Unbiased Modified Two-Parameter Estimator for the Linear Regression Model

A. O. Abidoye¹*, I. M. Ajayi², F. L. Adewale², J. O. Ogunjobi²

¹Department of Statistics, University of Ilorin, Ilorin, Kwara State, Nigeria
²Department of Social and Physical Sciences, Landmark University, Omu-Aran, Kwara State, Nigeria

Received 15 February 2022, accepted in final revised form 7 June 2022

Abstract

This study centers on estimating parameters in a linear regression model in the presence of multicollinearity. Multicollinearity poses a threat to the efficiency of the Ordinary Least Squares (OLS) estimator. Some alternative estimators have been developed as remedial measures to the earlier mentioned problem. This study introduces a new unbiased modified two-parameter estimator based on prior information. Its properties are also considered; the new estimator was compared with other estimators’ Mean Square Error (MSE). A numerical example and Monte Carlo simulation were used to illustrate the performance of the new estimator.

Keywords: Linear regression model; Multicollinearity; Ordinary least squares; Unbiased modified two-parameter; Prior information.

© 2022 JSR Publications. ISSN: 2070-0237 (Print); 2070-0245 (Online). All rights reserved. doi: http://dx.doi.org/10.3329/jsr.v14i3.58234

1. Introduction

The linear regression model is expressed as:

\[ Y = X\beta + \epsilon, \]  

where \( Y \) is an \( n \times 1 \) vector of observations on the dependent variable, \( X \) is an \( n \times p \) matrix of the predictor variables, \( \beta \) is a \( p \times 1 \) vector of unknown regression coefficients, \( \epsilon \) is an \( n \times 1 \) vector of random error with \( \epsilon_i \sim N(0,\sigma^2) \).

The Ordinary Least Square (OLS) estimator of \( \beta \) is given as:

\[ \hat{\beta}_{OLS} = (X'X)^{-1}(X'Y), \]

where \( \hat{\beta}_{OLS} \) is a \( p \times 1 \) vector of unknown regression coefficients, \( X'X \) is a \( p \times p \) orthogonal matrix, \( (X'Y) \) is a \( p \times 1 \) vector and \( \hat{\beta} \sim N(\beta, \sigma^2(X'X)^{-1}) \). The OLS estimator is unbiased and possesses minimum variance among other estimators. However, one of the notable limitations of this estimator occurs when the predictor variables are highly correlated. This is termed multicollinearity, in the presence of which the OLS becomes unstable and

*Corresponding author: abidoye@unilorin.edu.ng
gives misleading regression results. Several biased estimators have been proposed in the literature to overcome the problem of multicollinearity.

Hoerl et al. [1] proposed the Ridge Regression (RR) estimator.

$$\hat{\beta}_R = (X'X + kI)^{-1}X'Y,$$  \hspace{1cm} (3)

where $k$ is the ridge parameter (or biasing constant) and $0 \leq k \leq 1$.

Amin et al. [2] proposed the modified ridge regression estimator based on prior information. The estimator is given as:

$$\hat{\beta}_{MRR}(k, b) = (X'X + kI)^{-1}(X'Y + kb),$$  \hspace{1cm} (4)

where $b$ is prior information on $\beta$. The OLS is a special case of this estimator when $k = 0$. It tends to $b$ as $k$ tends to infinity.

Liu [3] proposed an estimator to overcome the limitations of RR estimator. He combined the benefit of both the estimators given by [1] and [4]. It is given as:

$$\hat{\beta}_{Liu} = (X'X + I)^{-1}(X'X + dB)\hat{\beta}_{OLS}$$  \hspace{1cm} (5)

where $0 < d < 1$.

Dorugade [5] and Dorugade [6] introduced a modified two-parameter estimator. The estimator is given as:

$$\hat{\beta}_{MTTP}(k, d) = (X'X + kdl)^{-1}(X'X)\hat{\beta}_{OLS},$$  \hspace{1cm} (6)

where $k > 0$, $0 < d < 1$.

This estimator is a general estimator which includes the OLS and RR estimator as special cases, when $k = 0$ or $d = 0$, it gives the OLS estimator and when $d = 1$, it gives the RR estimator.

Although these estimators solve the problem of multicollinearity, they are biased estimators. Unbiased estimators have also been proposed by some researchers. The major advantage of unbiased estimators over biased estimators is that they produce unbiased estimates with minimum variance.

Crouse et al. [7] proposed an unbiased ridge estimator with prior information $J$. The estimator is defined as:

$$\hat{\beta}(k, J) = (X'X + kI)^{-1}(X'Y + kJ),$$  \hspace{1cm} (7)

where $J \sim N(\beta, (\sigma^2/k)I)$ and $J$ is uncorrelated with $\hat{\beta}_{OLS}$.

Amin et al. [8] proposed an Almost Unbiased Two-Parameter (AUTP) estimator. AUTP estimator was compared with OLS estimator and Two-Parameter (TP) estimator based on MSE criterion. The estimator is given as:

$$\hat{\beta}(k, d) = \hat{\beta}_{TP} + k(1 - d)(X'X + kI)^{-1}\hat{\beta}_{TPp},$$  \hspace{1cm} (8)

where $\hat{\beta}_{TP} = (X'X + kI)^{-1}(X'X + kdI)\hat{\beta}_{OLS}$.  \hspace{1cm} (9)

Wu [9] introduced an Unbiased Two-Parameter (UTP) estimator with prior information based on the Two-Parameter (TP) estimator was by. This is defined as follows:

$$\hat{\beta}_{UTP}(F_{kd}, J) = F_{kd}\hat{\beta}_{OLS} + (I - F_{kd})J,$$  \hspace{1cm} (10)
where $F_{kd} = (X'X + kl)^{-1}(X'X + kdI)$. 

(11)

with $J$ being uncorrelated with $\hat{\beta}_{OLS}$ and $J \sim N(\beta, (\sigma^2/(k(1-d)))(S + kdI)S - 1)$ and $S = X'X$.

In this study, a new estimator referred to as an “Unbiased Modified Two-Parameter” (UMTP) Estimator to minimize the effect of multicollinearity in a linear regression model is introduced. The article is organized as follows. In section 2, the new estimator is proposed, and its properties are obtained. The proposed estimator is compared with other two-parameter estimators using the MSE criterion in section 3.

2. Materials and Methodology

2.1. The new estimator and its properties

Dorugade [5] introduced a Modified Two-Parameter (MTP) estimator, which was earlier defined in equation (7) as $\hat{\beta}_{MTP} = (X'X + kdI)^{-1}(X'X)\hat{\beta}_{OLS}$, where $R_{kd} = (X'X + kdI)^{-1}X'X$ and $0 < k < 1, 0 < d < 1$.

Considering the convex estimator below:

$$\hat{\beta}(C, J) = C\hat{\beta}_{OLS} + (I - C)J,$$

(12)

where $C$ is a $p \times p$ matrix, $I$ is a $p \times p$ Identity matrix and $\hat{\beta}(C, J)$ is an unbiased estimator of $\beta$.

We define the new estimator, the Unbiased Modified Two-Parameter (UMTP) Estimator, based on prior information as follows:

$$\hat{\beta}(R_{kd}, J) = R_{kd}\hat{\beta}_{OLS} + (I - R_{kd})J,$$

(13)

where $R_{kd} = (X'X + kdI)^{-1}X'X$ and $J \sim N(\beta, \sigma^2(kdI)^{-1})$ for $k > 0, 0 < d < 1$.

2.1.1. Determination of the variance of $J$ ($\text{Var}(J)$)

Recall from equation (12),

$$\hat{\beta}(C, J) = C\hat{\beta}_{OLS} + (I - C)J,$$

$$\text{MSE} \{\hat{\beta}(C, J)\} = \sigma^2CS^{-1}C' + (1 - C)\text{Var}(J)(1 - C)'.$$ 

where $S = X'X$

$$\frac{\partial\text{MSE}[\hat{\beta}(C, J)]}{\partial C} = 2C\sigma^2S^{-1} - 2(I - C)\text{Var}(J) = 0.$$

$C = \text{Var}(J)[\sigma^2S^{-1} + \text{Var}(J)]^{-1}$, similarly $\text{Var}(J) = \sigma^2(I - C)^{-1}CS^{-1}$.

Note that $R_{kd}$ in the proposed estimator corresponds to $C$ in the convex estimator. Therefore, for the proposed estimator, $\text{Var}(J) = \sigma^2(I - R_{kd})^{-1}R_{kd}S^{-1} = \sigma^2(kdI)^{-1}$.

2.1.2. Unbiasedness of the proposed estimator

$$E[\hat{\beta}(R_{kd}, J)] = E[R_{kd}\beta + (I - R_{kd})J],$$

$$= E[R_{kd}\beta] + E[(I - R_{kd})J],$$

$$= (S + kdI)^{-1}SE[\beta] + (S + kdI)^{-1}(kdI)E[J],$$
Unbiased Modified Two-Parameter Estimator

\[
\hat{\beta}(R_{kd, J}) = (S + kdI)^{-1}S\beta + (S + kdI)^{-1}(kdI)\beta,
\]

\[
E[\hat{\beta}(R_{kd, J})] = \beta
\]  

(15)

2.1.3. The variance of the proposed estimator

\[
\text{Var}[\hat{\beta}(R_{kd, J})] = \text{Var}[R_{kd}\hat{\beta} + (I - R_{kd})J],
\]

\[
= \text{Var}[(S + kdI)^{-1}S\hat{\beta} + \text{Var}[(S + kdI)^{-1}kdI])],
\]

\[
= ((S + kdI)^{-1}S)\text{var}[\hat{\beta}][((S + kdI)^{-1}S) + (kdI)^2(S + kdI)^{-1}\text{var}[J](S + kdI)^{-1},
\]

\[
= S(S + kdI)^{-1}\sigma^2S^{-1}(S + kdI)^{-1}S + (kdI)^2(S + kdI)^{-1}\sigma^2(kdI)^{-1}(S + kdI)^{-1},
\]

\[
= \sigma^2S(S + kdI)^{-2} + \sigma^2(S + kdI)^{-2}(kdI)^{-1},
\]

\[
= \sigma^2(S + kdI)^{-2}(S + kdI),
\]

\[
\text{Var}[\hat{\beta}(R_{kd, J})] = \sigma^2(S + kdI)^{-1}.
\]

Thus \( \hat{\beta}(R_{kd, J}) \sim N(\beta, \sigma^2(S + kdI)^{-1}) \).  

(16)

2.1.4. Mean Squared Error (MSE) of the proposed estimator

\[
\text{MSE}[\hat{\beta}(R_{kd, J})] = \text{Var}(\hat{\beta}(R_{kd, J})) + \text{Bias}(\hat{\beta}(R_{kd, J}))^2 = \text{Var}(\hat{\beta}(R_{kd, J})),
\]

where bias \( (\hat{\beta}(R_{kd, J}))^2 = 0 \).

\[
\text{MSE}[\hat{\beta}(R_{kd, J})] = \sigma^2(S + kdI)^{-1}.
\]  

(17)

2.2. Comparison of the proposed estimator with other existing estimators based on MSE criterion

**Lemma 1:** Let \( b_1 \) and \( b_2 \) be two estimators of \( \beta \). Then \( b_2 \) is said to be MSE superior to \( b_1 \) if and only if, \( \text{MSE}(b_1) - \text{MSE}(b_2) \leq 0 \).

**Lemma 2:** Let \( M \) be a positive definite matrix, such that \( M > 0 \), and let \( \alpha \) be some vector, then \( M - \alpha \alpha' \geq 0 \) if and only if \( \alpha M^{-1} \alpha \leq 1 \).

2.2.1. Comparison between \( \hat{\beta}(R_{kd, J}) \) and \( \hat{\beta}_{OLS} \) using MSE criterion

\[
\text{MSE}(\hat{\beta}_{OLS}) = \sigma^2(X'X)^{-1} = \sigma^2S^{-1}.
\]

\[
\text{MSE}(\hat{\beta}_{OLS}) - \text{MSE}(\hat{\beta}(R_{kd, J})) = \sigma^2S^{-1} - \sigma^2(S + kdI)^{-1},
\]

\[
= \sigma^2(S^{-1} - (S + kdI)^{-1}).
\]  

(18)

(19)

which is a non-negative definite matrix for \( k > 0 \) and \( 0 < d < 1 \). Thus, according to Lemma \( \hat{\beta}(R_{kd, J}) \) is superior to \( \hat{\beta}_{OLS} \).
2.2.2. Comparison between $\hat{\beta}(R_{kd}, J)$ and $\hat{\beta}_{TP}$ using MSE criterion

**Theorem 1:** $\hat{\beta}(R_{kd}, J)$ is superior to $\hat{\beta}_{TP}$ if and only if $\frac{\sigma^2}{k^2(1-d)^2} (C_1^{-1}C_1 - C_2^{-1}C_2) < \beta \beta'$. **Proof:**

\[
\text{Bias}(\hat{\beta}_{TP}) = k(1-d)(S + kd)\beta,
\]
where $C_1 = S + kd$ and $C_2 = S + kdl$

\[
\text{Var}(\hat{\beta}_{TP}) = \sigma^2(S + kd)^{-1}(S + kdl)(S + kd)\beta
\]

\[
\text{MSE}(\hat{\beta}_{TP}) = \sigma^2C_1^{-1}C_2S^{-1}C_2^{-1}.
\]

\[
\text{MSE}(\hat{\beta}(R_{kd}, J)) < \text{MSE}(\hat{\beta}_{TP}) = \sigma^2C_1^{-1}C_2S^{-1}C_2^{-1} + k^2(1-d)^2C_1^{-1}\beta^2C_1^{-1},
\]

\[
= \sigma^2C_2^{-1} - \sigma^2C_1^{-1}C_2S^{-1}C_2^{-1} < k^2(1-d)^2C_1^{-1}\beta^2C_1^{-1},
\]

\[
= \frac{\sigma^2}{k^2(1-d)^2} (C_1^{-1}C_1 - C_2^{-1}C_2) < \beta \beta'.
\]

2.2.3. Comparison between $\hat{\beta}(R_{kd}, J)$ and $\hat{\beta}_{UTP}$ using MSE criterion

\[
\text{Var}(\hat{\beta}_{UTP}) = \sigma^2(S + kd)^{-1}S^{-1}(S + kdl).
\]

\[
\text{Var}(\hat{\beta}_{UTP}) = \sigma^2C_1^{-1}S^{-1}C_2.
\]

\[
\text{MSE}(\hat{\beta}_{UTP}) = \sigma^2C_1^{-1}S^{-1}C_2.
\]

\[
\text{MSE}(\hat{\beta}(R_{kd}, J)) - \text{MSE}(\hat{\beta}_{UTP}) = \sigma^2C_1^{-1}S^{-1}C_2 - \sigma^2C_2^{-1} > 0,
\]

which is a non-negative definite matrix for $k > 0$ and $0 < d < 1$. Thus, according to Lemma 1 $\hat{\beta}(R_{kd}, J)$ is superior to $\hat{\beta}_{UTP}$.

2.2.4. Comparison between $\hat{\beta}(R_{kd}, J)$ and $\hat{\beta}_{MTP}(k, d)$ using MSE criterion

**Theorem 2:** The estimator $\hat{\beta}(R_{kd}, J)$ is superior to $\hat{\beta}_{MTP}(k, d)$ if and only if $\beta \beta' > \frac{\sigma^2}{kd}$ with $k > 0$ and $0 < d < 1$.

**Proof:**

\[
\text{Bias}(\hat{\beta}_{MTP}(k, d)) = (R_{kd} - 1)\beta.
\]

\[
\text{Var}(\hat{\beta}_{MTP}(k, d)) = \sigma^2R_{kd}S^{-1}R_{kd}.
\]

\[
\text{MSE}(\hat{\beta}_{MTP}(k, d)) = \sigma^2R_{kd}S^{-1}R_{kd} + (R_{kd} - 1)\beta \beta'(R_{kd} - 1)'.
\]

\[
\text{MSE}(\hat{\beta}(R_{kd}, J)) - \text{MSE}(\hat{\beta}_{MTP}(k, d)) = \sigma^2(S + kd)^{-1} - \sigma^2(S + kd)^{-1}S(S + kd)\beta \beta'(S + kd)^{-1} - (S + kd)^{-1}(S + kd)\beta \beta'(S + kd)^{-1} - (S + kd)^{-1}S(S + kd)^{-1},
\]

\[
= \sigma^2(S + kd)^{-1} - \sigma^2(S + kd)^{-1}S(S + kd)^{-1} - (S + kd)^{-1}k^2d^2 \beta \beta' (S + kd)^{-1},
\]

\[
= (S + kd)^{-1}k^2d^2 \beta \beta' (S + kd)^{-1},
\]

\[
= (S + kd)^{-1}k^2d^2 \beta \beta' (S + kd)^{-1}.
\]
Unbiased Modified Two-Parameter Estimator

\[ = kd(S + kdl)^{-1}[\sigma^2 - kd|\beta'|](S + kdl)^{-1}. \]  

(30)

Since \( k > 0 \) and \( 0 < d < 1 \), by lemma 2, we obtain that \( \text{MSE}(\hat{\beta}(R_{kd}, J)) - \text{MSE}(\hat{\beta}_{\text{MTP}}(k, d)) \) is a non-negative definite matrix if and only if \( \beta' \leq \frac{\sigma^2}{kd} \). Thus, we conclude that \( \hat{\beta}(R_{kd}, J) \) is superior to \( \hat{\beta}_{\text{MTP}}(k, d) \) if and only if \( \beta' > \frac{\sigma^2}{kd} \).

2.2.5. Selection of bias parameters \( k \) and \( d \)

The bias parameters \( k \) and \( d \) are used in estimating two-parameter estimators. They are very important and crucial as they play a vital role in controlling the regression bias towards the mean of the response variable [10].

To examine the performance of the new proposed estimator over other existing estimators, the bias parameters \( k \) and \( d \) are chosen. In this study, we choose

\[ k = \frac{p\sigma^2}{\sum_i a_i^2}. \]  

(31)

and

\[ d = \sum_i \left[ \frac{\lambda_i + k(\bar{a}_i^2 + \lambda_i\bar{a}_i^2) - \lambda_i}{ka_i^2} \right]. \]  

(32)

where parameters \( k \) and \( d \) are ridge parameters and constant.

3. Results and Discussion

3.1. Numerical example and stimulation study

3.1.2. Numerical example

The proposed estimator will be illustrated using Portland cement data that exhibit multicollinearity, where multicollinearity is a statistical concept with several independent variables in a model correlated. Multicollinearity occurs when two or more independent variables are highly correlated with one another in a regression model. This means that an independent variable can be predicted from another independent variable in a regression model. Multicollinearity generally occurs when there are high correlations between two or more predictor variables. In other words, one predictor variable can be used to predict the others. This creates redundant information, skewing the results in a regression model; for example, correlated predictor variables are also called multicollinear predictors. Portland cement data set that was initially used by [12]. We computed the regression coefficient and the MSE of the proposed unbiased modified two-parameter estimator \( \hat{\beta}(R_{kd}, J) \) and also that of the following estimators \( \hat{\beta}_{\text{OLS}}, \hat{\beta}_{\text{MRR}}, \hat{\beta}_{\text{MLIU}}, \hat{\beta}_{\text{MTP}}, \hat{\beta}_{\text{TP}}, \) and \( \hat{\beta}_{\text{UTP}}. \)

Note that \( \hat{\beta}_{\text{MRR}} \) is the MSE of Modified Ridge (MRR) estimator, \( \hat{\beta}_{\text{MLIU}} \) is the MSE of Modified Liu (MLIU) estimator, \( \hat{\beta}_{\text{MTP}} \) is the MSE of the Modified Two-Parameter (MTP) estimator, \( \hat{\beta}_{\text{TP}} \) is the MSE of Two-Parameter (TP) estimator and \( \hat{\beta}_{\text{UTP}} \) is the MSE of Unbiased Two-Parameter (UTP) estimator. The result is presented in Table 1.
Table 1. Estimated regression parameter and MSE Value.

| Parameter | $\hat{\beta}_{OLS}$ | $\hat{\beta}_{MRR}$ | $\hat{\beta}_{MLIU}$ | $\hat{\beta}_{MTP}$ | $\hat{\beta}_{TP}$ | $\hat{\beta}_{UTP}$ | $\hat{\beta}(R_{ka,l})$ |
|-----------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| $\hat{\beta}_0$ | 62.4054 | 0.1531 | 0.1611 | 0.1411 | 53.0636 | 0.1629 | 0.1270 |
| $\hat{\beta}_1$ | 1.5511 | 2.2305 | 2.4105 | 2.1805 | 1.6452 | 2.0822 | 2.1806 |
| $\hat{\beta}_2$ | 0.5102 | 1.4544 | 1.5644 | 1.1544 | 0.6069 | 1.0939 | 1.1546 |
| $\hat{\beta}_3$ | 0.1019 | 0.8490 | 0.9890 | 0.8039 | 0.1988 | 0.7192 | 0.4792 |
| $\hat{\beta}_4$ | -0.1441 | 0.4786 | 0.4907 | 0.4866 | -0.0494 | 0.4603 | 0.4867 |
| MSE | 4912.09 | 3582.51 | 3612.71 | 3912.74 | 3712.21 | 4361.11 | 101.26 |

From Table 1, we observed that the estimated MSE value of the new proposed estimator ($\hat{\beta}(R_{ka,l})$) is smaller than the MSE of OLS ($\hat{\beta}_{OLS}$), Modified Ridge estimator ($\hat{\beta}_{MRR}$), Modified Liu estimator ($\hat{\beta}_{MLIU}$), Modified Two-Parameter estimator ($\hat{\beta}_{MTP}$), Two-Parameter estimator ($\hat{\beta}_{TP}$) and Unbiased Two Parameter estimator ($\hat{\beta}_{UTP}$). This implies that the new proposed estimator performs better in the presence of multicollinearity.

3.2. Simulation study

In order to investigate the performance of the proposed estimator, the MSE of the proposed estimator is compared with those of some existing estimators. The simulation process uses a linear regression model with fixed independent variables such that there exist different levels of multicollinearity among the independent variables. Considering the regression model:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \ldots + \beta_p X_{pi} + \epsilon_i$$

where $i=1, 2, \ldots, n$ and $p=3, 6$.

The independent variables were generated by the simulation process used by [2,13,14] and [15-18] as follows:

$$X_{ij} = (1 - \rho^2)^{1/2} v_{ij} + \rho v_{ip} \quad i=1, 2, \ldots, n \text{ and } j=1, 2, \ldots, p$$

where $X_{ij}$ are the generated independent variables, $\rho$ is the correlation between any two independent variables, $v_{ij}$ are random numbers from standard normal distribution and $p$ is the number of independent variables. In this study, we take $p = 3, 6$ and $\rho = 0.8, 0.9, 0.95$, and 0.99. $X_{ij}$ are standardized such that $X'X$ are used to generate the dependent variables at specified value of $n, p, \sigma$ and $\epsilon$ with $\beta = (0.8, 0.1, 0.6)$ when $p = 3$ and $\beta = (0.8, 0.1, 0.6, 0.15, 0.19, 0.05)$ when $p = 6$.

The parameter values were chosen such that $\beta \beta' = 1$, which is a common restriction in simulation studies of this type [19-25]. The data set are simulated with sample sizes $n = 20, 50, 100$ and $\sigma = 1, 5, 10$. The process is replicated 2000 times. We obtained the estimated MSE values of the following estimators OLS, MRR, MLIU, MTP, TP, UTP, and the proposed UMTP, respectively. Their respective MSE is obtained by the following computation.

$$MSE(\hat{\beta}) = \frac{1}{2000} \sum_{j=1}^{2000} \sum_{i=1}^{p} (\hat{\beta}_{ij} - \beta_i) (\hat{\beta}_{ij} - \beta_i)$$

From Table 1, we observed that the estimated MSE value of the new proposed estimator ($\hat{\beta}(R_{ka,l})$) is smaller than the MSE of OLS ($\hat{\beta}_{OLS}$), Modified Ridge estimator ($\hat{\beta}_{MRR}$), Modified Liu estimator ($\hat{\beta}_{MLIU}$), Modified Two-Parameter estimator ($\hat{\beta}_{MTP}$), Two-Parameter estimator ($\hat{\beta}_{TP}$) and Unbiased Two Parameter estimator ($\hat{\beta}_{UTP}$). This implies that the new proposed estimator performs better in the presence of multicollinearity.
where $\hat{\beta}_{ij}$ is the estimate of the $i^{th}$ parameter in the $j^{th}$ replication and $\beta_i$ is the true parameter value. The estimated MSE values for different combinations of the $n$, $p$, and $\rho$ are presented in Tables 2 and 3.

Table 2. Estimated MSE values of the new estimator and various two-parameter estimators when $p = 3$, $\beta = (0.8, 0.1, 0.6)$.

| $\rho$ | N   | 20          | 50          | 100         |
|--------|-----|-------------|-------------|-------------|
| $\sigma$ |   | 1 | 5 | 10 | 1 | 5 | 10 | 1 | 5 | 10 | 1 | 5 | 10 |
| $\hat{\beta}_{\text{OLS}}$ | 0.883 | 17.219 | 68.270 | 0.420 | 5.709 | 22.230 | 0.304 | 2.765 | 10.457 |
| $\hat{\beta}_{\text{MRR}}$ | 0.759 | 14.119 | 55.868 | 0.416 | 5.569 | 21.670 | 0.303 | 2.749 | 10.393 |
| $\hat{\beta}_{\text{MLIU}}$ | 0.880 | 17.141 | 67.959 | 0.421 | 5.700 | 22.194 | 0.304 | 2.763 | 10.449 |
| $\hat{\beta}_{\text{MTF}}$ | 0.580 | 9.581 | 37.711 | 0.384 | 4.753 | 18.405 | 0.295 | 2.543 | 9.566 |
| $\hat{\beta}_{\text{TP}}$ | 0.824 | 15.753 | 62.407 | 0.415 | 5.535 | 21.537 | 0.302 | 2.726 | 10.298 |
| $\hat{\beta}_{\text{UTP}}$ | 0.669 | 11.805 | 46.606 | 0.396 | 5.004 | 19.403 | 0.296 | 2.506 | 9.406 |
| $\beta_{(R_{kd}, J)}$ | 0.571 | 9.367 | 36.854 | 0.383 | 4.738 | 18.347 | 0.295 | 2.541 | 9.559 |

| $\rho = 0.8$ |
|-----------|
| $\rho = 0.9$ |
| $\rho = 0.95$ |
| $\rho = 0.99$ |
Table 3. Estimated MSE values of the new estimator and various two-parameter estimators when $p = 6, \beta = (0.8, 0.1, 0.6, 0.15, 0.19, 0.05)$.

| $\rho$ | N   | 20   | 50   | 100  |
|--------|-----|------|------|------|
| $\sigma$ |     |      |      |      |
| 1      | 1.517 | 33.072 | 131.680 | 0.725 | 13.297 | 52.583 | 0.393 | 4.99 | 19.356 |
| 5      | 1.279 | 27.093 | 107.779 | 0.703 | 12.748 | 50.389 | 0.392 | 4.962 | 19.244 |
| 10     | 1.511 | 32.923 | 131.085 | 0.724 | 13.269 | 52.472 | 0.493 | 4.986 | 19.341 |
| 0.9    | 0.938 | 20.459 | 73203  | 0.601 | 10.454 | 41.129 | 0.377 | 4.591 | 17.759 |
| 1.561 | 1.405 | 30.264 | 120.447 | 0.714 | 12.773 | 50.489 | 0.390 | 4.929 | 19.070 |
| 21.442| 1.105 | 22.659 | 90.006  | 0.650 | 11.334 | 44.722 | 0.377 | 4.593 | 17.396 |
| 20     | 0.922 | 18.052 | 71.573  | 0.629 | 10.381 | 40.927 | 0.375 | 4.588 | 17.746 |

| $\rho$ | N   | 0.9  | 50   | 100  |
|--------|-----|------|------|------|
| $\sigma$ |     |      |      |      |
| 1      | 2.874 | 66.991 | 267.358 | 1.175 | 24.548 | 97.588 | 0.570 | 9.421 | 37.079 |
| 5      | 1.759 | 39.02 | 155.443 | 1.063 | 21.728 | 86.308 | 0.562 | 9.221 | 36.278 |
| 10     | 2.854 | 66.508 | 265.425 | 1.172 | 24.461 | 97.239 | 0.570 | 9.407 | 37.022 |
| 0.9    | 1.127 | 22.961 | 91.181  | 0.826 | 15.772 | 62.477 | 0.520 | 7.945 | 31.173 |
| 1.561 | 2.511 | 57.911 | 231.031 | 1.109 | 22.898 | 90.986 | 0.560 | 9.154 | 36.011 |
| 21.442| 1.416 | 30.288 | 120.401 | 0.943 | 18.669 | 74.060 | 0.528 | 8.287 | 32.530 |
| 20     | 1.083 | 21.831 | 86.657  | 0.817 | 15.550 | 61.587 | 0.511 | 7.924 | 31.089 |

| $\rho$ | N   | 0.95 | 50   | 100  |
|--------|-----|------|------|------|
| $\sigma$ |     |      |      |      |
| 1      | 5.660 | 136.664 | 546.048 | 2.043 | 46.482 | 184.324 | 0.925 | 18.291 | 72.558 |
| 5      | 2.075 | 46.692 | 186.105 | 1.561 | 34.185 | 136.132 | 0.875 | 17.032 | 67.566 |
| 10     | 5.607 | 148.33 | 540.712 | 2.032 | 45.973 | 184.072 | 0.923 | 18.241 | 72.358 |
| 0.9    | 1.228 | 25.046 | 99.470  | 1.064 | 21.646 | 85.964 | 0.722 | 13.164 | 52.048 |
| 1.561 | 4.663 | 111.694 | 446.159 | 1.850 | 41.399 | 164.991 | 0.887 | 17.342 | 68.765 |
| 21.442| 1.589 | 34.329 | 136.628 | 1.309 | 27.774 | 110.477 | 0.795 | 14.942 | 59.148 |
| 20     | 1.155 | 23.187 | 92.028  | 1.037 | 20.951 | 83.184 | 0.718 | 13.052 | 51.596 |

| $\rho$ | N   | 0.99 | 50   | 100  |
|--------|-----|------|------|------|
| $\sigma$ |     |      |      |      |
| 1      | 28.462 | 706.699 | 2826.192 | 8.789 | 212.657 | 850.024 | 3.757 | 89.087 | 355.745 |
| 5      | 2.157 | 47.236 | 188.116 | 2.448 | 56.091 | 223.723 | 2.152 | 48.845 | 194.760 |
| 10     | 28.085 | 697.267 | 2788.460 | 8.699 | 251.391 | 842.960 | 4.730 | 88.494 | 353.009 |
| 0.9    | 1.178 | 21.537 | 85.167  | 1.354 | 9.0224 | 108.160 | 1.305 | 27.159 | 108.940 |
| 1.561 | 21.442 | 531.082 | 2123.695 | 7.006 | 170.277 | 680.499 | 3.244 | 76.359 | 334.352 |
| 21.442| 1.561 | 32.676 | 127.777 | 1.796 | 39.944 | 157.111 | 1.699 | 37.286 | 150.485 |
| 20     | 1.078 | 19.805 | 76.422  | 1.221 | 26.555 | 97.477 | 1.244 | 25.601 | 103.400 |

It is observed in Tables 2 and 3 that the new proposed two-parameter estimator $\hat{P}(R_{kd}, J)$ is superior to $\hat{P}_{OLS}$ and other two-parameter estimators such as $\hat{P}_{MRR}, \hat{P}_{MLIU}, \hat{P}_{MTP}, \hat{P}_{TP}, \hat{P}_{UTP}$ because it possesses minimum MSE when compared to
others. The new proposed estimator performs better for a different number of independent variables and various levels of correlation among independent variables ($\rho = 0.8, 0.9, 0.95, 0.99$). It also performs better when $n$ is small ($n = 20$) and for various combinations of variance ($\sigma^2$) of the error term. The new estimator is an unbiased estimator that overcomes the problem of multicollinearity and can be used in the place of other estimators considered in this study.

4. Conclusion

A new estimator is proposed, called the Unbiased Modified Two-Parameter (UMTP) estimator, based on prior information to minimize the effect of multicollinearity for the linear regression model. A Monte Carlo simulation study across different combinations of $d, k, n, p, \rho$ and $\sigma$ are carried out, and the MSE criterion was used to examine the performance of the new estimator over the OLS and other existing two-parameter estimators reviewed in this study. Real-life data with multicollinearity problems were also used to evaluate the performance of the new estimator. It was observed that the newly proposed Unbiased Modified Two-Parameter (UMTP) estimator performs better than the existing estimators in the presence of multicollinearity.

References

1. A. E Hoerl, R. W. Kannard, and K. F Baldwin, Commun. Stat. -Theory Meth. 4, 105 (1975). https://doi.org/10.1080/03610927508827232
2. M. Amin, M. Qasim, S. Afzal, and K. Naveed, Commun. Stat. -Simul. Comput. 1 (2020). https://doi.org/10.1080/03610919.2020.1797794
3. K. Liu, Commun. Stat. -Theory Meth. 22, 505 (1993). https://doi.org/10.1080/03610929308831034
4. A. V. Dorugade, Stat. Transition New Series 2, (2019).
5. A. V. Dorugade, Sri Lankan J. Appl. Stat. 15, 31 (2014). https://doi.org/10.4038/sljstats.v15i1.6792
6. A. V. Dorugade, J. Assoc. Arab Univ. Basic Appl. Sci. 21, 96 (2016). https://doi.org/10.1016/j.jaubs.2015.04.002
7. R. H. Crouse, C. Jin, and R. C. Hanumara, Commun. Stat. -Theory Meth. 24, 2341 (1995). https://doi.org/10.1080/03610929508831620
8. M. Amin, M. N. Akram, and A. Majid, Commun. Stat. -Simul. Comput. 1 (2021). https://doi.org/10.1080/03610919.2020.1870694
9. J. Wu, The Scientific World J. 18 (2014).
10. M. R. Ozkale and S. Kaciranlar, Commun. Stat. -Theor. Meth. 36, 2707 (2007). https://doi.org/10.1080/03610920701386877
11. A. E. Hoerl, R. W. Kennard, and K. F Baldwin, Commun. Stat. -Theory Meth. 4, (1975). https://doi.org/10.1080/03610927508827232
12. H. Woods, H. H. Steinour, and H. R. Starke, J. Indust. Eng. Chem. 24, 1207 (1932). https://doi.org/10.1021/ie50275a002
13. G. C. McDonald and D. I. Galarneau, J. Am. Stat. Associat. 70, 407 (1975). https://doi.org/10.1080/01621459.1975.10479882
14. B. M. G. Kibria, Commun. Stat. -Simul. Comput. 32, (2003).
15. M. Amin, M. Qasim, M. Amanullah, and S. Afzal, Statistical Papers 61 (2020).
16. G. Muniz and B. M. G. Kibria, Commun. Stat.-Simul. Comput. 38, 621 (2009). https://doi.org/10.1080/03610910902592838
17. K. Liu, Commun. Statistics-Theory Meth. 32, 1009 (2003). https://doi.org/10.1081/STA-120019959
18. A. F. Lukman, K. Ayinde, B. Aladeitan, and R. Bamidele, Arab J. Basic Appl. Sci. 27, 45 (2020). https://doi.org/10.1080/25765299.2019.1706799
19. A. F. Lukman, B. Aladeitan, K. Ayinde, M. R. Abonazel, J. Appl. Stat. 1 (2021).
20. A. F. Lukman, E. Adewuyi, K. Månsson, and B. M. G. Kibria, Scientific Reports 11, 3732 (2021). https://doi.org/10.1038/s41598-021-82582-w
21. A. F. Lukman, K. Ayinde, and A. S. Ajiboye, J. Modern Appl. Statist. Meth. 16, 428 (2017). https://doi.org/10.22237/jmasm/1493598240
22. M. Nauman, A. M. Amin, and M. Qasim, J. Statist. Computat. Simul. 90 (2020).
23. I. Dawoud and B. M. G. Kibria, Int. J. Clin. Biostat. Biometrics 16 (2020).
24. B. Aladeitan, A. F. Lukman, E. Davids E. H. Oranye, and G. B. M. Kibria, J. F1000 Res. 10 (2021). https://doi.org/10.12688/f1000research.54990.1
25. A. V. Dorugade, Stat. Transition New Series 20 (2019). https://doi.org/10.21307/stattrans-2019-021