Two Stage Robust Dawoud – Kibria Estimator for handling multicollinearity and outliers in the linear regression model

By

Enas Goda Mohammed

Master's researcher in the Department of Statistics,

Faculty of Commerce, Zagazig University

enasgoda98@gmail.com

Faculty of Commerce -Zagazig University

Volume 47 - Issue 1 January 2025

link: https://zcom.journals.ekb.eg/

Abstract

In the linear regression model, the least-squares (LS) estimator is commonly used to estimate regression parameters. However, LS becomes unreliable and unfavorable when the model is affected by multicollinearity and outliers simultaneously. Numerous authors have proposed various estimators to address the challenges of multicollinearity and outliers in linear regression models. This paper introduces an alternative robust regression estimator, called the Two-Stage Robust Dawoud–Kibria estimator, designed to address the two issues simultaneously. We performed theoretical comparisons, conducted simulations under different scenarios to illustrate the effectiveness of the proposed estimator. Theoretical analysis and simulation results indicate that the proposed estimator outperforms other regression estimators under certain conditions when both multicollinearity and outlier issues are present, based on the mean squared error criterion.

Key words: Two-Stage Robust Dawoud–Kibria estimator **(MMDK)**, Robust Dawoud–Kibria estimator **(MDK)**, Robust Liu **(M-Liu),** Robust Ridge **(M-Ridge)**, Robust Özkale–Kaçiranlar **(MOK)**, M-estimator **(ME)**, Non-Robust estimators, multicollinearity, outliers.

1. Introduction

A linear regression model is given as follows:

$$Y = X\beta + \varepsilon, \qquad \textbf{(1)}$$

where y is an n × 1 vector of responses; X is a popular n × p matrix of explanatory variables; β is an p × 1 vector of unknown regression parameters; and ε is an n × 1 vector of disturbances with zero mean and variance–covariance matrix given by Cov(ε) = $\sigma^2 I_n$, where I_n is an identity matrix of order n × n. The least-squares (LS) estimator of β in model (1) is given by

$$\hat{\beta} = C^{-1}\acute{x}\acute{y} \quad \text{, where } C = \acute{x}x. \qquad \textbf{(2)}$$

Outliers in a dataset can significantly compromise the reliability and efficiency of the Least Squares (LS) estimator. To address this issue, various robust regression estimators have been developed. Among these, the MM-estimator, introduced by (Yohai,1987), the M-estimator developed by (Huber, 1981) are widely used to handle outliers specifically in the y-direction of the dataset. Another challenge that can arise is multicollinearity among explanatory variables, which can greatly inflate the variances of regression parameter estimates, leading to unreliable results when using the LS estimator. To mitigate this problem, several alternative estimators have been proposed.

The most widely used of these is the Ridge Regression (RR) estimator, which is defined as follows:

$$\boldsymbol{\hat{\beta}(k) = D\hat{\beta}}, \quad \boldsymbol{k \geq 0}, \qquad \textbf{(3)}$$

where $D = \left(I_p + kC^{-1}\right)^{-1}$, and k represents the biasing parameter (Hoerl & Kennard, 1970). Since the Ridge Regression (RR) estimator is also susceptible to outliers in the y-direction, Silvapulle (1991) introduced the robust Ridge Regression (RRR) estimator, defined as follows:

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{M}}(\boldsymbol{k}) = \boldsymbol{D}\widehat{\boldsymbol{\beta}}_{\boldsymbol{M}}, \quad \boldsymbol{k} \geq \mathbf{0}, \qquad (\mathbf{4})$$

where $\hat{\beta}_M$ is the M- estimation obtained as follows:

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{M}} = \boldsymbol{min} \sum_{i=1}^{n} \boldsymbol{\rho}\left(\frac{\varepsilon_i}{s}\right) = \boldsymbol{min} \sum_{i=1}^{n} \boldsymbol{\rho}\left(\frac{y_i - \acute{x}_i\beta}{s}\right), \qquad (\mathbf{5})$$

where ρ denotes a robust criterion function, and s represents a scale parameter estimate.

Kejian (1993) proposed the Liu estimator, a different approach aimed at reducing the impact of multicollinearity, defined as follows:

$$\widehat{\boldsymbol{\beta}}(\boldsymbol{d}) = \boldsymbol{L}\widehat{\boldsymbol{\beta}}, \quad \mathbf{0} < \boldsymbol{d} < \mathbf{1}, \qquad (\mathbf{6})$$

where $L = \left[C + I_p\right]^{-1}\left[C + dI_p\right], d$ is a biasing parameter of the Liu estimator. However, the Liu estimator is sensitive to outliers in the y-direction. In this respect, Arslan and Billor (2000) introduced the robust Liu (RL) estimator, which is defined as follows:

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{M}}(\boldsymbol{d}) = \boldsymbol{L}\widehat{\boldsymbol{\beta}}_{\boldsymbol{M}} \quad \mathbf{0} < \boldsymbol{d} < \mathbf{1}. \qquad (\mathbf{7})$$

Özkale and Kaçiranlar (2007) proposed a two-parameter estimator that encompasses OLS, ridge regression, Liu, and contraction estimators as special cases. The effectiveness of this estimator is determined by the biasing parameters (k) and (d). They defined their own two parameter as follows:

$$\widehat{\boldsymbol{\beta}}(\boldsymbol{OK}) = \boldsymbol{R}\widehat{\boldsymbol{\beta}}, \qquad (\mathbf{8})$$

where $R = \left(C + kI_p\right)^{-1}\left(C + kdI_p\right)$.

However, the OK estimator is sensitive to outliers in the y-direction. thus (Awwad et al., 2021) introduced MOK estimator defined as follows:

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{M}}(\boldsymbol{OK}) = \boldsymbol{R}\widehat{\boldsymbol{\beta}}_{\boldsymbol{M}}. \qquad (\mathbf{9})$$

To address the issue of multicollinearity, Dawoud and Kibria (2020) introduced a regression estimator known as the Dawoud–Kibria (DK) estimator, defined as follows:

$$\boldsymbol{\widehat{\beta}(DK) = F\widehat{\beta}}, \qquad (\mathbf{10})$$

where $F = \left[C + k(1+d)I_p\right]^{-1}\left[C - k(1+d)I_p\right]$.

However, the DK estimator is sensitive to outliers in the y-direction. Thus, Dawoud and Abonazel (2021) introduced another estimator, the MDK, which is defined as follows:

$$\boldsymbol{\widehat{\beta}_M(DK) = F\widehat{\beta}_M}. \qquad (\mathbf{11})$$

The rest of the article is organized as follows. Section 2 covers the statistical methodology. In Section 3, we discuss the simulation study setup and discussions of the results. Finally, Section 4 concludes the article.

2. statistical methodology

2.1. Alternative two-stage robust estimator

For solving the multicollinearity problem, Dawoud & Kibria, (2020) introduced an alternative regression estimator called the (DK) estimator, as defined in equation (10). However, the existence of outliers in the y-direction affects the DK estimator, which shrinks the LS estimator using the matrix F. Therefore, a two-stage robust Dawoud-Kibria (MMDK) estimator, donates as $\hat{\beta}_{MM}(DK)$, is proposed which is obtained by using $\hat{\beta}_{MM}$ instead of $\hat{\beta}$. The proposed MMDK estimator of β defined as

$$\boldsymbol{\widehat{\beta}_{MM}(DK) = F\widehat{\beta}_{MM}}. \qquad (\mathbf{12})$$

The (MM-DK) estimator is proposed for estimating the parameters of a regression model with both multicollinearity and non-normality problems.

2.1.1 The Steps of the MM-DK Estimator

Step 1

Consider the general linear regression model in equation (1) with the canonical form as follows:

$$\boldsymbol{y_i = h_{ij}\alpha_j + \varepsilon_i, \qquad i = 1, 2 \dots, n} \qquad \textbf{(13)}$$

y_i represents the observed dependent variable, h_{ij} denotes the elements of the matrix H, α_j are the coefficients, and ε_i represents the error term.

Where $H = \acute{X}T, \quad \alpha = \acute{T}\beta, \; T(\acute{X}X)\acute{T} = \acute{H}H = A = \Lambda$

The DK estimator after the transformation the linear model is defined as:

$$\hat{\alpha}(DK) = \left[A + k(1+d)I_p\right]^{-1}\left[A - k(1+d)I_p\right]\hat{\alpha}. \qquad \textbf{(14)}$$

Step 2:

MM-estimation is a special type of M-estimation developed by Yohai. MM-estimators combine the high asymptotic relative efficiency of M-estimators with the high breakdown of a class of estimators called S-estimators. It was among the first robust estimators to have these two properties simultaneously. The 'MM' refers to the fact that multiple M-estimation procedures are carried out in the computation of the estimator (Lukman et al., 2015). The MM-estimator is defined as:

$$\hat{\alpha}_{MM} = \sum_{i=1}^{n} x_{ij}\, \rho\left(\frac{y_i - \acute{h}_i\alpha}{s_n}\right) . \qquad \textbf{(15)}$$

Where ρ denotes a robust criterion function, and s represents a scale parameter estimate.

Step 3:

A two-stage Robust Dawoud Kibria estimator is proposed to address the simultaneous presence of multicollinearity and outliers. The new class of (MMDK) estimator for α by minimizing $(y-h\alpha)'(y-h\alpha)$, subject to $(\alpha+\hat{\alpha})'(\alpha+\hat{\alpha})$= c

where c is a constant,

$$(y-h\alpha)'(y-h\alpha)+k(1+d)\left[(\alpha+\hat{\alpha})'(\alpha+\hat{\alpha})-c\right].$$

In this case, k and 1 + d are the Lagrange multipliers. The solution provides the proposed MMDK estimator in the following format:

$$\hat{\alpha}_{MM}(DK)=\left(h'h+k(1+d)I_p\right)^{-1}\left(h'h-k(1+d)I_p\right)\hat{\alpha}_{MM},$$

$$=\left[A+k(1+d)I_p\right]^{-1}\left[A-k(1+d)I_p\right]\hat{\alpha}_{MM}. \qquad \mathbf{(16)}$$

Additionally, after transforming the linear model, the other estimators mentioned above are as follows:

$$\hat{\alpha}_{LS}=A^{-1}H'y, \qquad \mathbf{(17)}$$

$$\hat{\alpha}_{M}=min\sum_{i=1}^{n}\rho\left(\frac{\varepsilon_i}{s}\right)=min\sum_{i=1}^{n}\rho\left(\frac{y_i-h'_i\alpha}{s}\right), \qquad \mathbf{(18)}$$

$$\alpha_M(k)=\left[I_p+kA^{-1}\right]^{-1}\hat{\alpha}_M, \qquad \mathbf{(19)}$$

$$\hat{\alpha}_M(d)=\left[A+I_p\right]^{-1}\left[A+dI_p\right]\hat{\alpha}_M, \qquad \mathbf{(20)}$$

$$\hat{\alpha}_M(OK)=\left[A+KI_p\right]^{-1}\left[A+kdI_p\right]\hat{\alpha}_M, \qquad \mathbf{(21)}$$

$$\hat{\alpha}_M(DK)=\left[A+k(1+d)I_p\right]^{-1}\left[A-k(1+d)I_p\right]\hat{\alpha}_M \qquad \mathbf{(22)}$$

2.2. Properties of the estimators

2.2.1. The mean squared errors (MSEs)

Dawoud and Abonazel (2021) defined the mean squared error (MSE) for an estimator $\bar{\alpha}$ as:

$$\text{MSE}(\bar{\alpha}) = E(\bar{\alpha} - \alpha)'(\bar{\alpha} - \alpha) = tr\big(Cov\,(\bar{\alpha})\big) + bias(\bar{\alpha})'\,bias(\bar{\alpha}) \quad \textbf{(23)}$$

The (MSEs) for the robust estimators, including M, M-Ridge, M-Liu, MOK, MDK, and MMDK, are defined as follows:

$$\text{MSE}\,(\hat{\alpha}_M) = \sum_{j=1}^{p} \Omega_{jj}, \quad \textbf{(24)}$$

where Ω_{jj} represents the diagonal elements for $Cov(\hat{\alpha}_M) = \Omega$ that is finite.

$$\text{MSE}\big(\hat{\alpha}_M(k)\big) = \sum_{j=1}^{p} \frac{a_j^2\Omega_{jj} + k^2\alpha_j^2}{\left(a_j + k\right)^2} \quad (\mathbf{25})$$

$$\text{MSE}\big(\hat{\alpha}_M(d)\big) = \sum_{j=1}^{p} \frac{\left(a_j + d\right)^2\Omega_{jj} + (1-d)^2\alpha_j^2}{\left(a_j + 1\right)^2} \quad (\mathbf{26})$$

$$\text{MSE}\big(\hat{\alpha}_M(OK)\big) = \sum_{j=1}^{p} \frac{\left(a_j + kd\right)^2\Omega_{jj} + k^2(1-d)^2\alpha_j^2}{\left(a_j + k\right)^2} \quad (\mathbf{27})$$

$$\text{MSE}(\hat{\alpha}_M(DK)) = \sum_{j=1}^{p} \frac{\left(a_j - k(1+d)\right)^2\Omega_{jj} + 4k^2(1+d)^2\alpha_j^2}{\left(a_j + k(1+d)\right)^2} \quad \textbf{(28)}$$

Updating the last equation of robust Dawoud Kibria MSE, we can obtain the MSE of the Two-Stage Robust Dawoud-Kibria Estimator (MM-DK).

$$MSE\big(\hat{\alpha}_{MM}(DK)\big) = \sum_{j=1}^{p} \frac{\left(a_j - k(1+d)\right)^2\Omega_{jj(MM)} + 4k^2(1+d)^2\alpha_j^2}{\left(a_j + k(1+d)\right)^2}\,. \quad \textbf{(29)}$$

Where $\Omega_{jj(MM)}$ represents the diagonal elements $cov(\hat{\alpha}_{MM}) = \Omega$ that is finite.

Additionally, the (MSEs) of the non-robust estimators, including OLS, Ridge, Liu, OK, and DK, are defined as follows:

$$\text{MSE}(\hat{\alpha}) = \sum_{j=1}^{p} \frac{\sigma^2}{a_j}, \quad (30)$$

$$\text{MSE } \hat{\alpha}(k) = \sum_{j=1}^{p} \frac{a_j\sigma^2 + k^2\alpha_j^2}{\left(a_j + k\right)^2} \quad (31)$$

$$\text{MSE } \hat{\alpha}(d) = \sum_{j=1}^{p} \frac{\left(a_j + d\right)^2\sigma^2 + (1-d)^2 a_j\alpha_j^2}{a_j\left(a_j + 1\right)^2} \quad (32)$$

$$\text{MSE } \hat{\alpha}(OK) = \sum_{j=1}^{p} \frac{\left(a_j + kd\right)^2\sigma^2 + k^2(1-d)^2 a_j\alpha_j^2}{a_j\left(a_j + k\right)^2} \quad (33)$$

$$\text{MSE } \hat{\alpha}(DK) = \sum_{j=1}^{p} \frac{\left(a_j - k(1+d)\right)^2\sigma^2 + 4k^2(1+d)^2 a_j\alpha_j^2}{a_j\left(a_j + k(1+d)\right)^2} \quad (34)$$

2.2.2 Theorems of the two-stage robust Dawoud-Kibria (MM-DK) estimator:

To establish the superiority of the proposed MM-DK estimator, two main theorems are assumed based on specific conditions: the Nondecreasing and skew-symmetric ϕ. The presence of errors with Zero-mean and finite variance, and the finiteness of Ω .

Theorem 1:

If $\sum_{j=1}^{p} \Omega_{jj(MM)} < \sum_{j=1}^{p} \frac{\sigma^2}{a_j}$, MSE$(\hat{\alpha}_{MM}(DK)) <$ MSE $(\hat{\alpha}(DK))$, where Ω represent the diagonal elements of $\Omega_{jj(MM)}$.

Proof: The MSE difference between the DK and MM-DK estimators is given by

$$\Delta_1 = MSE\left(\hat{\alpha}_{MM}(DK)\right) - \text{MSE}(\hat{\alpha}(DK))$$

$$= \sum_{j=1}^{p} \frac{\left(a_j - k(1+d)\right)^2 (a_j \Omega_{jj(MM)} - \sigma^2)}{a_j \left(a_j + k(1+d)\right)^2} \qquad (35)$$

If $\left(a_j \Omega_{jj(MM)} - \sigma^2\right) < 0$ in the last equation, which implies that $\sum_{j=1}^{p} \Omega_{jj(MM)} < \sum_{j=1}^{p} \frac{\sigma^2}{a_j}$, $MSE\left(\hat{\alpha}_{MM}(DK)\right) <$ MSE$(\hat{\alpha}(DK))$, implying that MM-DK estimator is better than DK estimator if $\sum_{j=1}^{p} \Omega_{jj(MM)} < \sum_{j=1}^{p} \frac{\sigma^2}{a_j}$.

Theorem 2:

MSE $(\hat{\alpha}_{MM}(DK)) <$ MSE $(\hat{\alpha}_{M}(DK))$.

Proof: The difference $MSE\left(\hat{\alpha}_{MM}(DK)\right) - MSE(\hat{\alpha}_{M}(DK))$ is given by

$$\Delta_2 = MSE\left(\hat{\alpha}_{MM}(DK)\right) - MSE(\hat{\alpha}_{M}(DK)).$$

$$= \sum_{j=1}^{p} \left[\frac{\left(a_j - k(1+d)\right)^2 \Omega_{jj(MM)} + 4k^2(1+d)^2 \alpha_j^2}{\left(a_j + k(1+d)\right)^2} - \frac{\left(a_j - k(1+d)\right)^2 \Omega_{jj} + 4k^2(1+d)^2 \alpha_j^2}{\left(a_j + k(1+d)\right)^2} \right],$$

$$= \sum_{j=1}^{P} \frac{\left(a_j - k(1+d)\right)^2 (\Omega_{jj(MM)} - \Omega_{jj})}{\left(a_j + k(1+d)\right)^2}. \qquad (36)$$

2.3. Selection of shrinkage parameter

Assume that $\hat{\alpha}_{MM}$ follows a normal distribution with mean equals α and covariance matrix $B^2 A^{-1}$.This assumption holds nearly in practice when

$\sqrt{n}(\hat{\alpha}_{MM}^2 - \alpha) \rightarrow N(0, B^2A^{-1})$,Here$B^2 = k_0^2 E\,[\varphi^2(\varepsilon/k_0)]/[E\varphi'(\varepsilon/k_0)]^2$, and the scale estimates k_0. The $\alpha_{jj(MM)}^2$ unbiased estimator is equal to $\hat{\alpha}_{jj(MM)}^2$, and the $\Omega_{jj(MM)}$ unbiased estimator is asymptotically equal to $\hat{B}^2/a_j$, where $\hat{B}^2$ is written as $\hat{B}^2 = k^2(n-p)^{-1}\sum_{i=1}^{n}[\varphi(e_i/k)]^2/\sum_{i=1}^{n}\left[\frac{1}{n}\varphi'(e_i/k)\right]^2$.

"We state the estimates of the biasing parameters for the robust estimators as follows:"

1- As per Hoerl and Kennard, $\hat{k}$ of the RRR estimator is given by

$$\hat{k} = \frac{\hat{B}^2}{max\left(\hat{\alpha}_{Mj}^2\right)}. \qquad (37)$$

2- As per Liu, $\hat{d}$ of the RL estimator is given by

$$\hat{d} = \frac{\sum_{j=1}^{p} a_j(\hat{\alpha}_{Mj}^2 - \hat{B}^2)}{\sum_{j=1}^{p}\left(\hat{B}^2 + a_j\hat{\alpha}_{Mj}^2\right)}. \qquad (38)$$

3- $\hat{k}_{(OK)}$ *of the MOK estimator is given by* :

$$\hat{k}_{(OK)} = \frac{p\hat{B}^2}{\sum_{j=1}^{p}\left(\hat{\alpha}_{Mj}^2 - \hat{d}_{OK}\left((\hat{B}^2/a_i) + \hat{\alpha}_{Mj}^2\right)\right)} \quad , \qquad (39)$$

where, $\hat{d}_{OK}$ alternative biasing parameter following Özkale and Kaçiranlar is given by:

$$\hat{d}_{OK} = min\left[\frac{\hat{\alpha}_{Mj}^2}{\left(\hat{B}^2/a_i\right) + \hat{\alpha}_{Mj}^2}\right]_{j=1}^{p} \quad . \qquad (40)$$

4- $\hat{k}_{DK}$ *of the MDK estimator is given by* :

$$\hat{k}_{DK} = \frac{1}{p}\sum_{j=1}^{p}\frac{\hat{B}^2}{(1+\hat{d}_{TP})\left((\hat{B}^2/a_j) + 2\hat{\alpha}_{Mj}^2\right)} \qquad (41)$$

5- Dawoud and Kibria $\hat{k}$ of the proposed estimator MMDK is given by

$$\hat{k}_{MMDK} = \frac{1}{p}\sum_{j=1}^{p} \frac{\hat{B}^2}{(1+\hat{d}_{TP})\left(\left(\hat{B}^2/a_j\right)+2\hat{\alpha}_{jjMM}^2\right)}. \quad (42)$$

Where $\hat{d}_{TP}$ suggested by Ozkale and Kaciranlar is given by

$$\hat{d}_{TP} = min\left[\frac{\hat{\alpha}_{Mj}^2}{\left(\hat{B}^2/a_j\right)+\hat{\alpha}_{Mj}^2}\right]_{j=1}^{p}. \quad (43)$$

For non-robust estimators, we use small positive constants as estimates for the biasing parameters.

6- $\hat{k}$ of Ridge estimator is given by:

$$\hat{k} = \frac{\hat{B}^2}{\max\left(\hat{\alpha}_j^2\right)} \quad (44)$$

7- $\hat{d}$ of Liu estimator is given by:

$$\hat{d} = \frac{\sum_{j=1}^{p} a_j\left(\hat{\alpha}_j^2-\hat{B}^2\right)}{\sum_{j=1}^{p}\left(\hat{B}^2+a_j\hat{\alpha}_j^2\right)} \quad (45)$$

8- $\hat{k}$ *of Dawoud and Kibria estimator is given by*:

$$\hat{k} = \frac{1}{p}\sum_{j=1}^{p} \frac{\hat{B}^2}{(1+\hat{d}_{TP})\left((\hat{B}^2/a_j)+2\hat{\alpha}_j^2\right)} \quad (46)$$

9- $\hat{k}$ *of Ozkale and kaciranlar estimator is given by*:

$$\hat{k} = \frac{p\hat{B}^2}{\sum_{j=1}^{p}\left(\hat{\alpha}_{Mj}^2-\hat{d}_{OK}\left((\hat{B}^2/a_i)+\hat{\alpha}_{Mj}^2\right)\right)} \quad (47)$$

3. Monte Carlo simulation

We performed a Monte Carlo simulation study to evaluate the performance of the LS, M-estimator (ME), Ridge, RRidge, Liu, MLiu, OK, MOK, DK, MDK, and the proposed MMDK estimators. The simulation study was implemented using the R programming language, as referenced by Abonazel (2018).

3.1. Experimental setup

We simulated the datasets using equation (1) with the following settings: According to Gibbons (1981), and Kibria (2003), the independent variables in the Monte Carlo experiment were generated with varying degrees of multicollinearity, following a specific procedure.

$$x_{ij} = z_{ij}\sqrt{1-\rho^2} + \rho z_{ip} \quad ; \; z_{ij} \sim N(0,1) \qquad (\mathbf{48})$$

Where $i = 1,2,\ldots\ldots,n \;\; ; j = 1,2,\ldots\ldots,p \; and \; \rho^2$ represents the correlation between the independent variables, and $z_{ij}{}'s$ are independent random variables obtained from the standard normal distribution.

1- In accordance with the studies conducted by Kaçıranlar and Dawoud (2018), Elgohary et al. (2019), Abonazel (2019), Lukman et al. (2021), Farghali et al. (2021), Abonazel et al., 2023, Abdelwahab et al., 2024, and Abonazel et al., 2024, the parameter values were specified with the constraint that $\acute{\beta}\beta = 1 \; and \; \beta_1 = \cdots = \beta_p$, which is typical for simulation research of this kind.

2- The number of independent variables p where (p = 3, 6, and 9).

3- The sample size n where (n= 30, 50, 100, and 200).

4- The correlation among the independent variables (ρ) was selected as an influential parameter in our simulation analysis, with values set at (ρ = .85, .90, and .95).

5- the generated errors follow a normal distribution with mean 0 and variance $\sigma^2 = 1$ and 5

6- To introduce outliers, we randomly replaced a specified percentage (τ) of values from the error vector with a value calculated as (third quartile of the errors + 5 × interquartile range of the errors). This process was repeated for different values of τ (5%, 10%, and 20%) (Dawoud et al., 2022) and (Awwad et al., 2021).

7- The Monte Carlo simulation experiments are replicated 1000 times to ensure reliable and robust results.

In the simulation investigation, we use MSE to compare the performance of the suggested methods. MSE is a familiar criterion as it depends on the mean, which has good statistical properties. The MSE has been used in several studies, such as Dawoud and Abonazel (2021) the MSE is calculated as follow:

$$MSE(\hat{\alpha}) = \frac{1}{L}\sum_{l=1}^{L}\left[(\hat{\alpha}_l - \alpha)'(\hat{\alpha}_l - \alpha)\right] \qquad \textbf{(49)}$$

Where $\hat{\alpha}_l$ indicates the vector of the estimated parameter at 1^{th} iteration and α is the vector of the real parameter values. The experiment was repeated 1,000 times. We present the estimated mean squared errors for each of the estimators in tables 1 through 9, using the relevant bias parameters for both of which have been evaluated.

3.2. Simulation results

We evaluate the performance of 11 estimators 5 non-robust and 6 robust and propose one of the robust estimators, the MM-DK, as a modified estimator. From the simulation results in Tables 1–9, we make the following observations:

1- The LS estimator performs the worst when both multicollinearity and outlier issues occur simultaneously.

2- As the error variance σ^2, the number of explanatory variables p, the degree of multicollinearity ρ, and the percentage of outliers τ increase, the MSE values for the estimators also increase.
3- The MSE values of the estimators decrease as the sample size increases.
4- The MMDK estimator outperforms the MDK estimator in most cases across various values of, τ, σ^2, ρ, p and n.
5- Finally, the MMDK estimator generally performs better than both robust and non-robust estimators when multicollinearity and outlier problems occur simultaneously.

Table 1: MSE of different estimators when p = 3 and outliers = 5%

σ^2	*n*	*ρ*	Non-robust estimators					Robust estimators					
			OLS	**Ridge**	**Liu**	**OK**	**DK**	**ME**	**RRidge**	**MLiu**	**MOK**	**MDK**	**MMDK**
1	**30**	**0.85**	4.220	2.250	3.503	1.637	0.633	0.423	0.306	0.350	0.254	0.164	0.143
		0.90	7.941	4.176	5.431	2.936	1.253	0.888	0.527	0.589	0.421	0.210	0.187
		0.95	19.323	9.744	10.005	6.762	4.248	2.218	1.256	1.063	0.981	0.358	0.302
	50	**0.85**	2.132	1.165	1.846	0.916	0.451	0.277	0.215	0.238	0.179	0.103	0.088
		0.90	3.553	1.864	2.837	1.382	0.497	0.473	0.326	0.373	0.269	0.173	0.156
		0.95	5.401	2.822	3.808	2.028	0.471	0.779	0.469	0.532	0.377	0.222	0.196
	100	**0.85**	1.277	0.756	1.195	0.603	0.360	0.138	0.123	0.130	0.107	0.067	0.056
		0.90	1.814	0.962	1.619	0.737	0.378	0.223	0.179	0.197	0.150	0.081	0.074
		0.95	4.313	2.167	3.220	1.522	0.426	0.528	0.342	0.391	0.274	0.133	0.122
	200	**0.85**	0.583	0.386	0.563	0.315	0.206	0.067	0.063	0.065	0.058	0.039	0.035
		0.90	0.717	0.428	0.682	0.343	0.204	0.087	0.079	0.083	0.070	0.041	0.038
		0.95	2.152	1.147	1.886	0.863	0.372	0.254	0.200	0.223	0.166	0.087	0.078
5	**30**	**0.85**	29.978	15.604	21.165	10.249	10.668	3.346	1.828	2.466	1.377	0.553	0.433
		0.90	50.159	25.521	28.518	16.159	27.670	5.512	2.956	3.382	2.107	0.728	0.541
		0.95	84.867	43.268	43.454	27.886	35.980	9.155	4.881	5.013	3.409	1.009	0.756
	50	**0.85**	13.100	6.858	10.648	4.655	2.446	1.847	1.021	1.512	0.784	0.365	0.349
		0.90	11.440	5.865	9.750	4.087	1.309	1.657	0.934	1.412	0.735	0.407	0.377
		0.95	28.840	14.470	19.733	9.587	7.393	3.852	1.974	2.685	1.431	0.435	0.408
	100	**0.85**	6.490	3.493	5.972	2.466	0.896	0.762	0.478	0.703	0.390	0.240	0.222
		0.90	8.700	4.498	7.803	3.143	1.723	1.018	0.594	0.912	0.477	0.258	0.235
		0.95	19.559	9.858	15.520	6.680	5.259	2.265	1.173	1.804	0.884	0.363	0.355
	200	**0.85**	2.518	1.311	2.432	0.989	0.449	0.295	0.221	0.285	0.185	0.114	0.104
		0.90	4.487	2.352	4.201	1.656	0.471	0.525	0.347	0.492	0.279	0.154	0.137
		0.95	8.943	4.678	8.062	3.253	0.675	0.948	0.554	0.857	0.444	0.268	0.240

Table 2: MSE of different estimators when p = 3 and outliers = 10%

σ^2	*n*	*ρ*	Non-robust estimators					Robust estimators					
			OLS	**Ridge**	**Liu**	**OK**	**DK**	**ME**	**RRidge**	**MLiu**	**MOK**	**MDK**	**MMDK**
1	**30**	**0.85**	5.997	3.145	5.001	2.257	0.864	0.492	0.343	0.411	0.285	0.193	0.146
		0.90	9.481	5.075	7.126	3.653	1.602	0.803	0.506	0.597	0.418	0.224	0.151
		0.95	28.132	13.822	12.348	9.031	10.397	2.435	1.295	1.104	0.968	0.368	0.274
	50	**0.85**	4.748	2.405	4.027	1.698	0.600	0.372	0.263	0.316	0.216	0.131	0.105
		0.90	8.163	4.677	6.535	3.310	1.404	0.556	0.372	0.445	0.307	0.165	0.124
		0.95	13.028	6.393	8.362	4.174	2.092	1.024	0.564	0.668	0.442	0.229	0.190
	100	**0.85**	2.041	1.014	1.870	0.772	0.381	0.180	0.149	0.166	0.129	0.081	0.069
		0.90	2.960	1.520	2.652	1.108	0.419	0.246	0.193	0.221	0.162	0.087	0.070
		0.95	6.718	3.233	5.326	2.281	0.649	0.534	0.348	0.422	0.282	0.148	0.115
	200	**0.85**	1.108	0.635	1.063	0.493	0.267	0.082	0.074	0.079	0.067	0.043	0.040
		0.90	1.811	0.979	1.691	0.734	0.307	0.138	0.117	0.128	0.101	0.051	0.043
		0.95	3.779	1.995	3.342	1.463	0.451	0.265	0.202	0.235	0.168	0.087	0.058
5	**30**	**0.85**	58.544	28.802	34.640	18.043	26.725	5.099	2.626	3.216	1.895	1.173	0.817
		0.90	39.902	19.815	30.256	13.048	10.695	3.637	1.980	2.825	1.491	0.559	0.434
		0.95	103.127	50.921	57.237	32.695	41.236	8.725	4.518	5.113	3.169	0.670	0.486
	50	**0.85**	17.015	8.863	15.039	5.824	3.208	1.313	0.766	1.177	0.609	0.353	0.288
		0.90	29.180	14.779	23.265	9.418	9.244	2.546	1.393	2.093	1.059	0.433	0.400
		0.95	69.955	35.186	45.215	22.742	28.750	5.977	3.210	4.008	2.330	0.445	0.385
	100	**0.85**	12.126	6.002	10.956	3.926	2.922	1.025	0.602	0.936	0.479	0.268	0.229
		0.90	13.982	6.671	12.525	4.346	1.873	1.205	0.683	1.084	0.535	0.319	0.277
		0.95	34.412	18.137	27.859	12.082	12.278	2.506	1.352	2.063	1.024	0.396	0.328
	200	**0.85**	4.906	2.460	4.742	1.743	0.573	0.375	0.273	0.364	0.224	0.149	0.116
		0.90	7.029	3.583	6.668	2.464	0.707	0.552	0.368	0.525	0.298	0.174	0.144
		0.95	17.530	8.920	15.322	5.725	4.860	1.266	0.697	1.117	0.537	0.245	0.228

Table 3: MSE of different estimators when p = 3 and outliers = 20%

σ^2	*n*	ρ	Non-robust estimators					Robust estimators					
			OLS	**Ridge**	**Liu**	**OK**	**DK**	**ME**	**RRidge**	**MLiu**	**MOK**	**MDK**	**MMDK**
1	**30**	**0.85**	11.519	5.883	9.567	4.044	1.698	1.039	0.646	0.880	0.522	0.326	0.170
		0.90	22.078	11.488	14.941	7.582	5.983	2.193	1.399	1.565	1.115	0.328	0.179
		0.95	38.870	20.530	23.966	13.710	11.290	3.897	2.367	2.585	1.827	0.559	0.350
	50	**0.85**	8.177	3.836	6.857	2.544	1.203	0.666	0.426	0.565	0.345	0.205	0.134
		0.90	12.529	6.292	9.702	4.074	2.980	0.978	0.584	0.765	0.460	0.222	0.144
		0.95	27.636	14.131	19.105	9.785	3.878	2.173	1.235	1.513	0.983	0.470	0.234
	100	**0.85**	3.303	1.668	3.087	1.206	0.525	0.236	0.187	0.221	0.157	0.094	0.063
		0.90	5.921	2.816	5.195	1.926	0.837	0.428	0.289	0.377	0.236	0.141	0.094
		0.95	12.847	6.388	10.332	4.249	1.049	0.877	0.531	0.711	0.423	0.256	0.148
	200	**0.85**	1.721	0.898	1.663	0.694	0.331	0.112	0.100	0.108	0.089	0.061	0.041
		0.90	3.053	1.579	2.901	1.177	0.500	0.195	0.162	0.186	0.138	0.085	0.047
		0.95	6.530	3.272	5.835	2.279	0.448	0.424	0.303	0.381	0.246	0.139	0.075
5	**30**	**0.85**	98.444	48.339	66.698	30.535	39.984	8.962	4.692	6.493	3.429	1.218	0.482
		0.90	126.628	65.240	82.838	41.847	54.974	13.421	8.039	9.456	5.801	3.015	0.613
		0.95	335.282	201.780	168.749	136.554	143.779	58.569	44.915	35.013	36.262	12.794	0.932
	50	**0.85**	36.635	17.064	30.659	10.516	10.348	2.961	1.602	2.558	1.219	0.516	0.339
		0.90	79.260	42.369	62.264	28.362	21.312	6.066	3.452	4.870	2.602	0.814	0.426
		0.95	148.371	72.432	87.681	46.035	93.596	11.058	5.658	6.925	3.875	2.240	0.586
	100	**0.85**	17.392	8.051	16.136	5.182	3.113	1.241	0.686	1.160	0.544	0.339	0.241
		0.90	30.074	14.788	26.519	9.354	8.092	1.979	1.032	1.777	0.792	0.376	0.262
		0.95	65.765	31.277	50.023	19.988	27.558	4.649	2.351	3.573	1.660	0.492	0.337
	200	**0.85**	9.562	4.511	9.152	2.976	1.747	0.614	0.373	0.589	0.301	0.185	0.119
		0.90	14.535	6.892	13.656	4.472	2.806	0.938	0.517	0.884	0.410	0.258	0.180
		0.95	25.432	11.686	22.523	7.292	9.093	1.729	0.876	1.537	0.654	0.286	0.223

Table 4: MSE of different estimators when p = 6 and outliers = 5%

σ^2	*n*	*ρ*	Non-robust estimators					Robust estimators					
			OLS	**Ridge**	**Liu**	**OK**	**DK**	**ME**	**RRidge**	**MLiu**	**MOK**	**MDK**	**MMDK**
1	**30**	**0.85**	13.983	7.968	9.038	3.946	1.651	1.591	1.028	1.030	0.621	0.257	0.217
		0.90	33.426	18.483	11.568	8.297	10.941	4.023	2.345	1.399	1.215	1.163	0.904
		0.95	34.286	19.446	15.274	9.348	3.476	3.808	2.296	1.750	1.303	0.513	0.421
	50	**0.85**	4.615	2.664	3.667	1.472	0.740	0.655	0.489	0.518	0.314	0.150	0.135
		0.90	8.826	5.010	5.767	2.515	1.048	1.234	0.810	0.820	0.492	0.216	0.204
		0.95	18.772	10.593	8.480	5.043	3.319	2.657	1.593	1.210	0.871	0.468	0.411
	100	**0.85**	3.059	1.763	2.678	0.985	0.461	0.364	0.302	0.318	0.209	0.102	0.096
		0.90	4.488	2.638	3.808	1.420	0.548	0.526	0.422	0.447	0.278	0.104	0.094
		0.95	8.940	5.027	6.092	2.493	1.357	1.052	0.701	0.714	0.413	0.188	0.158
	200	**0.85**	1.524	0.960	1.416	0.557	0.237	0.180	0.163	0.168	0.126	0.064	0.062
		0.90	2.124	1.273	1.909	0.710	0.330	0.258	0.223	0.232	0.160	0.072	0.066
		0.95	4.565	2.576	3.674	1.321	0.723	0.548	0.418	0.441	0.266	0.111	0.098
5	**30**	**0.85**	68.482	38.507	46.304	18.591	15.108	7.790	4.601	5.333	2.515	0.913	0.786
		0.90	126.097	70.530	62.996	33.014	42.159	14.580	8.761	7.444	4.690	2.308	1.672
		0.95	237.719	130.818	79.470	61.472	115.565	27.126	15.730	9.136	7.824	5.836	3.867
	50	**0.85**	28.543	15.974	20.697	7.406	6.205	4.131	2.413	2.993	1.287	0.726	0.651
		0.90	45.084	25.747	29.676	12.365	7.455	6.556	3.791	4.333	2.000	0.829	0.746
		0.95	78.665	43.618	37.699	19.699	22.770	11.727	6.679	5.691	3.299	1.875	1.597
	100	**0.85**	13.594	7.480	11.890	3.635	2.032	1.667	1.040	1.452	0.595	0.299	0.273
		0.90	22.612	12.810	18.573	6.240	3.449	2.570	1.519	2.111	0.852	0.432	0.371
		0.95	49.583	27.606	33.264	12.863	10.244	6.043	3.445	4.027	1.769	0.786	0.697
	200	**0.85**	7.132	3.991	6.651	2.004	0.995	0.843	0.587	0.785	0.359	0.177	0.159
		0.90	11.543	6.794	10.475	3.392	1.336	1.307	0.870	1.187	0.520	0.219	0.187
		0.95	23.409	13.496	19.134	6.491	3.045	2.710	1.632	2.216	0.909	0.404	0.353

Table 5: MSE of different estimators when p = 6 and outliers = 10%

σ^2	*n*	*ρ*	Non-robust estimators					Robust estimators					
			OLS	**Ridge**	**Liu**	**OK**	**DK**	**ME**	**RRidge**	**MLiu**	**MOK**	**MDK**	**MMDK**
1	**30**	**0.85**	21.519	12.458	14.039	6.309	2.870	2.118	1.370	1.374	0.824	0.324	0.221
		0.90	34.495	19.595	16.960	9.317	6.652	3.370	2.105	1.665	1.203	0.579	0.390
		0.95	63.110	35.279	21.751	16.628	24.963	6.860	4.215	2.385	2.326	1.268	0.718
	50	**0.85**	11.028	6.279	8.657	3.174	1.308	0.969	0.686	0.762	0.438	0.194	0.147
		0.90	21.236	12.757	14.258	6.266	2.324	1.670	1.078	1.131	0.635	0.221	0.149
		0.95	31.625	17.196	16.192	8.090	5.153	2.725	1.594	1.407	0.887	0.457	0.330
	100	**0.85**	5.050	2.840	4.483	1.519	0.680	0.392	0.324	0.350	0.224	0.104	0.094
		0.90	8.812	4.908	7.099	2.468	1.153	0.709	0.516	0.570	0.319	0.136	0.103
		0.95	18.075	10.345	12.696	5.152	2.292	1.353	0.865	0.944	0.516	0.223	0.161
	200	**0.85**	2.808	1.672	2.653	0.959	0.413	0.211	0.192	0.200	0.147	0.063	0.053
		0.90	4.700	2.684	4.215	1.428	0.631	0.347	0.290	0.311	0.198	0.078	0.065
		0.95	9.090	5.422	7.453	2.778	1.373	0.631	0.470	0.515	0.294	0.120	0.086
5	**30**	**0.85**	87.233	50.137	59.851	24.081	28.967	8.678	5.233	5.940	2.815	1.247	0.722
		0.90	136.943	81.852	86.177	41.156	36.447	12.881	7.943	8.125	4.335	1.473	0.940
		0.95	268.653	148.664	101.622	68.096	156.915	27.111	16.091	10.650	8.100	5.428	2.677
	50	**0.85**	76.407	45.025	55.809	21.901	15.410	6.389	3.890	4.663	2.093	0.803	0.539
		0.90	118.456	72.851	74.821	35.417	41.535	8.758	5.200	5.526	2.641	1.484	0.936
		0.95	153.675	86.543	83.993	41.207	38.332	13.723	8.002	7.662	4.212	1.758	1.236
	100	**0.85**	31.577	18.133	26.993	8.565	6.956	2.440	1.478	2.082	0.821	0.424	0.305
		0.90	39.880	21.999	33.762	10.656	4.920	3.244	1.935	2.726	1.082	0.467	0.356
		0.95	91.635	52.463	63.907	25.481	23.771	6.791	3.908	4.793	2.061	0.836	0.626
	200	**0.85**	13.799	7.984	13.057	4.101	1.387	1.053	0.743	0.997	0.467	0.208	0.160
		0.90	17.841	9.880	16.337	4.762	2.106	1.358	0.880	1.243	0.517	0.236	0.194
		0.95	39.043	21.653	32.266	10.039	5.542	3.121	1.828	2.580	0.994	0.442	0.345

Table 6: MSE of different estimators when p = 6 and outliers = 20%

σ^2	*n*	*ρ*	Non-robust estimators					Robust estimators					
			OLS	**Ridge**	**Liu**	**OK**	**DK**	**ME**	**RRidge**	**MLiu**	**MOK**	**MDK**	**MMDK**
1	**30**	**0.85**	42.264	23.904	27.983	11.948	6.872	7.902	5.233	5.319	3.190	1.321	0.256
		0.90	105.103	62.915	43.987	30.622	40.097	19.594	13.443	8.794	8.000	4.211	0.666
		0.95	197.754	123.401	58.640	60.943	98.469	42.844	30.139	12.393	16.853	16.442	1.218
	50	**0.85**	27.053	15.778	20.773	8.380	4.384	2.588	1.700	1.983	1.063	0.528	0.184
		0.90	40.346	24.360	28.140	12.061	8.550	3.490	2.234	2.423	1.284	0.426	0.176
		0.95	58.835	33.644	33.191	15.719	11.942	5.198	3.208	2.952	1.726	0.732	0.276
	100	**0.85**	8.982	4.807	8.000	2.325	0.969	0.636	0.470	0.568	0.297	0.129	0.089
		0.90	18.551	10.513	14.840	5.000	2.849	1.252	0.822	1.000	0.481	0.242	0.129
		0.95	32.799	17.847	22.653	8.247	4.242	2.258	1.327	1.572	0.748	0.335	0.190
	200	**0.85**	4.266	2.241	4.017	1.151	0.548	0.288	0.246	0.272	0.173	0.089	0.070
		0.90	9.102	5.339	8.109	2.522	1.136	0.546	0.415	0.487	0.262	0.118	0.070
		0.95	15.293	8.297	12.265	3.654	2.148	1.003	0.670	0.805	0.382	0.180	0.099
5	**30**	**0.85**	209.147	117.433	132.851	57.255	71.870	51.691	36.891	33.759	22.990	9.770	1.121
		0.90	373.112	233.349	190.290	111.198	186.870	71.587	51.394	38.758	29.952	19.811	1.244
		0.95	907.145	477.102	196.345	196.338	663.875	136.139	84.370	35.173	42.283	70.830	13.336
	50	**0.85**	109.935	60.888	86.032	28.587	27.752	9.527	5.677	7.473	3.033	1.066	0.510
		0.90	169.572	96.082	119.770	46.172	54.102	14.608	8.591	10.505	4.565	1.675	0.701
		0.95	377.276	204.436	159.585	87.020	216.914	31.438	17.794	13.777	8.402	8.337	2.898
	100	**0.85**	76.234	48.128	66.041	23.471	11.610	4.723	3.011	4.103	1.641	0.575	0.267
		0.90	87.537	45.451	66.419	19.591	25.550	6.500	3.611	4.926	1.799	1.069	0.598
		0.95	173.348	98.710	120.912	47.376	59.236	11.904	6.913	8.188	3.554	1.707	0.838
	200	**0.85**	25.992	14.828	24.524	7.215	3.735	1.652	1.043	1.560	0.629	0.293	0.169
		0.90	38.651	21.466	35.560	10.334	6.147	2.581	1.548	2.366	0.883	0.386	0.210
		0.95	86.509	48.647	71.391	23.123	27.057	5.639	3.247	4.611	1.705	0.767	0.410

Table 7: MSE of different estimators when p = 9 and outliers = 5%

σ^2	*n*	*ρ*	Non-robust estimators					Robust estimators					
			OLS	**Ridge**	**Liu**	**OK**	**DK**	**ME**	**RRidge**	**MLiu**	**MOK**	**MDK**	**MMDK**
1	**30**	**0.85**	22.859	13.752	13.858	5.872	3.459	2.719	1.821	1.684	0.969	0.477	0.363
		0.90	42.575	25.441	18.781	10.723	7.873	5.515	3.573	2.533	1.803	0.848	0.660
		0.95	109.967	63.283	23.488	24.817	40.680	13.716	8.597	3.245	3.838	4.002	2.893
	50	**0.85**	9.545	5.718	6.309	2.478	1.965	1.384	0.952	0.920	0.491	0.332	0.295
		0.90	12.143	7.362	7.941	3.260	1.683	1.855	1.258	1.203	0.656	0.300	0.277
		0.95	30.194	17.953	12.457	7.321	6.195	4.382	2.684	1.832	1.242	0.901	0.788
	100	**0.85**	5.206	3.238	4.394	1.486	0.795	0.622	0.506	0.525	0.294	0.133	0.117
		0.90	8.075	4.980	6.141	2.193	1.472	0.897	0.670	0.689	0.362	0.202	0.161
		0.95	16.079	9.657	9.511	4.064	3.476	1.865	1.229	1.107	0.604	0.430	0.368
	200	**0.85**	2.594	1.666	2.369	0.824	0.415	0.300	0.268	0.275	0.178	0.087	0.084
		0.90	3.571	2.199	3.159	1.069	0.576	0.403	0.345	0.356	0.213	0.098	0.091
		0.95	7.006	4.153	5.460	1.843	1.271	0.835	0.630	0.647	0.339	0.171	0.144
5	**30**	**0.85**	163.908	99.508	79.614	40.502	55.738	20.093	12.706	9.917	5.667	3.931	2.801
		0.90	275.951	161.544	93.100	63.662	130.841	35.708	22.309	12.004	9.740	9.816	7.212
		0.95	387.313	221.444	112.814	87.570	167.954	46.802	28.813	13.648	12.653	9.281	6.839
	50	**0.85**	45.228	27.438	33.081	11.649	5.825	6.683	4.177	4.888	1.999	0.804	0.701
		0.90	74.616	44.246	39.690	17.544	23.795	11.225	6.809	5.953	2.914	2.864	2.506
		0.95	154.507	91.504	60.102	35.638	52.281	22.634	13.651	8.843	5.615	4.918	4.145
	100	**0.85**	23.871	14.450	20.498	6.110	2.978	2.868	1.838	2.466	0.909	0.429	0.374
		0.90	37.754	22.887	29.269	9.562	6.035	4.512	2.759	3.514	1.312	0.746	0.663
		0.95	72.598	42.918	46.460	17.835	13.691	8.867	5.298	5.623	2.370	1.631	1.384
	200	**0.85**	10.246	6.075	9.531	2.684	1.508	1.252	0.903	1.164	0.483	0.236	0.210
		0.90	18.385	10.964	16.130	4.597	2.748	2.157	1.395	1.891	0.686	0.365	0.324
		0.95	34.661	20.473	27.046	8.412	5.961	4.122	2.523	3.224	1.186	0.723	0.616

Table 8: MSE of different estimators when p = 9 and outliers = 10%

σ^2	*n*	*ρ*	Non-robust estimators					Robust estimators					
			OLS	**Ridge**	**Liu**	**OK**	**DK**	**ME**	**RRidge**	**MLiu**	**MOK**	**MDK**	**MMDK**
1	**30**	**0.85**	32.469	19.428	19.841	8.289	5.098	3.760	2.514	2.330	1.358	0.660	0.391
		0.90	54.000	32.471	25.544	13.565	12.990	5.946	3.849	2.780	1.871	1.102	0.605
		0.95	109.652	65.313	33.937	27.301	31.323	13.003	8.466	4.221	4.193	2.224	1.199
	50	**0.85**	19.621	12.056	14.146	5.223	2.934	1.667	1.164	1.223	0.617	0.283	0.233
		0.90	35.235	22.625	21.533	9.653	5.747	2.885	1.951	1.804	1.000	0.473	0.276
		0.95	66.219	39.836	28.173	16.133	16.381	5.541	3.361	2.454	1.557	1.104	0.661
	100	**0.85**	10.099	6.095	8.540	2.731	1.360	0.803	0.628	0.674	0.353	0.167	0.135
		0.90	12.202	7.227	9.825	3.144	1.732	0.995	0.739	0.796	0.393	0.176	0.141
		0.95	28.398	17.025	18.413	6.932	4.075	2.209	1.441	1.443	0.712	0.330	0.257
	200	**0.85**	4.758	2.930	4.310	1.366	0.978	0.350	0.300	0.317	0.191	0.132	0.102
		0.90	6.645	3.992	5.855	1.791	1.120	0.494	0.409	0.438	0.246	0.122	0.092
		0.95	15.576	9.348	11.564	3.988	3.047	1.153	0.819	0.858	0.425	0.234	0.185
5	**30**	**0.85**	244.988	145.832	118.125	58.225	90.462	24.487	15.212	11.989	6.706	5.146	3.012
		0.90	303.373	182.960	136.894	75.615	121.353	33.441	21.410	15.520	9.935	5.627	2.872
		0.95	660.519	388.020	161.136	147.740	393.952	82.669	53.551	20.773	24.118	33.664	14.437
	50	**0.85**	119.602	71.511	84.233	29.747	31.261	10.713	6.588	7.521	3.027	1.201	0.828
		0.90	184.437	116.157	108.269	47.334	50.882	15.334	9.660	9.275	4.316	2.308	1.519
		0.95	289.648	168.748	134.652	66.855	112.370	26.346	16.051	12.333	7.020	3.618	2.593
	100	**0.85**	38.963	24.076	35.092	11.071	4.292	3.005	1.953	2.701	1.025	0.449	0.313
		0.90	76.532	45.101	57.590	18.400	16.873	6.268	3.815	4.704	1.759	1.156	0.842
		0.95	151.860	89.522	90.984	36.011	48.600	12.230	7.335	7.249	3.138	2.488	1.822
	200	**0.85**	20.819	12.548	19.308	5.282	3.007	1.605	1.110	1.490	0.580	0.310	0.230
		0.90	32.090	19.579	28.696	8.448	4.907	2.327	1.514	2.080	0.755	0.425	0.299
		0.95	71.949	43.961	56.600	18.595	12.652	5.284	3.232	4.150	1.503	0.879	0.644

Table 9: MSE of different estimators when p = 9 and outliers = 20%

σ^2	*n*	*ρ*	Non-robust estimators					Robust estimators					
			OLS	**Ridge**	**Liu**	**OK**	**DK**	**ME**	**RRidge**	**MLiu**	**MOK**	**MDK**	**MMDK**
1	**30**	**0.85**	63.616	38.606	39.785	16.925	9.974	25.109	17.686	16.245	9.855	3.991	0.525
		0.90	84.223	47.368	42.402	18.920	21.453	25.447	17.182	13.654	8.741	5.031	0.779
		0.95	280.594	170.434	64.437	66.184	158.635	102.531	71.475	25.261	33.706	44.340	2.070
	50	**0.85**	38.336	22.919	28.016	9.598	5.457	4.353	2.893	3.241	1.521	0.654	0.214
		0.90	48.994	28.274	32.442	11.397	7.685	5.753	3.836	3.887	1.978	0.840	0.278
		0.95	128.512	80.469	58.815	34.508	41.048	16.148	10.859	7.838	5.503	2.903	0.698
	100	**0.85**	24.210	15.971	20.125	6.604	3.052	1.568	1.128	1.339	0.595	0.179	0.129
		0.90	26.494	16.409	22.235	7.622	2.795	1.906	1.334	1.602	0.725	0.265	0.126
		0.95	63.435	37.069	36.802	15.661	16.151	5.072	3.093	2.872	1.461	1.161	0.538
	200	**0.85**	7.379	4.262	6.866	1.887	1.066	0.517	0.432	0.481	0.263	0.129	0.091
		0.90	12.151	7.324	10.958	3.201	1.481	0.775	0.610	0.703	0.347	0.123	0.088
		0.95	26.798	15.812	20.808	6.647	4.506	1.734	1.150	1.353	0.588	0.312	0.186
5	**30**	**0.85**	495.153	293.042	229.350	121.613	198.004	188.119	128.331	89.747	62.393	58.406	4.193
		0.90	527.043	335.959	264.916	143.865	182.219	196.463	140.900	103.471	71.954	45.615	2.877
		0.95	1438.595	874.888	359.918	357.054	992.507	482.842	337.193	135.789	169.873	256.087	11.189
	50	**0.85**	209.942	128.606	143.876	49.928	97.322	25.643	17.016	18.491	8.353	5.753	1.038
		0.90	324.130	186.900	181.369	73.753	158.065	35.909	22.238	20.539	9.975	8.933	2.342
		0.95	723.469	426.349	248.292	157.436	428.645	95.477	65.011	33.396	29.327	36.259	6.038
	100	**0.85**	93.806	52.905	77.149	20.698	21.982	7.098	4.224	5.810	1.885	1.214	0.629
		0.90	140.541	85.544	111.352	34.860	50.031	10.166	6.282	8.160	2.835	1.244	0.594
		0.95	325.799	210.462	196.488	84.036	145.411	22.992	14.936	14.400	6.454	3.357	1.280
	200	**0.85**	36.673	21.377	34.430	8.695	4.799	2.572	1.650	2.410	0.826	0.384	0.218
		0.90	60.065	36.544	54.414	15.578	6.707	3.874	2.433	3.513	1.195	0.561	0.274
		0.95	117.144	68.250	93.916	28.207	27.012	7.949	4.742	6.289	2.119	1.365	0.733

4. Conclusion

Outliers can destabilize Least Squares (LS) estimators in regression models, leading to inefficiency. While robust estimators address this instability, they may still falter under multicollinearity. To tackle both multicollinearity and outliers, we introduce a modified robust estimator, the MMDK, derived from the Dawoud-Kibria estimator. Our theoretical analysis shows that the MMDK estimator outperforms both the Dawoud-Kibria and RDK estimators. Simulation studies across various scenarios confirm that the MMDK estimator consistently excels in environments with both multicollinearity and outliers. A numerical example further validates its effectiveness, aligning with our theoretical predictions. Selecting an appropriate biasing parameter is essential for optimal performance of the MMDK estimator. Thus, we recommend the MMDK estimator for practitioners facing the challenges of both multicollinearity and outliers in regression models.

References

1. **Abdelwahab, M. M., Abonazel, M. R., Hammad, A. T., & El-Masry, A. M. (2024). Modified Two-Parameter Liu Estimator for Addressing Multicollinearity in the Poisson Regression Model. *Axioms*, *13*(1), 46–46. https://doi.org/10.3390/axioms13010046**
2. **Abonazel MR. A practical guide for creating Monte Carlo simulation studies using R. Int J Math Comput Sci. 2018;4(1):18–33**
3. **Abonazel, M. R. (2019). New ridge estimators of SUR model when the errors are serially correlated. *International Journal of Mathematical Archive*, *10*(7), 53-62.**
4. **Abonazel, M. R., Rashash, A., Ashrakat Adel Saber, Dawoud, I., Elsayed Tageldin, & Azazy, A. R. (2024). Developing ridge estimators for the extended Poisson-Tweedie regression model: Method, simulation, and application. *Scientific African*, *23*, e02006–e02006. https://doi.org/10.1016/j.sciaf.2023.e02006**
5. **Abonazel, M. R., Saber, A. A., & Awwad, F. A. (2023). Kibria–Lukman estimator for the Conway–Maxwell Poisson regression model: Simulation and applications. *Scientific African*, *19*, e01553. https://doi.org/10.1016/j.sciaf.2023.e01553**
6. **Arslan, O., & Billor, N. (2000). Robust Liu estimator for regression based on an M-estimator. *Journal of Applied Statistics*, *27*(1), 39–47. https://doi.org/10.1080/02664760021817**
7. **Awwad, F. A., Dawoud, I., & Abonazel, M. R. (2021). Development of robust Özkale–Kaçiranlar and Yang–Chang estimators for regression models in the presence of multicollinearity and outliers. *Concurrency and Computation: Practice and Experience*. https://doi.org/10.1002/cpe.6779**
8. **Dawoud, I., & Abonazel, M. R. (2021). Robust Dawoud–Kibria estimator for handling multicollinearity and outliers in the linear regression model. *Journal of Statistical Computation and***

Simulation, *91*(17), 3678–3692. https://doi.org/10.1080/00949655.2021.1945063

9. Dawoud, I., & Kibria, B. M. G. (2020). A New Biased Estimator to Combat the Multicollinearity of the Gaussian Linear Regression Model. *Stats*, *3*(4), 526–541. https://doi.org/10.3390/stats3040033
10. Dawoud, I., Awwad, F. A., Tag Eldin, E., & Abonazel, M. R. (2022). New Robust Estimators for Handling Multicollinearity and Outliers in the Poisson Model: Methods, Simulation and Applications. *Axioms*, *11*(11), 612. https://doi.org/10.3390/axioms11110612
11. Farghali, R. A., Qasim, M., Kibria, B. M. G., & Abonazel, M. R. (2021). Generalized two-parameter estimators in the multinomial logit regression model: methods, simulation and application. *Communications in Statistics - Simulation and Computation*, *52*(7), 3327–3342. https://doi.org/10.1080/03610918.2021.1934023
12. Gibbons, D. I. (1981). *A Simulation Study of Some Ridge Estimators*. *76*(373), 131–139. https://doi.org/10.1080/01621459.1981.10477619
13. Hoerl, A. E., & Kennard, R. W. (1970). Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics*, *12*(1), 55–67. https://doi.org/10.1080/00401706.1970.10488634
14. Huber, P. J. (1981). Robust Statistics. In *Wiley Series in Probability and Statistics*. John Wiley & Sons, Inc. https://doi.org/10.1002/0471725250
15. Kaçıranlar, S., & Dawoud, I. (2018). On the performance of the poisson and the negative binomial ridge predictors. *Communications in Statistics - Simulation and Computation*, *47*(6), 1751–1770. https://doi.org/10.1080/03610918.2017.1324978

16. **Kejian, L. (1993). A new class of biased estimate in linear regression.** ***Communications in Statistics - Theory and Methods*, *22*****(2), 393–402. https://doi.org/10.1080/03610929308831027**

17. **Kibria, B. M. G. (2003). Performance of Some New Ridge Regression Estimators.** ***Communications in Statistics - Simulation and Computation*, *32*****(2), 419–435. https://doi.org/10.1081/sac-120017499**

18. **Lukman, A. F., Aladeitan, B., Ayinde, K., & Abonazel, M. R. (2021). Modified ridge-type for the Poisson regression model: simulation and application.** ***Journal of Applied Statistics*, 1–13. https://doi.org/10.1080/02664763.2021.1889998**

19. **Lukman, A. F., Osowole, O. I., & Ayinde, K. (2015). Two Stage Robust Ridge Method in a Linear Regression Model.** ***Journal of Modern Applied Statistical Methods*, *14*****(2), 53–67. https://doi.org/10.22237/jmasm/1446350820**

20. **M. Elgohary, M., R. Abonazel, M., M. Helmy, N., & R. Azazy, A. (2019). New robust-ridge estimators for partially linear model.** ***International Journal of Applied Mathematical Research*, *8*****(2), 46. https://doi.org/10.14419/ijamr.v8i2.29932**

21. **Özkale, M. R., & Kaçiranlar, S. (2007). The Restricted and Unrestricted Two-Parameter Estimators.** ***Communications in Statistics - Theory and Methods*, *36*****(15), 2707–2725. https://doi.org/10.1080/03610920701386877**

22. **SILVAPULLE, M. J. (1991). ROBUST RIDGE REGRESSION BASED ON AN M-ESTIMATOR.** ***Australian Journal of Statistics*, *33*****(3), 319–333. https://doi.org/10.1111/j.1467-842x.1991.tb00438.x**

23. **Yohai, V. J. (1987). High Breakdown-Point and High Efficiency Robust Estimates for Regression.** ***The Annals of Statistics*, *15*****(2), 642-656. https://doi.org/10.1214/aos/1176350366**

مقدر داود كيبريا القوي (المحسن) ذو مرحلتين لمعالجة التعدد الخطي والقيم المتطرفة في نموذج الانحدار الخطي

الملخص العربي:

يتناول البحث تطوير واختبار تقديرات إحصائية لمعالجة مشكلة التعدد الخطي والتأثيرات السلبية للقيم المتطرفة في نماذج الانحدار الخطي، ولذلك، تم عرض تقدير بديل يعرف بتقدير داوود وكبريا المحسن ذو المرحلتين (Two Stage Robust Dawoud-Kibria Estimator) بهدف تحسين دقة مقدرات النماذج في وجود القيم المتطرفة والتعدد الخطي معا، مع مقارنة التقدير البديل المقترح مع تقديرات أخرى من حيث الأداء في حالة زيادة كلا من نسب القيم المتطرفة والتعدد الخطي. وتشير نتائج البحث إلى أن المقدر البديل يوفر دقة أفضل من المقدرات الأخرى عند زيادة القيم المتطرفة للبيانات والتعدد الخطي معا.

Augmented Alternative Estimation Methods for Handling Multicollinearity and Non-Normality Problems Simultaneously.