}\right) = E\left(\tilde{\beta}\right)-\beta$ is the biasing vector.
Let $\tilde{\beta}_1$ and $\tilde{\beta}_2$be any two estimators of parameter β. Then, $\tilde{\beta}_2$ is superior to $\tilde{\beta}_1$ with respect to the MMSE criterion if and only if $MMSE\left(\tilde{\beta}_1\right)-MMSE\left(\tilde{\beta}_2\right)$ is a positive definite (pd) matrix. If $MMSE\left(\tilde{\beta}_1\right)-MMSE\left(\tilde{\beta}_2\right)$ is a non-negative definite matrix, then $SMSE\left(\tilde{\beta}_1\right)-SMSE\left(\tilde{\beta}_2\right)\geq 0$. But, the reverse is not always true (Theobald (1974)). Because of the relation of $\alpha = Q'\beta$; $\hat{\beta}_{OLS}, \hat{\beta}_{RE}, \hat{\beta}_{LE}, \hat{\beta}_{NLTE}, \hat{\beta}_{SK}\left(k,d\right), \hat{\beta}_{YC}\left(k,d\right)$ and $\hat{\beta}_{LKLR}\left(k\right)$ have the same mean squared error values as $\hat{\alpha}_{OLS}, \hat{\alpha}_{RE}, \hat{\alpha}_{LE}$, $\hat{\alpha}_{NLTE}$, $\hat{\alpha}_{SK}\left(k,d\right)$, $\hat{\alpha}_{YC}\left(k,d\right)$, and $\hat{\alpha}_{LKLR}\left(k\right)$, respectively.
In general, the theorems used to compare the two biased estimators are given below.

Theorem 2.1. *Farebrother (2022): Let A be a positive definite matrix, namely $A>0$, and c be a nonzero vector. Then, $A-cc'>0$ if and only if $c'A^{-1}c<1$.*

Theorem 2.2. *Trenkler and Toutenburg (1990): Let $\tilde{\beta}_l = B_lY,\ \ l=1,2$ be two homogeneous linear estimators of β and C be a positive definite matrix, where $B_1B_1'-B_2B_2'$. Then $MMSE\left(\tilde{\beta}_1\right)-MMSE\left(\tilde{\beta}_2\right)>0$ iff $bias\left(\tilde{\beta}_2\right)'\left(\sigma^2C+bias\left(\tilde{\beta}_1\right)bias\left(\tilde{\beta}_1\right)'\right)^{-1}bias\left(\tilde{\beta}_2\right)<1$.*

3. SUPERIORITY OF THE PROPOSED ESTIMATOR

In this section, the proposed estimator is compared with OLS, RE, LE, and KL estimators based on the MMSE sense. However, a more general theorem is given here by considering the NLTE which includes OLS, RE, LE, and KL estimators. To compare KLKR and NLTE estimators, let us first calculate the MMSE matrices of both estimators.
The *MMSE* of $\hat{\alpha}_{NLTE} = A_1Y$ and $\hat{\alpha}_{LKLR} = A_2Y$ are given as follows:

$$\begin{aligned} & MMSE\left(\hat{\alpha}_{NLTE}\right) = var\left(\hat{\alpha}_{NLTE}\right)+bias\left(\hat{\alpha}_{NLTE}\right)bias\left(\hat{\alpha}_{NLTE}\right)' \\ & = \sigma^2A_1A_1'+\left(A_1Z-I\right)\alpha\alpha'\left(A_1Z-I\right)' \\ & = \sigma^2\left(\Lambda+kI\right)^{-1}\left(\Lambda+f\left(k\right)I\right)\Lambda^{-1}\left(\Lambda+f\left(k\right)I\right)\left(\Lambda+kI\right)^{-1} \\ & +\left(f\left(k\right)-k\right)^2\left(\Lambda+kI\right)^{-1}\alpha\alpha'\left(\Lambda+kI\right)^{-1} \end{aligned} \tag{21}$$

$$\begin{aligned} & MMSE\left(\hat{\alpha}_{LKLR}\right) = var\left(\hat{\alpha}_{LKLR}\right)+bias\left(\hat{\alpha}_{LKLR}\right)bias\left(\hat{\alpha}_{LKLR}\right)' \\ & = \sigma^2A_2A_2'+\left(A_2Z-I\right)\alpha\alpha'\left(A_2Z-I\right)' \\ & = \sigma^2\left(\Lambda+I\right)^{-1}\left(\Lambda+\left(\Lambda+kI\right)^{-1}\left(\Lambda-kI\right)\left(\Lambda+kI\right)^{-1}\Lambda\right)\Lambda^{-1}\left(\Lambda+\left(\Lambda+kI\right)^{-1}\left(\Lambda-kI\right)\left(\Lambda+kI\right)^{-1}\Lambda\right)\left(\Lambda+I\right)^{-1} \\ & +\left(\left(\Lambda+I\right)^{-1}\left(\Lambda+\left(\Lambda+kI\right)^{-1}\left(\Lambda-kI\right)\left(\Lambda+kI\right)^{-1}\Lambda\right)-I\right)\alpha\alpha'\left(\left(\Lambda+I\right)^{-1}\left(\Lambda+\left(\Lambda+kI\right)^{-1}\left(\Lambda-kI\right)\left(\Lambda+kI\right)^{-1}\Lambda\right)-I\right) \end{aligned} \tag{22}$$

Then, we can give the following theorem:

Theorem 3.1. *Let be $k>0$ and $\left|\lambda_j+f\left(k\right)\right|\left(\lambda_j+1\right)\left(\lambda_j+k\right)>\lambda_j\left|\left(\lambda_j+k\right)^2+\lambda_j-k\right|$ where $j=1,2,...,p+1$. Then, $MMSE\left(\hat{\alpha}_{NLTE}\right)-MMSE\left(\hat{\alpha}_{LKLR}\right)>0$ if and only if*

$$bias\left(\hat{\alpha}_{NLTE}\right)'\left[\sigma^2\left(A_1A_1'-A_2A_2'\right)+bias\left(\hat{\alpha}_{LKLR}\right)bias\left(\hat{\alpha}_{LKLR}\right)'\right]^{-1}bias\left(\hat{\alpha}_{NLTE}\right)<1 \tag{23}$$

where $bias\left(\hat{\alpha}_{NLTE}\right) = \left(f\left(k\right)-k\right)\left(\Lambda+kI\right)^{-1}\alpha$.

Proof. Using (21) and (22), we obtain

$$
\begin{aligned}
& var\left(\hat{\alpha}_{NLTE}\right) - var\left(\hat{\alpha}_{LKLR}\right) = \sigma^2\left[A_1A_1' - A_2A_2'\right] \\
& = \sigma^2\left[(\Lambda + kI)^{-1}\left(\Lambda + f(k)I\right)\Lambda^{-1}\left(\Lambda + f(k)I\right)(\Lambda + kI)^{-1}\right. \\
& \left. - (\Lambda + I)^{-1}\left(\Lambda + (\Lambda + kI)^{-1}(\Lambda - kI)(\Lambda + kI)^{-1}\Lambda\right)\Lambda^{-1}\left(\Lambda + (\Lambda + kI)^{-1}(\Lambda - kI)(\Lambda + kI)^{-1}\Lambda\right)(\Lambda + I)^{-1}\right] \\
& = \sigma^2\, diag\left\{\frac{\left(\lambda_j + f(k)\right)^2}{\lambda_j\left(\lambda_j + k\right)^2} - \frac{\lambda_j\left(\left(\lambda_j + k\right)^2 + \lambda_j - k\right)^2}{\left(\lambda_j + 1\right)^2\left(\lambda_j + k\right)^4}\right\}_{j=1}^{p+1}.
\end{aligned}
$$

We can observe that $A_1A_1' - A_2A_2' > 0$ if and only if $\left(\lambda_j + f(k)\right)^2\left(\lambda_j + 1\right)^2\left(\lambda_j + k\right)^2 - \lambda_j^2((\lambda_j + k)^2 + \lambda_j - k)^2 > 0$ where $j = 1, 2, ..., p+1$. Therefore, $A_1A_1' - A_2A_2'$ is the pd matrix. By Theorem 2.2, the proof is completed.

4. SELECTION OF BIASING PARAMETER

In general, the performance of estimators depends on the biasing parameters. There are many techniques for estimating biasing parameters. However, among researchers, values that can minimize the SMSE function are often suggested as estimators of the biasing parameter. Firstly, to find the optimal biasing parameter k, we take the derivative of $h(k) = SMSE\left(\hat{\beta}_{LKLR}\right)$ with respect to k where $SMSE\left(\hat{\beta}_{LKLR}\right)$ is given as follows:

$$
SMSE\left(\hat{\beta}_{LKLR}\right) = \sum_{j=1}^{p+1}\frac{\left(\lambda_j\left(\lambda_j + k\right)^2 + \left(\lambda_j - k\right)\lambda_j\right)^2\sigma^2}{\lambda_j\left(\lambda_j + 1\right)^2\left(\lambda_j + k\right)^2} + \left(\frac{\lambda_j\left(\lambda_j + k\right)^2 + \left(\lambda_j - k\right)\lambda_j}{\left(\lambda_j + 1\right)\left(\lambda_j + k\right)^2} - 1\right)^2\alpha_j^2
$$

Then, we find $h'(k)$ as follows differentiating $h(k)$ with respect to k:

$$
h'(k) = \sum_{j=1}^{p+1}\frac{2\lambda_j\left(k - 3\lambda_j\right)\left(-k\alpha_j^2\left(k + 3\lambda_j\right) + \sigma^2\left((-1 + k)k + (1 + 2k)\lambda_j + \lambda_j^2\right)\right)}{\left(1 + \lambda_j\right)^2\left(k + \lambda_j\right)^5}
$$

When it is accepted $h'(k) = 0$, we have:

$$
\begin{aligned}
k_1 &= 3\lambda_j \\
k_2 &= \frac{\sigma^2 - 2\sigma^2\lambda_j + 3\alpha_j^2\lambda_j - \sqrt{\sigma^4 - 8\sigma^4\lambda_j + 10\sigma^2\alpha_j^2\lambda_j - 8\sigma^2\alpha_j^2\lambda_j^2 + 9\alpha_j^4\lambda_j^2}}{2\left(\sigma^2 - \alpha_j^2\right)} \\
k_3 &= \frac{\sigma^2 - 2\sigma^2\lambda_j + 3\alpha_j^2\lambda_j + \sqrt{\sigma^4 - 8\sigma^4\lambda_j + 10\sigma^2\alpha_j^2\lambda_j - 8\sigma^2\alpha_j^2\lambda_j^2 + 9\alpha_j^4\lambda_j^2}}{2\left(\sigma^2 - \alpha_j^2\right)}
\end{aligned}
$$

where $i = 1, 2, ..., p+1$. Unfortunately, the k value depends on σ^2 and α_j^2. For practical purposes, we replace them with their unbiased estimators $\hat{\sigma}^2$ and $\hat{\alpha}_j^2$ to find the estimators of the biasing parameter k. Based on the simulation results, we can use the following estimators to estimate the biasing parameter k: $\hat{k}_{LKLR\ I} = \frac{3\max\left(\lambda_j\right)}{p}$, $\hat{k}_{LKLR\ II} = 3\ median\left(\lambda_j\right)$, $\hat{k}_{LKLR\ III} = \frac{\hat{\sigma}^2}{\left(\prod_{j=1}^{p+1}\hat{\alpha}_j^2\right)^{\frac{1}{p+1}}}$ where $\hat{\sigma}^2 = \frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}{n-p-1}$.

5. THE MONTE CARLO SIMULATION STUDIES

In this section, the performance of the proposed biased estimator is compared with other existing estimators using two different Monte Carlo simulation designs. In the first design, we investigated the effects of sample size (n), the degree of the collinearity (ρ), the number of explanatory variables (p), and the variance $\left(\sigma^2\right)$ on the performances of OLS, RE, LE, LTE, SK, YC, KL, TSS, STO, LKL, and LKLR estimators. In the second simulation design, we examined RE, LE, KL, and LKLR performances for each of n, p, ρ, and σ^2 values at certain values of k. For both simulation designs, we generate the explanatory variables by following McDonald and Galarneau (1975) and Kibria (2003) as

$$
x_{ij} = \left(1 - \rho^2\right)^{1/2}u_{ij} + \rho u_{ip+1}, \quad i = 1, 2, .., n, \quad j = 1, 2, ..., p \tag{24}
$$

where u_{ij} are independent standard normal pseudo-random numbers. ρ is specified so that the correlation between any two variables is given by ρ^2. These variables are standardized such that $X'X$ is a correlation matrix. Investigations are conducted on

Table 1. The EMSE values of the estimators for the model when $p = 4$.

σ^2	n	ρ	OLS	RE	LE	YCI	YCII	SK	LTE	KL	TSS	TPE	STO	LKL	LKLRI	LKLRII	LKLRIII
1	50	0.8	7.73	1.455	0.962	4.417	3.178	2.898	3.058	4.386	7.387	3.104	338.930	4.386	0.644**	**0.632***	0.883***
5	50	0.8	37.874	5.058	4.35	21.046	15.353	13.585	14.322	21.185	35.526	14.056	469.622	21.185	2.949**	**2.935***	3.802***
10	50	0.8	74.525	8.791	8.385	40.745	29.257	25.808	27.145	41.073	69.530	26.512	5596.336	41.073	5.655**	**5.630***	7.02***
1	50	0.9	17.546	2.339	0.633	9.476	6.847	6.179	6.426	9.485	16.794	6.466	22.261	9.485	0.392**	**0.275***	0.486***
5	50	0.9	84.246	7.491	2.725	44.749	32.293	28.924	29.885	44.912	78.449	29.614	4357299.439	44.912	1.764**	**1.348***	1.868***
10	50	0.9	174.597	13.952	5.543	94.22	67.023	60.561	62.517	94.497	162.223	61.774	761.113	94.497	3.588***	**2.738***	3.568**
1	50	0.95	39.226	3.87	0.44	20.961	15.063	13.596	14.051	21.005	37.523	14.045	43.351	21.005	**0.302***	0.328**	0.368***
5	50	0.95	196.097	13.716	1.914	104.112	74.761	67.173	69.443	104.3	182.348	69.066	423.743	104.3	**1.375***	1.675**	1.704***
10	50	0.95	413.446	26.2	3.705	225.051	163.637	146.813	152.39	225.5	384.367	151.422	56193.545	225.5	**2.617***	3.213***	2.964**
1	100	0.8	8.749	1.487	0.867	4.833	3.413	3.119	3.235	4.789	8.313	3.305	168.014	4.789	0.556**	**0.511***	0.73***
5	100	0.8	43.477	5.28	3.927	23.714	16.744	15.269	15.649	23.771	40.315	15.479	150.473	23.771	2.553**	**2.436***	3.121***
10	100	0.8	88.483	9.815	7.809	48.594	34.656	31.326	32.189	48.788	81.857	31.732	10580.962	48.788	5.024**	**4.784***	6.248***
1	100	0.9	18.385	2.316	0.618	9.773	6.936	6.32	6.501	9.768	17.498	6.558	39.603	9.768	0.39**	**0.272***	0.435***
5	100	0.9	91.472	8.055	2.701	48.369	34.702	31.355	32.02	48.448	84.475	31.810	859.464	48.448	1.745***	**1.286***	1.725**
10	100	0.9	188.796	15.185	5.451	102.537	72.716	66.116	67.663	102.714	174.134	67.057	2450.396	102.714	3.56***	**2.645***	3.438**
1	100	0.95	34.568	3.559	0.484	18.518	13.351	12.01	12.454	18.554	33.1	12.456	57.270	18.554	0.319**	**0.285***	0.375***
5	100	0.95	168.844	11.045	2.01	90.465	64.932	57.955	60.116	90.679	157.235	59.742	2785.722	90.679	1.392***	**1.360***	1.383**
10	100	0.95	338.382	20.712	4.089	181.26	129.603	116.701	120.421	181.718	314.091	119.547	3498.758	181.718	2.879**	2.881***	**2.853***
1	200	0.8	8.405	1.469	0.859	4.589	3.234	2.963	3.078	4.549	7.987	3.149	772.531	4.549	0.556**	**0.516***	0.739***
5	200	0.8	42.94	5.009	4.005	23.356	16.556	14.94	15.398	23.432	39.881	15.228	330.363	23.432	2.619**	**2.510***	3.195***
10	200	0.8	85.782	9.478	7.802	46.688	32.904	29.733	30.638	46.879	79.325	30.147	4292.141	46.879	5.054**	**4.846***	6.073***
1	200	0.9	16.174	2.16	0.662	8.6	6.11	5.568	5.728	8.585	15.385	5.783	41.577	8.585	0.41**	**0.294***	0.468***
5	200	0.9	84.006	7.882	2.974	45.71	32.927	29.939	30.554	45.76	77.806	30.355	465.148	45.76	1.897**	**1.405***	1.963***
10	200	0.9	167.663	14.1	5.887	91.208	65.68	59.698	60.985	91.367	154.677	60.403	611.768	91.367	3.778**	**2.858***	3.78***
1	200	0.95	28.83	3.187	0.518	15.428	11.019	10.025	10.28	15.442	27.526	10.302	74.288	15.442	0.331**	**0.259***	0.348***
5	200	0.95	139.313	10.602	2.199	73.591	52.831	47.801	48.904	73.712	128.824	48.608	3536.723	73.712	1.495***	**1.283***	1.455**
10	200	0.95	286.894	19.693	4.397	153.373	109.879	98.912	101.520	153.66	264.558	100.755	471.594	153.66	2.993***	**2.568***	2.915**

four distinct sets of correlations that correspond to $\rho = 0.8, 0.9$ and 0.95. The response variable is generated by

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_p x_{pi} + \varepsilon_i, \quad i = 1, 2, \ldots, n$$

where $\varepsilon_i \sim N\left(0, \sigma^2\right)$ and β_0 is equal to zero. The values of σ^2 are 1, 5, and 10 for various comparisons of the error term. For each set of explanatory variables, the parameter vector β is chosen as the normalized eigenvector corresponding to the largest eigenvalue of $X'X$ so that $\beta'\beta = 1$. The sample sizes n are 50, 100, and 200. The number of explanatory variables is chosen as $p = 4,\ 8,$ and 12.

For the simulation and application sections, we use the estimator proposed by Kibria (2003) to estimate the parameter k in RE, as follows: $\hat{k}_{RE} = \frac{\hat{\sigma}^2}{\left(\prod_{j=1}^{p+1} \hat{\alpha}_j^2\right)^{\frac{1}{p+1}}}$ where $\hat{\sigma}^2 = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n-p-1}$. Based on the results given by Qasim et al. (2020), we use the best estimation of d in LE as $\hat{d}_{LE} = \max\left(0, \min\left(\frac{\hat{\alpha}_j^2 - \hat{\sigma}^2}{\max\left(\frac{\hat{\sigma}^2}{\lambda_j}\right) + \hat{\alpha}_{\max}^2}\right)\right)$. Moreover, we used the best estimators and iterative techniques recommended by Liu (2003), Özkale and Kaçıranlar (2007), Sakallıoğlu and Kaçıranlar (2008), Yang and Chang (2010), Huang and Yang (2014), Kibria and Lukman (2020), Üstündağ et al. (2021), Qasim et al. (2022), Idowu et al. (2023) to determine the estimates of the biasing parameters for the LTE, SK, YC, KL, TSS, STO, TPE, and LKL estimators.

The performance of the estimated MSEs (EMSEs) is used as a basis for comparing the proposed estimators, which are calculated for an estimator $\hat{\beta}$ of β as

$$EMSE\left(\hat{\beta}\right) = \frac{1}{N}\sum_{r=1}^{N}\left(\hat{\beta}_r - \beta\right)'\left(\hat{\beta}_r - \beta\right) \tag{25}$$

where $\left(\hat{\beta}_r - \beta\right)$ is the difference between the estimated and true parameter vectors at rth replication and N is the number of replications. The experiment was repeated 2000 times for each case of n, p, σ^2, and ρ by creating response variables. The computations were performed in R programming language. The results are given in Tables 1-3 where the first, second, and third best EMSE values in each row are indicated by the signs (*), (**), and (***).

In all 81 scenarios in Tables 1-3, the proposed estimator outperformed all other available estimators according to the EMSE criterion. However, all considered estimators exhibited different behaviors in different scenarios. The following observations can be obtained from Tables 1-3:

1. When the number of variables in the model is gradually increased while keeping ρ, n, and σ^2 constant, an increase is observed in the EMSE values of all estimators in general. However, this increase is much lower in the proposed estimator.
2. When the correlation ρ between the variables in the model is increased while keeping $n, p,$ and σ^2 constant, the EMSE values of some estimators increased while the EMSE values of some estimators systematically decreased. The EMSE of the proposed estimator tends to decrease as the correlation coefficient increases.

Table 2. The EMSE values of the estimators for the model when $p = 8$.

σ^2	n	ρ	OLS	RE	LE	YCI	YCII	SK	LTE	KL	TSS	TPE	STO	LKL	LKLRI	LKLRII	LKLRIII
1	50	0.8	17.927	1.901	1.615	10.916	5.893	5.524	5.87	10.975	17.701	5.902	7751.854	10.975	0.966**	**0.950***	1.186***
5	50	0.8	89.699	7.356	7.997	53.873	29.153	27.141	28.682	54.309	87.72	28.115	56086.283	54.309	4.827**	**4.756***	5.610***
10	50	0.8	177.684	13.807	15.598	105.388	57.253	52.721	56.191	106.26	173.343	54.862	4699.803	106.26	9.382**	**9.241***	10.812***
1	50	0.9	41.839	3.719	1.065	25.14	13.327	12.367	13.389	25.312	41.371	13.231	27424.305	25.312	0.570**	**0.320***	0.576***
5	50	0.9	206.925	13.664	5.103	123.075	65.337	59.351	65.216	124.027	202.052	64.073	17058.325	124.027	2.739***	**1.531***	2.432**
10	50	0.9	412.283	25.184	10.189	245.835	130.042	117.648	129.814	247.748	401.849	127.342	10383.83	247.748	5.530***	**3.176***	4.754**
1	50	0.95	93.092	7.164	0.666	56.037	30.042	27.194	30.369	56.422	92.182	29.828	510.059	56.422	0.371**	**0.353***	0.524***
5	50	0.95	459.191	24.519	3.156	277.59	147.993	131.557	148.453	279.665	449.73	145.97	565.976	279.665	1.772**	**1.700***	1.912***
10	50	0.95	917.846	45.27	6.246	551.874	291.829	258.909	293.122	555.681	896.501	288.224	2039.893	555.681	3.512***	**3.393***	3.455**
1	100	0.8	24.214	2.32	1.384	14.402	7.547	7.074	7.554	14.48	23.922	7.546	152315.676	14.48	0.771**	**0.650***	0.866***
5	100	0.8	123.027	9.379	6.943	73.791	38.853	36.518	38.553	74.252	120.277	37.947	8893.613	74.252	3.853**	**3.249***	4.176***
10	100	0.8	244.546	16.161	13.779	146.193	76.354	71.107	75.404	147.09	238.722	74.086	6325.787	147.09	7.649**	**6.472***	8.049***
1	100	0.9	37.573	3.45	1.061	22.402	11.361	11.022	11.334	22.445	37.105	11.365	2624.188	22.445	0.561***	**0.273***	0.469**
5	100	0.9	194.263	13.369	5.364	117.333	60.514	58.269	60.176	117.596	189.324	59.737	6023.231	117.596	2.868***	**1.414***	2.148**
10	100	0.9	386.218	26.721	10.581	232.783	121.565	117.205	120.897	233.363	375.623	119.85	11982.27	233.363	5.612***	**2.715***	4.449**
1	100	0.95	69.842	5.487	0.709	41.58	20.968	20.245	20.912	41.705	69.033	20.833	553.66	41.705	0.384***	**0.191***	0.256**
5	100	0.95	359.95	22.867	3.521	218.104	112.674	108.295	112.261	218.593	351.106	111.435	706.995	218.593	1.931***	**0.975***	1.209**
10	100	0.95	716.945	40.935	6.949	431.688	223.182	213.159	222.355	432.848	697.615	220.667	1132.808	432.848	3.829***	**1.975***	2.289**
1	200	0.8	17.411	1.884	1.591	10.496	5.377	5.245	5.366	10.5	17.132	5.493	24916.018	10.5	0.912**	**0.866***	1.067***
5	200	0.8	88.456	7.555	7.965	53.407	27.722	26.929	27.416	53.526	86.118	27.25	1506.584	53.526	4.589**	**4.380***	5.231***
10	200	0.8	177.523	14.336	15.978	107.477	55.476	53.841	54.734	107.653	172.516	54.213	1215.601	107.653	9.220**	**8.804***	10.45***
1	200	0.9	40.78	3.737	1.04	24.526	12.633	12.329	12.596	24.552	40.305	12.638	55943.504	24.552	0.551***	**0.240***	0.427**
5	200	0.9	201.348	13.602	5.054	120.901	61.766	59.923	61.307	121.119	196.316	60.954	1500.696	121.119	2.710***	**1.195***	1.866**
10	200	0.9	409.144	26.535	10.183	245.851	126.485	122.861	125.571	246.342	398.247	124.691	825.275	246.342	5.428***	**2.368***	3.712**
1	200	0.95	75.512	5.908	0.703	45.297	23.2	22.203	23.202	45.427	74.645	23.093	171.987	45.427	0.384***	**0.221***	0.274**
5	200	0.95	380.004	23.253	3.476	228.965	117.798	113.149	117.327	229.502	370.325	116.474	4790.433	229.502	1.934***	**1.108***	1.223**
10	200	0.95	753.565	41.767	6.857	451.33	229.927	221.459	228.576	452.452	732.852	226.918	1004.939	452.452	3.796***	**2.139***	2.252**

Table 3. The EMSE values of the estimators for the model when $p = 12$.

σ^2	n	ρ	OLS	RE	LE	YCI	YCII	SK	LTE	KL	TSS	TPE	STO	LKL	LKLRI	LKLRII	LKLRIII
1	50	0.8	36.987	2.821	2.182	23.621	11.626	10.354	11.637	23.865	36.751	11.434	297772.445	23.865	1.206**	**1.134***	1.403***
5	50	0.8	181.743	11.791	10.837	115.281	56.311	50.591	55.959	116.44	179.451	54.33	6352.978	116.44	6.019**	**5.678***	6.724***
10	50	0.8	359.223	22.14	21.439	226.952	109.898	98.497	109.114	229.446	354.314	105.785	36819.019	229.446	11.937**	**11.268***	13.324***
1	50	0.9	101.156	6.493	1.236	64.183	30.847	26.135	31.514	64.93	100.689	30.597	25701.481	64.930	0.623***	**0.277***	0.587**
5	50	0.9	504.845	26.812	6.173	321.946	156.031	131.887	158.225	325.534	499.227	153.804	188521.021	325.534	3.121***	**1.388***	2.349**
10	50	0.9	1023.08	51.434	12.291	653.041	316.387	263.471	320.556	660.255	1010.527	311.849	5239.590	660.255	6.206***	**2.743***	4.713**
1	50	0.95	162.466	10.121	0.902	103.233	50.001	41.96	51.299	104.375	161.757	49.694	718183.886	104.375	0.459**	**0.411***	0.719***
5	50	0.95	816.49	40.696	4.435	515.358	245.943	203.666	250.67	520.759	807.093	243.701	10870.123	520.759	2.262**	**2.036***	2.762***
10	50	0.95	1634.04	81.385	9.039	1043.321	501.717	417.399	510.434	1053.802	1613.713	496.017	12131.320	1053.802	4.620**	**4.095***	5.652***
1	100	0.8	32.046	2.643	2.227	20.287	9.363	8.959	9.364	20.385	31.819	9.398	8129.905	20.385	1.192**	**1.081***	1.300***
5	100	0.8	160.081	10.991	11.079	101.783	47.11	44.972	46.948	102.207	157.865	46.311	3839.132	102.207	5.970**	**5.427***	6.343***
10	100	0.8	320.861	20.747	22.252	204.985	94.66	90.554	94.2	205.84	316.083	92.863	88772.183	205.84	11.983**	**10.897***	12.690***
1	100	0.9	80.278	5.873	1.342	51.14	23.535	22.366	23.737	51.347	79.841	23.512	46913.183	51.347	0.667***	**0.228***	0.409**
5	100	0.9	395.126	22.718	6.544	250.47	114.084	107.973	114.771	251.494	389.874	113.33	1915.609	251.494	3.266***	**1.125***	1.742**
10	100	0.9	803.981	46.594	13.376	514.05	236.623	224.291	237.897	515.851	792.452	234.817	4297.482	515.851	6.694***	**2.329***	3.744**
1	100	0.95	142.346	9.444	0.909	90.366	41.55	38.237	42.151	91.012	141.738	41.353	2048.883	91.012	0.457***	**0.306***	0.423**
5	100	0.95	709.597	38.951	4.464	448.784	208.094	188.448	210.461	452.13	701.586	206.643	946.410	452.13	2.281***	**1.593***	1.840**
10	100	0.95	1429.302	70.908	8.882	903.696	418.151	376.611	422.328	909.787	1411.304	414.951	2386.412	909.787	4.493***	**3.114***	3.381**
1	200	0.8	32.059	2.627	2.222	20.415	9.35	9.113	9.354	20.47	31.803	9.452	7463.637	20.470	1.177**	**1.039***	1.235***
5	200	0.8	158.478	10.304	10.911	100.461	45.229	44	45.137	100.726	155.998	44.809	8605.372	100.726	5.795**	**5.124***	5.953***
10	200	0.8	314.225	19.312	21.62	198.306	88.923	86.551	88.606	198.833	308.752	87.761	27821.477	198.833	11.457**	**10.126***	11.752***
1	200	0.9	71.274	5.27	1.422	45.459	20.749	19.923	20.892	45.581	70.877	20.792	485542.076	45.581	0.703***	**0.231***	0.407**
5	200	0.9	347.253	20.583	6.95	221.085	99.149	95.303	99.532	221.68	342.463	98.611	6795816.344	221.680	3.457***	**1.159***	1.822**
10	200	0.9	710.097	39.336	14.265	453.976	205.42	197.836	205.874	455.257	699.539	203.932	306580.259	455.257	7.107***	**2.401***	3.650**
1	200	0.95	134.639	8.946	0.87	85.746	38.286	36.883	38.54	85.943	133.913	38.279	2037.410	85.943	0.439***	0.319**	**0.303***
5	200	0.95	670.351	36.382	4.317	423.068	189.037	182.964	189.833	423.85	660.96	188.341	663.181	423.850	2.185***	1.566**	**1.281***
10	200	0.95	1381.718	69.995	8.804	884.071	396.959	381.753	398.948	886.431	1360.706	395.945	5310.917	886.431	4.451***	3.242**	**2.565***

Table 4. The estimated parameter values and the estimated variance values of the estimators

	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$var\,(\hat{\beta})$	$SMSE\,(\hat{\beta})$
$\hat{\beta}_{OLS}$	62.4054	1.5511	0.5102	0.1019	–0.1441	4912.0902	
$\hat{\beta}_{RE}\left(\hat{k}_{RE}=\ 1.4250\right)$	0.1003	2.1725	1.1568	0.7435	0.4882	0.0673	5.07197
$\hat{\beta}_{LE}\left(\hat{d}_{LE}=0\right)$	0.1230	2.1781	1.1552	0.7473	0.4871	0.0715	5.06501
$\hat{\beta}_{LTE}\left(\hat{k}_{LTE}=1.4250,\ \hat{d}_{LTE}=\ -0.6291\right)$	27.6066	1.8982	0.8713	0.4602	0.2091	959.5019	961.0631
$\hat{\beta}_{SK}\left(\hat{k}_{SK}=1.4250,\ \hat{d}_{SK}=\ 493.7504\right)$	26.4790	8.5996	–0.6618	5.2740	–0.7883	878.0997	2620.2491
$\hat{\beta}_{YC\ I}\ (\hat{K}_1=0.0015,\ \hat{D}_1=0.9992)$	27.6068	1.9090	0.8688	0.4680	0.2075	959.5030	961.0595
$\hat{\beta}_{YC\ II}\ (\hat{K}_2=0.0008,\ \hat{D}_2=0.7206)$	27.6067	1.9052	0.8697	0.4653	0.2080	959.5027	961.0598
$\hat{\beta}_{TSS}\left(\hat{k}_{TSS}=0.5509\times10^{-3},\ \hat{d}_{TSS}=\ 0.7920\right)$	27.6068	1.9091	0.8688	0.468	0.2075	959.5030	961.05953
$\hat{\beta}_{LKL}\left(\hat{k}_{LKL}=0.4714\times10^{-3},\ \hat{d}_{LKL}=\ 1\right)$	27.6068	1.9091	0.8688	0.468	0.2075	959.5030	961.0595
$\hat{\beta}_{TPE}\left(\hat{k}_{TPE}=37.9673,\ \hat{d}_{TPE}=\ 0.4420\right)$	27.6046	1.6898	0.9184	0.3167	0.2396	959.5464	962.9542
$\hat{\beta}_{STO}\left(\hat{k}_{STO}=29.4052,\ \hat{d}_{STO}=\ 49148.7380\right)$	62.3251	1.2323	0.5835	–0.1196	–0.0962	4900.1245	4904.0676
$\hat{\beta}_{KL}\left(\hat{k}_{KL}=0.4714\times10^{-3}\right)$	27.6068	1.9091	0.8688	0.468	0.2075	959.5030	961.0595
$\hat{\beta}_{LKLR}\left(\hat{k}_{LKLR\ I}=26805.7236\right)$	0.1230	2.1780	1.1552	0.7473	0.4872	0.0715	5.0651
$\hat{\beta}_{LKLR}\left(\hat{k}_{LKLR\ II}=2429.8562\right)$	0.1230	2.1775	1.1554	0.7471	0.4872	0.0714	5.0655
$\hat{\beta}_{LKLR}\left(\hat{k}_{LKLR\ III}=1.4250\right)$	0.0701	2.1918	1.1527	0.7574	0.4857	0.0659	5.0606

3. The impact of model variance on the performance of estimators is quite high. In scenarios where n, p, and ρ are kept constant and the variance is increased, it is observed that the EMSE values of all existing estimators, including our proposed estimator, increase. However, the dramatic increase in model variance does not significantly reduce the performance of the proposed estimator.
4. It is observed that the change in the number of observations n does not have a significant effect on the estimators. The EMSE values of all estimators, including the proposed estimator, do not change significantly when the number of observations is increased.

As a result, the proposed LKLR estimator is not significantly affected by an increase in model variance, correlation between variables, or the number of variables in the model.

In the second simulation scheme, we investigate the performances of RE, LE, KL, and LKLR for each $n,\ p,\ \rho$, and σ^2. The purpose of this simulation is to examine the performances of RE, LE, KL, and LKLR at various values of the biasing parameter k according to EMSE values given in (25). The biasing parameter k is not estimated in the second simulation scheme. Only the EMSE values obtained by increasing k values in the range $[0.1,\ 1]$ by 0.1 are compared. We only consider the cases $\rho = 0.8,\ 0.9$, $n = 50,\ \ 200$, and $p = 4,\ \ 12$, and $\sigma^2 = 1,\ \ 10$. Depending on these $n, \rho, p,$ and σ^2 values, the explanatory variables are generated according to equation (24). For every value of k, the simulation is run 2000 times. The results are collectively presented graphically in Figures 1 and 2.

Figures 1 and 2 clearly show the effects of varying the biasing parameter between 0.1 and 1 on the EMSE values of the estimators. According to Figures 1 -2 , we can obtain the following results depending on each $\left(n, \rho, p, \sigma^2\right)$.

1. The RE tends to decrease as k increases. But the decrease is lagging behind the other estimators for small values of the parameter k.
2. The EMSE values of LE have the best EMSE value at small values of the biasing parameter d, while it is observed that there is an increase with increasing values of d.
3. The EMSE values of the KL estimator first decrease and then increase as k values increase.
4. The proposed LKLR estimator has smaller EMSE values with increasing correlation between variables.

6. NUMERICAL EXAMPLE

In this section, we reconsider the dataset on Portland cement data which was analyzed by Hald (2022), Liu (2003), Sakallıoğlu and Kaçıranlar (2008), Yang and Chang (2010), Kurnaz and Akay (2018). In this data, the following four compounds are independent variables: tricalcium aluminate (x_1), tetracalcium silicate (x_2), tetracalcium alumino ferrite (x_3), and dicalcium silicate (x_4). The dependent variable y is the heat evolved in calories per gram of cement. We fit a linear regression model with an intercept to the data. Then, the eigenvalues of $X'X$ are $\lambda_1 = 44676.2059$, $\lambda_2 = 5965.4221$, $\lambda_3 = 809.9521$, $\lambda_4 = 105.4187$, and $\lambda_5 = 0.0012$. The condition number is approximately 3.66×10^7, therefore the matrix X is quite ill-conditioned.

The numerical results are summarized in Table 4. In addition, $\hat{\alpha}_{OLS}$ is substituted for α in order to calculate SMSE values. From Table 4, it can be observed that the estimated variance values and the SMSE values of LKLR I, LKLR II, and LKLR III yield appropriate results compared to other existing estimators.

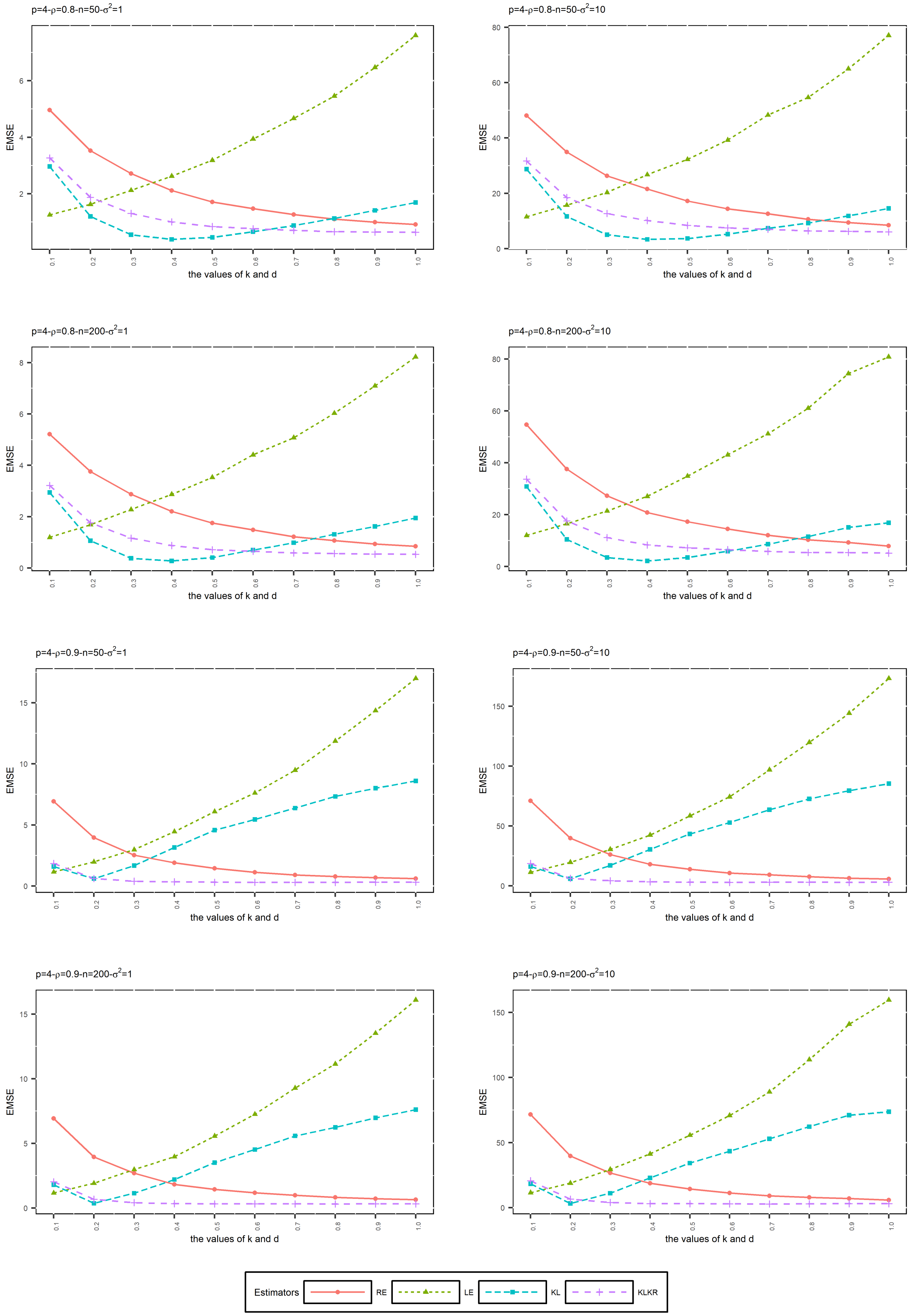

Figure 1. The EMSE values of RE, LE, KL, and LKLR as a function k and d where $p = 4$

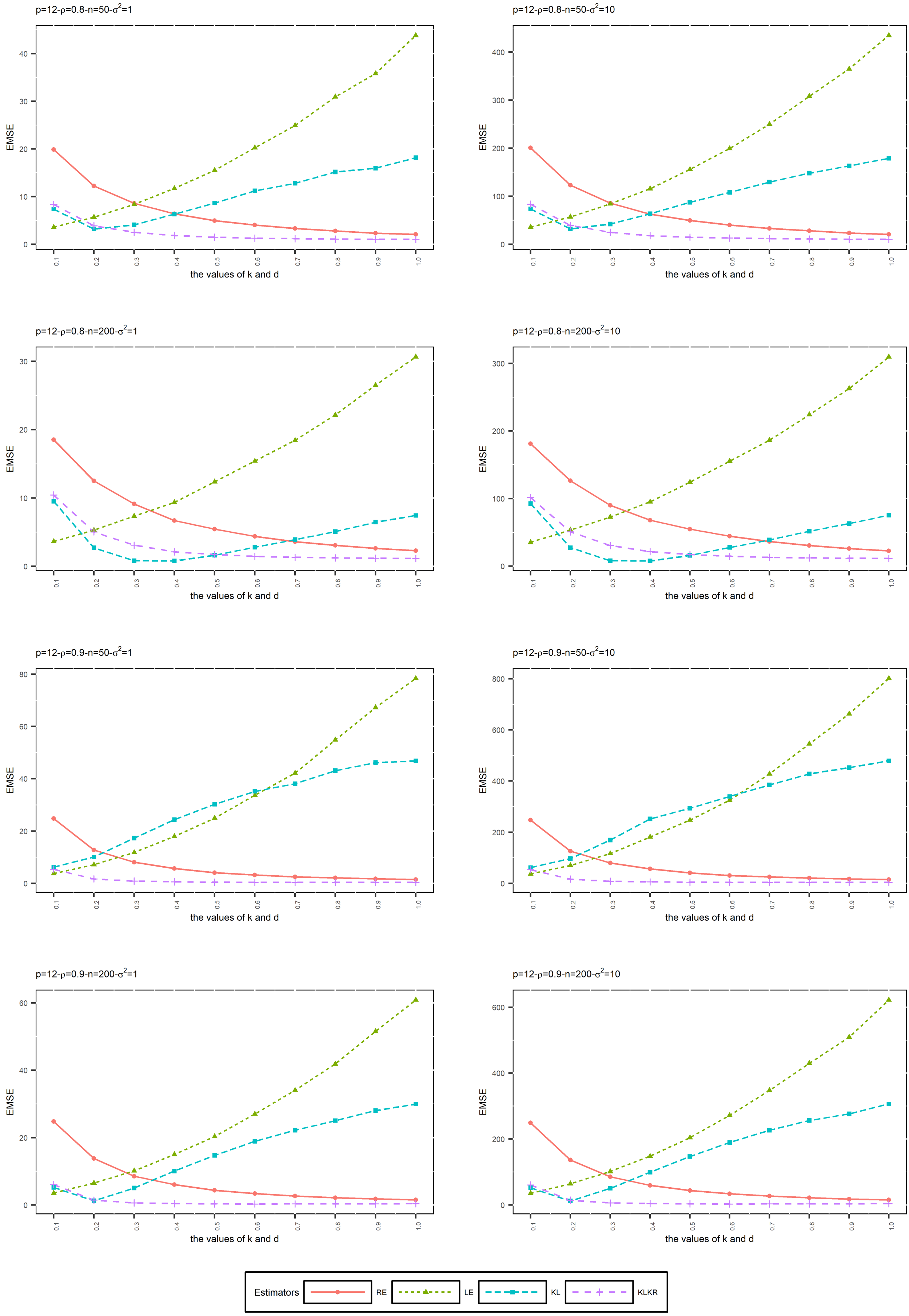

Figure 2. The EMSE values of RE, LE, KL, and LKLR as a function k and d where $p = 12$

7. CONCLUSION

In this study, a new biased estimator called LKLR is proposed in the presence of multicollinearity. This estimator has one biasing parameter as an alternative to estimators with two biasing parameters. New estimators are proposed to estimate the biasing parameter of the LKLR estimator. Simulation results show that the LKLR estimator performs better than standard estimators. In particular, $\hat{k}_{LKLR\ II}$ gave better results than other proposed biasing parameter estimators. We also examined the overall performance of other estimators with a single biasing parameter when k is in the range $[0.1, 1]$. Furthermore, the performance of the LKLR on Portland data is analyzed together with other existing estimators. Based on the results, a more robust estimator is obtained for increasing variance, variables, correlation, and number of observations than estimators with two biasing parameters. Finally, the LKLR is recommended when there is multicollinearity in the linear regression models.

Peer Review: Externally peer-reviewed.
Author Contribution: All authors have contributed equally.
Conflict of Interest: Authors declared no conflict of interest.
Financial Disclosure: Authors declared no financial support.

ACKNOWLEDGMENTS

The authors would like to thank the journal editor and the referees for their valuable comments and suggestions that contributed to the improvement of the manuscript.

LIST OF AUTHOR ORCIDS

K. U. Akay https://orcid.org/0000-0002-8668-2879
E. Ertan https://orcid.org/0000-0002-6020-8749
A. Erkoç https://orcid.org/0000-0003-4597-4282

REFERENCES

Ahmad, S., and Aslam, M. (2022). Another proposal about the new two-parameter estimator for linear regression model with correlated regressors. *Communications in Statistics-Simulation and Computation*, **51**(6), 3054–3072.

Aslam, M., and Ahmad, S. (2022). The modified Liu-ridge-type estimator: a new class of biased estimators to address multicollinearity. *Communications in Statistics-Simulation and Computation*, **51**(11), 6591–6609.

Babar, I., and Chand, S. (2022). Weighted ridge and Liu estimators for linear regression model. *Concurrency and Computation: Practice and Experience*, e7343.

Dawoud, I., Lukman, A. F., Haadi, A. R. (2022). A new biased regression estimator: Theory, simulation and application. *Scientific African*, **15**, e01100.

Farebrother, R. W. (1976).Further results on the mean square error of ridge regression. *J R Stat Soc B*, **28**, 248–250.

Hald, A. (1952).*Statistical theory with engineering applications.* Wiley, New York.

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

Huang, J., Yang, H. (2014).A two-parameter estimator in the negative binomial regression model. *Journal of Statistical Computation and Simulation*, **84**(1), 124–134.

Idowu, J. I., Oladapo, O. J., Owolabi, A. T., Ayinde, K., Akinmoju, O. (2023). Combating Multicollinearity: A New Two-Parameter Approach. *Nicel Bilimler Dergisi*, **5**(1): 90–100.

Kibria, B. G. (2003). Performance of some new ridge regression estimators. *Communications in Statistics: Simulation and Computation*, **32**(2), 419–435.

Kibria, G. B. M. and A. F. Lukman. (2020). A new ridge-type estimator for the linear regression model: Simulations and applications. *Scientifica*, 2020:9758378. doi:10.1155/2020/9758378.

Kurnaz, F. S., Akay, K. U. (2015). A new Liu-type estimator. *Stat Papers*, **56**, 495–517.

Kurnaz, F. S., Akay, K. U. (2018). Matrix mean squared error comparisons of some biased estimators with two biasing parameters. *Commun Stat Theory Methods*, **47**(8), 2022–2035.

Liu, K. (1993). A new class of biased estimate in linear regression. *Commun Stat Theory Methods*, **22**(2), 393–402.

Liu, K. (2003). Using Liu-type estimator to combat collinearity. *Commun Stat Theory Methods*, **32**(5), 1009–1020.

Lukman, A. F., Ayinde, K., Siok Kun, S., & Adewuyi, E. T. (2019). A modified new two-parameter estimator in a linear regression model. *Modelling and Simulation in Engineering*. vol. 2019, Article ID 6342702.

Lukman, A. F., Kibria, B. G., Ayinde, K., & Jegede, S. L. (2020). Modified one-parameter Liu estimator for the linear regression model. *Modelling and Simulation in Engineering*, 2020:1-17.

McDonald G.C., Galarneau D.I. (1975). A Monte Carlo evaluation of some ridge-type estimators. *J Am Stat Assoc* **70**(350):407–416.

Özkale M.R., Kaçiranlar S. (2007). The restricted and unrestricted two-parameter estimators. *Commun Stat Theory Methods* **36**(10) 2707–2725.

Qasim, M., Amin, M., Omer, T. (2020). Performance of some new Liu parameters for the linear regression model. *Communications in Statistics-Theory and Methods* **49**(17) 4178–4196.

Qasim, M., Månsson, K., Sjölander, P., Kibria, B. G. (2022). A new class of efficient and debiased two-step shrinkage estimators: method and application. *Journal of Applied Statistics*, **49**(16) 4181–4205.

Sakallioğlu S., Kaçiranlar S. (2008). A new biased estimator based on ridge estimation. *Stat Pap* **49** 4178–4196.

Shewa G. A. and Ugwuowo F. I. (2023). A new hybrid estimator for linear regression model analysis: Computations and simulations. *Scientific African*, **19** e01441.

Stein C. (1956). *Inadmissibility of the usual estimator for the mean of a multivariate normal distribution, vol. 1.* In "Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability", 197–206.

Theobald, C. M. (1974). Generalizations of mean square error applied to ridge regression. *J R Stat So B* **36** 103–106.

Trenkler G., Toutenburg H. (1990). Mean squared error matrix comparisons between biased estimator- an overview of recent results. *Stat Pap* **31** 165–179.

Üstündağ Ş. G., Toker, S., Özbay, N. (2021). Defining a two-parameter estimator: a mathematical programming evidence. *Journal of Statistical Comput. and Simul.* **91**(11) 2133–2152.

Yang H., Chang X. (2010). A new two-parameter estimator in linear regression. *Commun Stat Theory Methods* **39**(6) 923–934.

Zeinal, A., and Azmoun Zavie Kivi, M. R. (2023). The generalized new two-type parameter estimator in linear regression model. *Comm. Statist. Simulation Comput.* **52**(1) 98–109.