ISSN: 0974-5823 Vol. 7 No. 1 January, 2022
International Journal of Mechanical Engineering

# Quantile-Based Estimation of Biasing Parameter for Kibria-Lukman Regression Estimator: Simulation and Application to Portland Cement Data

Issam Dawoud[1], Mohamed R. Abonazel[2]*, Abeer R. Azazy[3]

[1]Department of Mathematics, Al-Aqsa University, Gaza, Palestine.

[2*] Department of Applied Statistics and Econometrics, Faculty of Graduate Studies for Statistical Research, Cairo University, Giza, Egypt

[3]Al-Alsun Higher Institute for Tourism, Hotels and Computer, Cairo, Egypt.

**Abstract**

In regression models, the multicollinearity effects on the ordinary least squares estimator performance make it inefficient. To solve this, several estimators are proposed for coping with multicollinearity problem in regression models as ridge, Liu, and the recent one is the Kibria-Lukman estimator (KLE). And because the important role of the biasing parameters in these estimators in determining their performances, so several estimators of the biasing parameters are proposed and examined in recent papers. But the existing estimators do not perform very well due to their mean squared error for multicollinearity high noise. Therefore, this paper proposes a new robust estimator of the biasing parameter of the KLE and compares its performance with some existing estimators. The new estimator of the biasing parameter is based on the quantile of the error variance ratio to the canonical parameter. Furthermore, the simulation study and the numerical example are performed for studying the performance of the proposed estimator. The main results indicate that our proposed estimator has a better performance when the multicollinearity degree and the error variance of the model are large.

**Keywords**: Linear regression; Least Squares; multicollinearity; Kibria-Lukman estimator; biasing parameter; MSE

## 1. Introduction

The multiple linear regression model in the matrix form is given as

$$y = Xb + c, \qquad (1)$$

where $y$ is an $n\times 1$ given responses vector, $X$ is defined as an $n\times p$ explanatory variables matrix, $b$ is called as an $p\times 1$ unknown regression parameter vector, and $c$ is known as an $n\times 1$ errors vector such that its mean is equal to zero and its variance-covariance matrix, $Cov(c)=\sigma^2 I_n$, $\sigma^2$ error variance and $I_n$ is an $n\times n$ identity matrix. So, the least squares estimator (LSE) of $b$ is

$$\hat{b}_{LSE} = (X'X)^{-1}X'y.$$

Theoretically, it is assumed that the explanatory variables in the regression model should be independent but practically this assumption may be violated which means there is relationships between the explanatory variables and this caused multicollinearity (Kibria 2003). The multicollinearity effects on the LSE is very serious such that the LSE becomes inefficient as well unstable by giving incorrect signs of the regression parameter estimates and other trouble issues, see Hoerl and Kennard (1970) and Ullah et al. (2019). As a remedy of that, various estimators are proposed and the basic one is the ridge regression estimator (RRE) (Hoerl and Kennard 1970). Then, the recent one is the Kibria-Lukman estimator (KLE) and it is given as:

$$\hat{b}_{KLE} = (X'X + k\mathrm{I}_p)^{-1}(X'X - k\mathrm{I}_p)\hat{b}_{LSE}$$

where $k \ge 0$ is the biasing parameter (Kibria and Lukman 2020). As any regression estimator, the determination of the biasing parameter in the recent KLE is important to show its performance. Different studies proposed and investigated the biasing parameter estimators for the RRE such as (Hoerl and Kennard 1970), (Hoerl et al. 1975), (Lawless and Wang 1976), (Hocking et al. 1976), (Kibria 2003), (Khalaf and Shukur 2005), (Alkhamisi et al. 2006), (Alkhamisi and Shukur 2008), (Muniz and Kibria 2009), (Muniz et al. 2012), (Abonazel and Farghali 2019), and (Abonazel and Taha 2021). In addition to, Dawoud (2021a) proposed and examined the new biasing parameter estimators for the KLE in the linear regression model. Since there is a little discussion of the multicollinearity with the noise parameter which means a high multicollinearity degree with a high variance of the error challenge and affect the existing biasing parameter estimators' performances; so, we introduce a new estimator of the biasing parameter, k, in the KLE to handle the above issue.

This paper sections are: the statistical methodology is stated in Sec. 2. In Sec. 3, we state the new estimator of the biasing parameter. A simulation study and a real data are stated in Sec. 4. In final, conclusion is stated in Sec. 5.

## 2. Statistical methodology

### 2.1. Canonical form

The popular canonical form of the model in (1) is

$$y = Z\alpha + c, \tag{2}$$

where $Z = XR$, $\alpha = R'b$, $R$ is called as an orthogonal matrix such that $Z'Z = R'X'XR = T = diag\,(t_1, t_2, \ldots, t_p)$. The LSE of $\alpha$ is as

$$\hat{\alpha}_{LSE} = T^{-1}Z'y. \tag{3}$$

The matrix mean squares error (MMSE) and the scalar mean squares error (SMSE) of the LSE are

$$MMSE(\hat{\alpha}_{LSE}) = Cov\,(\hat{\alpha}_{LSE}) = \sigma^2 T^{-1} \tag{4}$$

and

$$SMSE(\hat{\alpha}_{LSE}) = trace\,(MMSE\,(\hat{\alpha}_{LSE})) = \sigma^2 \sum_{i=1}^{p} \frac{1}{t_i}. \tag{5}$$

The generalized KLE of $\alpha$ (Dawoud 2021b) is given as

$$\hat{\alpha}_{GKLE} = WM\,\hat{\alpha}_{LSE}, \tag{6}$$

where $W = [I_p + KT^{-1}]^{-1}$, $M = [I_p - KT^{-1}]$, $K = diag\,(k_1, k_2, \ldots, k_p)$, $k_i > 0$, and $k_i = \dfrac{\sigma^2}{2\alpha_i^2 + (\sigma^2 / t_i)}$.

and the MMSE and the SMSE are given as

$$\begin{aligned} MMSE(\hat{\alpha}_{GKLE}) &= Cov\,(\hat{\alpha}_{GKLE}) + Bias\,(\hat{\alpha}_{GKLE})\,Bias\,(\hat{\alpha}_{GKLE})' \\ &= \sigma^2\, WM\, T^{-1} M' W' + [WM - I_p]\,\alpha\alpha'[WM - I_p]' \end{aligned} \tag{7}$$

and

$$\begin{aligned} SMSE(\hat{\alpha}_{GKLE}) &= trace\,(MMSE\,(\hat{\alpha}_{GKLE})) \\ &= \sigma^2 \sum_{i=1}^{p} \frac{(t_i - k_i)^2}{t_i\,(t_i + k_i)^2} + 4\sum_{i=1}^{p} \frac{k_i^2 \alpha_i^2}{(t_i + k_i)^2}. \end{aligned} \tag{8}$$

### 2.2. Biasing parameters

We state some existing estimators of the biasing parameter for the KLE are as follows:

$$\hat{k}_1 = \min\left\{\frac{\hat{\sigma}^2}{2\hat{\alpha}_i^2 + (\hat{\sigma}^2 / t_i)}\right\}. \quad \text{(Kibria and Lukman 2020)} \quad (9)$$

$$\hat{k}_2 = \frac{p\,\hat{\sigma}^2}{\sum_{i=1}^{p}\left(2\hat{\alpha}_i^2 + (\hat{\sigma}^2 / t_i)\right)}. \quad \text{(Kibria and Lukman 2020)} \quad (10)$$

$$\hat{k}_3 = \frac{p\,\hat{\sigma}^2}{\sum_{i=1}^{p}\left(2t_i\,\hat{\alpha}_i^2 + \hat{\sigma}^2\right)}. \quad \text{(Dawoud 2021a)} \quad (11)$$

$$\hat{k}_4 = \frac{\hat{\sigma}^2 \sum_{i=1}^{p}\left(2t_i\,\hat{\alpha}_i + \hat{\sigma}^2\right)^2}{\left(\sum_{i=1}^{p}\left(2t_i\,\hat{\alpha}_i^2 + \hat{\sigma}^2\right)\right)^2}. \quad \text{(Dawoud 2021a)} \quad (12)$$

$$\hat{k}_5 = \frac{1}{p}\sum_{i=1}^{p}\frac{\hat{\sigma}^2}{\left(2\hat{\alpha}_i^2 + (\hat{\sigma}^2 / t_i)\right)}. \quad \text{(Dawoud 2021a)} \quad (13)$$

$$\hat{k}_6 = \frac{\hat{\sigma}^2}{\left(\prod_{i=1}^{p}\left(2\hat{\alpha}_i^2 + (\hat{\sigma}^2 / t_i)\right)\right)^{\frac{1}{p}}}. \quad \text{(Dawoud 2021a)} \quad (14)$$

$$\hat{k}_7 = Median\left\{\frac{\hat{\sigma}^2}{\left(2\hat{\alpha}_i^2 + (\hat{\sigma}^2 / t_i)\right)}\right\}. \quad \text{(Dawoud 2021a)} \quad (15)$$

## 3. The New Estimator

Consider that $(\hat{k}_1, \hat{k}_2, \ldots, \hat{k}_p)$ are the real values of $\hat{k}_i$ that defined in equation (6). Writing $(\hat{k}_1, \hat{k}_2, \ldots, \hat{k}_p)$ in the magnitude ascending order as

$$\hat{k}_{(1)} \le \hat{k}_{(2)} \le \Lambda \le \hat{k}_{(p)} \quad (16)$$

where $\hat{k}_{(1)}$ and $\hat{k}_{(p)}$ are the minimum and the maximum of $(\hat{k}_1, \hat{k}_2, \ldots, \hat{k}_p)$ respectively, and $\{\hat{k}_{(1)}, \hat{k}_{(2)}, \ldots, \hat{k}_{(p)}\}$ is the order statistics set for $(\hat{k}_1, \hat{k}_2, \ldots, \hat{k}_p)$ as well $\hat{k}_{(i)}$, $i = 1, 2, \ldots, p$ is defined as the observation of i-th order. By following Suhail et al. (2020), we consider $KQ_\tau$, $0 < \tau < 1$, which it is the $100\,\tau$ quantile of the set $\{\hat{k}_{(1)}, \hat{k}_{(2)}, \ldots, \hat{k}_{(p)}\}$; therefore, the new quantile estimator is given as:

$$KQ_\tau = \{\hat{k}_{(i)}\}_\tau = \{\hat{k}_{(1)}, \hat{k}_{(2)}, \ldots, \hat{k}_{(p)}\}_\tau, \quad (17)$$

and

$$\Pr(\hat{k} < KQ_{\tau}) = \tau, \quad (18)$$

Now, considering three levels following Suhail et al. (2020) i.e., low level ( $\tau = 0.10$ and 0.25), moderate level ( $\tau = 0.50$ ) and high level ( $\tau = 0.75$, 0.90 and 0.95). So, the new six estimators are denoted as $\hat{k}_1^* = KQ_{0.10}$, $\hat{k}_2^* = KQ_{0.25}$, $\hat{k}_3^* = KQ_{0.50}$, $\hat{k}_4^* = KQ_{0.75}$, $\hat{k}_5^* = KQ_{0.90}$ and $\hat{k}_6^* = KQ_{0.95}$. It is noted that when $\tau$ equals 0.50 in equation (18), we get equation (15), that means $\hat{k}_7$ estimator is special case of the new quantile approach.

## 4. Applications
### 4.1 Simulation Study

The impossibility of performing the theoretical comparisons among estimators in this case gives us a motivation to conduct a massive simulation study using various factors to give a useful view about the biasing parameter new defined estimators' performances in the KLE. As known, the MSE criterion is used for measuring the new defined estimators effects. The MATLAB software is used for the computational procedures. By following the authors (Gibbons 1981) and (Kibria 2003), explanatory variables are generated using:

$$x_{ji} = (1-\kappa^2)^{1/2} z_{ji} + \kappa z_{j,p+1}, \qquad j = 1,2,\ldots,n, \quad i = 1,2,\ldots,p \quad (19)$$

where $z_{ji}$ are known as the independent numbers that distributed as a standard normal and here $\kappa = 0.80, 0.90$ and 0.99 that is defined as any two explanatory variables correlation degree. Considering $p = 3$ and $p = 7$ in simulation where the variables are standardized. The response variable $y_i$ are defined as:

$$y_j = b_1 x_{j1} + b_2 x_{j2} + \text{K} + b_p x_{jp} + c_j, \quad j = 1,2,\text{K}, n \quad (20)$$

where $c_j$ are known as an $i.i.d\ N(0,\sigma^2)$. And $b'b = 1$ as in Dawoud and Abonazel (2021), Lukman et al (2021), Algamal and Abonazel (2021), Awwad et al. (2021), and Abonazel et al. (2022a,b). The number of replications in the performed simulation is 2000 times with $n = 50$, 100 and 150 and $\sigma^2 = 1$, 25, and 100. The estimated MSE (EMSE) of estimators are measured in each replicate and computed as:

$$EMSE(\alpha^*) = \frac{1}{2000} \sum_{i=1}^{2000} (\alpha_i^* - \alpha)'(\alpha_i^* - \alpha) \quad (21)$$

where $\alpha^*$ is an estimator of the true parameter is $\alpha$.

We have the following comments due to Tables 1 – 6 results:

1. In case of the factors $\sigma$, $\kappa$ and $p$ have an increase, the EMSEs of the estimators have an increase but our suggested estimator with higher quantiles in special has the least increase while the EMSEs of the estimators have a decrease in case of the sample size $n$ has an increase.
2. As expected, the LSE gives the highest EMSEs in multicollinearity existence.
3. Generally, the KLE with new quantile estimators of biasing parameter has better performance by giving lower EMSEs than LSE and other available estimators. So, our suggested new quantile estimator has a good impact on KLE performance.
4. The new quantile estimator has a better performance in high degrees of collinearity with high variance of errors.
5. The estimators $\hat{k}_5$, $\hat{k}_6$, $\hat{k}_3^*$, $\hat{k}_4^*$, $\hat{k}_5^*$ and $\hat{k}_6^*$ for the KLE have less EMSEs than others in general and specially $\hat{k}_4^*$ has the best EMSEs in many different values of factors.
6. The KLE with estimators $\hat{k}_5$, $\hat{k}_6$, $\hat{k}_3^*$, $\hat{k}_4^*$, $\hat{k}_5^*$ and $\hat{k}_6^*$ has better performance than with others for $p = 3$ and different values of $\sigma$, $n$, and $\kappa$.
7. The KLE with estimators $\hat{k}_5$, $\hat{k}_6$, $\hat{k}_3^*$ and $\hat{k}_4^*$ has better performance than with others especially with $\hat{k}_4^*$ for $p = 7$ and different values of $\sigma$, $n$, and $\kappa$.
8. The performance of KLE with $\hat{k}_1$ is approximately the same with $\hat{k}_1^*$.

**Table 1.** EMSEs of KLE when $n = 50$ and $p = 3$

| Estimator | $\kappa = 0.80$ | | | $\kappa = 0.90$ | | | $\kappa = 0.99$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | $\sigma = 1$ | $\sigma = 5$ | $\sigma = 10$ | $\sigma = 1$ | $\sigma = 5$ | $\sigma = 10$ | $\sigma = 1$ | $\sigma = 5$ | $\sigma = 10$ |
| **LSE** | 0.1207 | 3.0183 | 12.0732 | 0.2178 | 5.4443 | 21.777 | 1.9842 | 49.6047 | 198.419 |
| $KLE\,(\hat{k}_1)$ | 0.1055 | 1.6834 | 6.0367 | 0.1758 | 2.732 | 10.1093 | 1.0881 | 21.0614 | 83.1397 |
| $KLE\,(\hat{k}_2)$ | 0.0956 | 1.1239 | 3.96 | 0.1457 | 1.7372 | 6.4818 | 0.6546 | 13.1115 | 52.043 |
| $KLE\,(\hat{k}_3)$ | 0.1198 | 2.9136 | 11.609 | 0.2137 | 5.1478 | 20.5337 | 1.6966 | 42.5185 | 169.929 |
| $KLE\,(\hat{k}_4)$ | 0.0854 | 2.0777 | 7.8941 | 0.1211 | 4.0378 | 15.3956 | 1.4512 | 36.0104 | 139.213 |
| $KLE\,(\hat{k}_5)$ | 0.0742 | 0.6641 | 2.2657 | 0.0813 | 1.0806 | 3.9868 | 0.4732 | 10.081 | 40.3855 |
| $KLE(\hat{k}_6)$ | 0.0844 | 0.811 | 2.938 | 0.109 | 1.2313 | 4.7528 | 0.4182 | 9.5538 | 38.2458 |
| $KLE\,(\hat{k}_1^*)$ | 0.1055 | 1.6834 | 6.0367 | 0.1758 | 2.732 | 10.1093 | 1.0881 | 21.0614 | 83.1397 |
| $KLE\,(\hat{k}_2^*)$ | 0.1008 | 1.2987 | 4.891 | 0.1577 | 2.1318 | 8.3775 | 0.7647 | 17.9249 | 71.6266 |
| $KLE\,(\hat{k}_3^*)$ | 0.0895 | 0.7834 | 3.1026 | 0.1213 | 1.2715 | 5.472 | 0.3546 | 12.0614 | 49.1073 |
| $KLE\,(\hat{k}_4^*)$ | 0.0762 | 0.6542 | 1.8591 | 0.0773 | 1.0454 | 3.4567 | 0.4733 | 8.6962 | 34.3459 |
| $KLE\,(\hat{k}_5^*)$ | 0.0824 | 0.7178 | 1.9393 | 0.086 | 1.1245 | 3.6098 | 0.5044 | 8.9102 | 35.1155 |
| $KLE\,(\hat{k}_6^*)$ | 0.0824 | 0.7178 | 1.9393 | 0.086 | 1.1245 | 3.6098 | 0.5044 | 8.9102 | 35.1155 |

**Table 2:** EMSEs of KLE when $n = 100$ and $p = 3$

| Estimator | $\kappa = 0.80$ | | | $\kappa = 0.90$ | | | $\kappa = 0.99$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | $\sigma = 1$ | $\sigma = 5$ | $\sigma = 10$ | $\sigma = 1$ | $\sigma = 5$ | $\sigma = 10$ | $\sigma = 1$ | $\sigma = 5$ | $\sigma = 10$ |
| **LSE** | 0.059 | 1.4753 | 5.9012 | 0.1081 | 2.7016 | 10.8063 | 1.0084 | 25.2105 | 100.842 |
| $KLE\,(\hat{k}_1)$ | 0.0543 | 0.8985 | 3.02 | 0.094 | 1.454 | 5.0891 | 0.6262 | 10.9583 | 42.932 |
| $KLE\,(\hat{k}_2)$ | 0.0516 | 0.6016 | 1.9618 | 0.0846 | 0.9079 | 3.2016 | 0.3955 | 6.7285 | 26.5124 |
| $KLE\,(\hat{k}_3)$ | 0.0589 | 1.4529 | 5.7877 | 0.1077 | 2.6316 | 10.4895 | 0.94 | 23.0052 | 92.0974 |
| $KLE\,(\hat{k}_4)$ | 0.053 | 1.161 | 2.9301 | 0.0833 | 1.3464 | 4.669 | 0.7079 | 13.7723 | 52.7101 |
| $KLE\,(\hat{k}_5)$ | 0.0473 | 0.367 | 1.175 | 0.0619 | 0.6086 | 2.1938 | 0.2597 | 5.7313 | 22.1847 |
| $KLE(\hat{k}_6)$ | 0.0499 | 0.4053 | 1.3841 | 0.0757 | 0.5876 | 2.1994 | 0.2232 | 4.6758 | 18.5959 |
| $KLE\,(\hat{k}_1^*)$ | 0.0543 | 0.8985 | 3.02 | 0.094 | 1.454 | 5.0891 | 0.6262 | 10.9583 | 42.932 |
| $KLE\,(\hat{k}_2^*)$ | 0.0533 | 0.6731 | 2.4134 | 0.0898 | 1.0835 | 4.153 | 0.4371 | 9.2154 | 36.8403 |
| $KLE\,(\hat{k}_3^*)$ | 0.0506 | 0.3986 | 1.5506 | 0.0791 | 0.6094 | 2.7159 | 0.2131 | 6.1927 | 25.4408 |
| $KLE\,(\hat{k}_4^*)$ | 0.042 | 0.4333 | 1.2197 | 0.0462 | 0.6967 | 2.3019 | 0.2647 | 5.5923 | 21.4297 |
| $KLE\,(\hat{k}_5^*)$ | 0.04 | 0.4929 | 1.3408 | 0.0419 | 0.7726 | 2.4915 | 0.2844 | 5.8779 | 22.5878 |
| $KLE\,(\hat{k}_6^*)$ | 0.04 | 0.4929 | 1.3408 | 0.0419 | 0.7726 | 2.4915 | 0.2844 | 5.8779 | 22.5878 |

**Table 3:** EMSEs of KLE when $n = 150$ and $p = 3$

| Estimator | $\kappa = 0.80$ | | | $\kappa = 0.90$ | | | $\kappa = 0.99$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | $\sigma = 1$ | $\sigma = 5$ | $\sigma = 10$ | $\sigma = 1$ | $\sigma = 5$ | $\sigma = 10$ | $\sigma = 1$ | $\sigma = 5$ | $\sigma = 10$ |
| **LSE** | 0.0416 | 1.04 | 4.1602 | 0.076 | 1.9005 | 7.6019 | 0.7053 | 17.6326 | 70.5303 |
| $KLE(\hat{k}_1)$ | 0.039 | 0.6705 | 2.194 | 0.0681 | 1.0768 | 3.6383 | 0.4645 | 7.6395 | 29.5464 |
| $KLE(\hat{k}_2)$ | 0.0376 | 0.4558 | 1.4089 | 0.0631 | 0.6678 | 2.2464 | 0.304 | 4.6004 | 17.9934 |
| $KLE(\hat{k}_3)$ | 0.0416 | 1.03 | 4.1044 | 0.0759 | 1.8666 | 7.4415 | 0.6733 | 16.452 | 65.7495 |
| $KLE(\hat{k}_4)$ | 0.0364 | 0.5373 | 2.1876 | 0.0495 | 0.9334 | 4.084 | 0.6105 | 13.4849 | 60.4232 |
| $KLE(\hat{k}_5)$ | 0.0362 | 0.281 | 0.8576 | 0.0536 | 0.4515 | 1.5404 | 0.1867 | 3.9647 | 15.4788 |
| $KLE(\hat{k}_6)$ | 0.037 | 0.3029 | 0.9806 | 0.0595 | 0.4209 | 1.5234 | 0.1642 | 3.174 | 12.5563 |
| $KLE(\hat{k}_1^*)$ | 0.039 | 0.6705 | 2.194 | 0.0681 | 1.0768 | 3.6383 | 0.4645 | 7.6395 | 29.5464 |
| $KLE(\hat{k}_2^*)$ | 0.0385 | 0.5021 | 1.7074 | 0.0661 | 0.7812 | 2.902 | 0.3325 | 6.3379 | 25.2711 |
| $KLE(\hat{k}_3^*)$ | 0.0372 | 0.2933 | 1.0377 | 0.0608 | 0.4051 | 1.805 | 0.1667 | 4.1532 | 17.3193 |
| $KLE(\hat{k}_4^*)$ | 0.0339 | 0.3499 | 0.9202 | 0.0439 | 0.5388 | 1.6354 | 0.187 | 3.9197 | 15.0196 |
| $KLE(\hat{k}_5^*)$ | 0.0329 | 0.4068 | 1.0219 | 0.0405 | 0.608 | 1.7761 | 0.2015 | 4.1154 | 15.7884 |
| $KLE(\hat{k}_6^*)$ | 0.0329 | 0.4068 | 1.0219 | 0.0405 | 0.608 | 1.7761 | 0.2015 | 4.1154 | 15.7884 |

**Table 4.** EMSEs of KLE when $n = 50$ and $p = 7$

| Estimator | $\kappa = 0.80$ | | | $\kappa = 0.90$ | | | $\kappa = 0.99$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | $\sigma = 1$ | $\sigma = 5$ | $\sigma = 10$ | $\sigma = 1$ | $\sigma = 5$ | $\sigma = 10$ | $\sigma = 1$ | $\sigma = 5$ | $\sigma = 10$ |
| **LSE** | 0.4032 | 10.0789 | 40.3155 | 0.7581 | 18.9517 | 75.8069 | 7.184 | 179.6 | 718.401 |
| $KLE(\hat{k}_1)$ | 0.3038 | 5.3179 | 20.411 | 0.5233 | 9.7166 | 37.942 | 3.847 | 89.2727 | 356.153 |
| $KLE(\hat{k}_2)$ | 0.2061 | 2.5539 | 9.7585 | 0.3079 | 4.5992 | 17.9422 | 1.8164 | 41.7882 | 166.697 |
| $KLE(\hat{k}_3)$ | 0.3943 | 9.5387 | 38.0572 | 0.7254 | 17.4924 | 69.8365 | 5.9019 | 146.66 | 586.606 |
| $KLE(\hat{k}_4)$ | 0.27 | 4.8755 | 16.1642 | 0.5228 | 7.9308 | 31.073 | 4.2861 | 100.554 | 401.586 |
| $KLE(\hat{k}_5)$ | 0.0936 | 1.3676 | 5.3384 | 0.1425 | 2.8826 | 11.2967 | 1.2941 | 30.904 | 122.625 |
| $KLE(\hat{k}_6)$ | 0.1391 | 1.5747 | 6.0901 | 0.1868 | 2.7985 | 11.0081 | 1.0748 | 25.2713 | 101.007 |
| $KLE(\hat{k}_1^*)$ | 0.2942 | 4.9019 | 18.9366 | 0.498 | 8.9762 | 35.2192 | 3.5307 | 82.8656 | 330.874 |
| $KLE(\hat{k}_2^*)$ | 0.2469 | 3.4079 | 13.305 | 0.3824 | 6.2469 | 24.6847 | 2.4217 | 57.8447 | 231.077 |
| $KLE(\hat{k}_3^*)$ | 0.1572 | 1.7915 | 6.9952 | 0.205 | 3.2398 | 12.8334 | 1.2381 | 29.6706 | 118.416 |
| $KLE(\hat{k}_4^*)$ | 0.0715 | 1.0233 | 3.9791 | 0.0969 | 1.8178 | 7.2272 | 0.6842 | 16.617 | 66.5108 |
| $KLE(\hat{k}_5^*)$ | 0.1387 | 2.5745 | 9.9069 | 0.2128 | 5.0493 | 19.8595 | 2.0193 | 49.6915 | 198.621 |
| $KLE(\hat{k}_6^*)$ | 0.1595 | 2.8192 | 10.7976 | 0.2374 | 5.4589 | 21.4393 | 2.1611 | 53.2063 | 212.755 |

**Table 5:** EMSEs of KLE when $n=100$ and $p=7$

| Estimator | $\kappa=0.80$ | | | $\kappa=0.90$ | | | $\kappa=0.99$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | $\sigma=1$ | $\sigma=5$ | $\sigma=10$ | $\sigma=1$ | $\sigma=5$ | $\sigma=10$ | $\sigma=1$ | $\sigma=5$ | $\sigma=10$ |
| **LSE** | 0.175 | 4.3749 | 17.4996 | 0.3291 | 8.2277 | 32.9108 | 3.1258 | 78.1443 | 312.577 |
| $KLE(\hat{k}_1)$ | 0.1515 | 2.7016 | 10.0451 | 0.2677 | 4.8433 | 18.4594 | 1.9686 | 43.0603 | 171.053 |
| $KLE(\hat{k}_2)$ | 0.1274 | 1.3373 | 4.7852 | 0.1996 | 2.2993 | 8.6295 | 0.9706 | 19.8569 | 78.8527 |
| $KLE(\hat{k}_3)$ | 0.1743 | 4.2861 | 17.1022 | 0.326 | 7.9699 | 31.8147 | 2.8586 | 70.4645 | 281.792 |
| $KLE(\hat{k}_4)$ | 0.1017 | 1.4707 | 6.1501 | 0.1424 | 3.2936 | 12.8927 | 2.0573 | 45.4045 | 181.911 |
| $KLE(\hat{k}_5)$ | 0.0842 | 0.512 | 1.8438 | 0.0976 | 1.0552 | 3.9042 | 0.5111 | 11.4657 | 45.8617 |
| $KLE(\hat{k}_6)$ | 0.1112 | 0.7819 | 2.8027 | 0.1496 | 1.2971 | 4.9029 | 0.5299 | 10.9901 | 43.6567 |
| $KLE(\hat{k}_1^*)$ | 0.1494 | 2.4856 | 9.2574 | 0.2612 | 4.4414 | 17.0326 | 1.8123 | 39.8263 | 158.569 |
| $KLE(\hat{k}_2^*)$ | 0.1388 | 1.6724 | 6.227 | 0.2285 | 2.9551 | 11.4262 | 1.2083 | 26.6638 | 106.445 |
| $KLE(\hat{k}_3^*)$ | 0.1176 | 0.7754 | 2.7968 | 0.1623 | 1.3077 | 5.0048 | 0.5469 | 11.4372 | 45.7652 |
| $KLE(\hat{k}_4^*)$ | 0.0752 | 0.2823 | 1.0154 | 0.0639 | 0.4477 | 1.7092 | 0.169 | 3.7174 | 14.8952 |
| $KLE(\hat{k}_5^*)$ | 0.0467 | 0.8163 | 2.9662 | 0.0664 | 1.6232 | 6.2577 | 0.6442 | 16.1903 | 64.7467 |
| $KLE(\hat{k}_6^*)$ | 0.0479 | 0.9239 | 3.32 | 0.0742 | 1.7853 | 6.8681 | 0.6983 | 17.4954 | 69.9776 |

**Table 6:** EMSEs of KLE when $n=150$ and $p=7$

| Estimator | $\kappa=0.80$ | | | $\kappa=0.90$ | | | $\kappa=0.99$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | $\sigma=1$ | $\sigma=5$ | $\sigma=10$ | $\sigma=1$ | $\sigma=5$ | $\sigma=10$ | $\sigma=1$ | $\sigma=5$ | $\sigma=10$ |
| **LSE** | 0.1105 | 2.7613 | 11.0451 | 0.2079 | 5.1972 | 20.7889 | 1.9753 | 49.3826 | 197.53 |
| $KLE(\hat{k}_1)$ | 0.0991 | 1.7669 | 6.4172 | 0.1769 | 3.1368 | 11.7323 | 1.3015 | 27.2287 | 107.865 |
| $KLE(\hat{k}_2)$ | 0.0888 | 0.8954 | 3.0245 | 0.1445 | 1.4937 | 5.4013 | 0.6693 | 12.3433 | 48.7847 |
| $KLE(\hat{k}_3)$ | 0.1103 | 2.7276 | 10.886 | 0.2071 | 5.0946 | 20.3396 | 1.8682 | 45.9118 | 183.53 |
| $KLE(\hat{k}_4)$ | 0.0887 | 2.6194 | 9.2518 | 0.123 | 4.2014 | 10.6091 | 1.6023 | 28.3791 | 111.585 |
| $KLE(\hat{k}_5)$ | 0.0727 | 0.3461 | 1.2185 | 0.0868 | 0.6921 | 2.6147 | 0.3413 | 7.8224 | 31.1758 |
| $KLE(\hat{k}_6)$ | 0.0832 | 0.5232 | 1.7679 | 0.1215 | 0.8418 | 3.0616 | 0.3674 | 6.8039 | 26.9145 |
| $KLE(\hat{k}_1^*)$ | 0.0983 | 1.6328 | 5.8862 | 0.1741 | 2.8716 | 10.7772 | 1.2086 | 25.0984 | 99.7006 |
| $KLE(\hat{k}_2^*)$ | 0.094 | 1.1102 | 3.9478 | 0.16 | 1.909 | 7.2138 | 0.8265 | 16.7789 | 66.7259 |
| $KLE(\hat{k}_3^*)$ | 0.0856 | 0.5223 | 1.7832 | 0.1301 | 0.8614 | 3.1683 | 0.3736 | 7.285 | 28.9814 |
| $KLE(\hat{k}_4^*)$ | 0.0682 | 0.17 | 0.5692 | 0.0714 | 0.2553 | 0.9291 | 0.1023 | 1.9627 | 7.7578 |
| $KLE(\hat{k}_5^*)$ | 0.0444 | 0.5533 | 1.9507 | 0.0401 | 1.0504 | 4.0607 | 0.3982 | 10.5532 | 42.2238 |
| $KLE(\hat{k}_6^*)$ | 0.0422 | 0.6303 | 2.1856 | 0.0408 | 1.1614 | 4.4462 | 0.4318 | 11.3382 | 45.3725 |

## 4.2 Real-Life data: Portland cement data

In this section, we used the Portland cement data that adopted by Woods et al. (1932) to show the performance of the new suggested quantile estimator for the biasing parameter for the KLE. Many authors have used this data such as (Kaciranlar et al.

1999); (Li and Yang 2012); (Lukman et al. 2019), (Kibria and Lukman 2020), (Dawoud and Kibria 2020) and (Dawoud 2021a). The model of this data is written as

$$y_j = b_1 x_1 + b_2 x_2 + b_3 x_3 + b_4 x_4 + c_j, \quad j = 1, 2, \ldots, 13. \tag{22}$$

To learn about this data and its variables, see (Woods et al. 1932). Some of the measures are computed to clarify the multicollinearity happen as the variance inflation factors (VIFs) in which their values are as: 38.50, 254.42, 46.87 and 282.51. And the eigenvalues of $X'X$ are $t_1 = 44676.206$, $t_2 = 5965.422$, $t_3 = 809.952$, $t_4 = 105.419$, as well the condition number (CN) of $X'X$ is near to 20.58. The VIFs, eigenvalues, and CN tell us a severe multicollinearity happen in the data. In Table 7, we state the estimated parameter with the estimators' EMSEs. Table 7 appears that the available estimator $\hat{k}_1$ and the suggested new quantile estimator $\hat{k}_1^*$ for the KLE have the same performance and they outperform better than others for this data.

**Table 7:** SMSEs of LSE and KLE using different biasing estimators

| Estimator | $\hat{b}_1$ | $\hat{b}_2$ | $\hat{b}_3$ | $\hat{b}_4$ | $\hat{k}$ | $SMSE$ |
|---|---|---|---|---|---|---|
| **LSE** | 2.1930 | 1.1533 | 0.7585 | 0.4863 | ----- | 0.0638 |
| $KL(\hat{k}_1)$ | 2.1764 | 1.1572 | 0.7465 | 0.4888 | 0.6042 | **0.0629** |
| $KL(\hat{k}_2)$ | 2.1473 | 1.1639 | 0.7257 | 0.4932 | 1.6741 | 0.0636 |
| $KL(\hat{k}_3)$ | 2.1930 | 1.1533 | 0.7585 | 0.4863 | 0.00075 | 0.0638 |
| $KL(\hat{k}_4)$ | 1.8960 | 1.2215 | 0.5475 | 0.5308 | 11.8634 | 0.1826 |
| $KL(\hat{k}_5)$ | 2.0590 | 1.1842 | 0.6626 | 0.5065 | 5.0564 | 0.0824 |
| $KL(\hat{k}_6)$ | 2.1116 | 1.1721 | 0.7001 | 0.4986 | 3.0186 | 0.0682 |
| $KL(\hat{k}_1^*)$ | 2.1764 | 1.1572 | 0.7465 | 0.4888 | 0.6042 | **0.0629** |
| $KL(\hat{k}_2^*)$ | 2.1548 | 1.1622 | 0.7310 | 0.4921 | 1.3978 | 0.0632 |
| $KL(\hat{k}_3^*)$ | 2.0955 | 1.1758 | 0.6886 | 0.5010 | 3.6335 | 0.0716 |
| $KL(\hat{k}_4^*)$ | 1.9691 | 1.2048 | 0.5989 | 0.5199 | 8.7151 | 0.1272 |
| $KL(\hat{k}_5^*)$ | 1.8849 | 1.2240 | 0.5398 | 0.5324 | 12.3545 | 0.1924 |
| $KL(\hat{k}_6^*)$ | 1.8849 | 1.2240 | 0.5398 | 0.5324 | 12.3545 | 0.1924 |

## 5. Conclusions

In regression model, the performance of KLE depends on the biasing parameter determination. We propose a new quantile estimator of the biasing parameter for the KLE. Our suggested quantile estimator with a useful quantile level choice has a better performance than others with using different values of factors, especially in severe multicollinearity and from moderate to high variance of the error. Within different available and suggested estimators of the biasing parameter, we have focused on giving the one/ones outperform better by using a simulation study and a real data. Then, we have investigated the performance of the suggested quantile estimator with different quantile levels by giving various correlation levels between the explanatory variables, various variance of error, various sample sizes, various numbers of explanatory variables in our simulation study. In general, the KLE has a better performance with the suggested quantile estimator of the biasing parameter. The available $\hat{k}_5$, $\hat{k}_6$ and the suggested $\hat{k}_3^*$, $\hat{k}_4^*$, $\hat{k}_5^*$, $\hat{k}_6^*$ estimators of the biasing parameter in the KLE are better than others in various different cases in the simulation study. Also, the KLE has a better performance than the LSE through various available and suggested estimators and it has the least EMSEs with the available $\hat{k}_1$ and the suggested $\hat{k}_1^*$ in the real data. Finally, the KLE with the available $\hat{k}_1$, $\hat{k}_5$, $\hat{k}_6$ and the suggested $\hat{k}_1^*$, $\hat{k}_3^*$, $\hat{k}_4^*$, $\hat{k}_5^*$, $\hat{k}_6^*$ is highly recommended to the practitioners.

### References


1. Abonazel, M.R., Algamal, Z.Y., Awwad, F.A., and Taha, I.M. 2022a. A New Two-Parameter Estimator for Beta Regression Model: Method, Simulation, and Application. Frontiers in Applied Mathematics and Statistics 7:780322. doi: 10.3389/fams.2021.780322.

2. Abonazel, M.R., Dawoud, I., Awwad, F.A., and Lukman, A.F. 2022b. Dawoud–Kibria Estimator for Beta Regression Model: Simulation and Application. Frontiers in Applied Mathematics and Statistics. doi: 10.3389/fams.2022.775068.
3. Abonazel, M.R., and Farghali, R.A. 2019. Liu-type multinomial logistic estimator. Sankhya B 81(2): 203-225.
4. Abonazel, M.R., and Taha, I.M. 2021. Beta ridge regression estimators: simulation and application. Communications in Statistics-Simulation and Computation, 1-13.
5. Algamal, Z. Y., and Abonazel, M. R. 2021. Developing a Liu-type estimator in beta regression model. *Concurrency and Computation: Practice and Experience*, e6685.
6. 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*, e6779.
7. Dawoud, I., and 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.
8. Alkhamisi, M., Khalaf, G., and Shukur, G. 2006. Some modifications for choosing ridge parameters. *Communications in Statistics-Theory and Methods* 35(11): 2005–2020.
9. Alkhamisi, M., and Shukur, G. 2008. Developing ridge parameters for SUR model. *Communications in Statistics-Theory and Methods* 37(4): 544–564.
10. Dawoud, I., and Kibria, B.M.G. 2020. A new biased estimator to combat the multicollinearity of the Gaussian linear regression model. *Stats* 3(4): 526-541.
11. Dawoud, I. 2021a. Performance of Some New Biasing Parameter Estimates of Kibria-Lukman Estimator in Linear Regression Model. *Al-Aqsa University Scientific Journal* In appear.
12. Dawoud, I. 2021b. A General Form of the Ridge-Type Regression Estimator. *Communications in Statistics - Simulation and Computation* Under review.
13. Gibbons. D.G. 1981. A simulation study of some ridge estimators. *Journal of the American Statistical Association* 76(373): 131–139.
14. Hocking, R.R., Speed, F.M., and Lynn, M.J. 1976. A class of biased estimators in linear regression. *Technometrics* 18(4): 425–437.
15. Hoerl, A.E., and Kennard, R.W. 1970. Ridge regression: biased estimation for nonorthogonal problems. *Technometrics* 12(1): 55–67.
16. Hoerl, A.E., Kannard, R.W., and Baldwin, K.F. 1975. Ridge regression: some simulations. *Communications in Statistics* 4(2):105–123.
17. Kaciranlar, S., Sakallioglu, S., Akdeniz, F., Styan, G.P.H., and Werner, H.J. 1999. A new biased estimator in linear regression and a detailed analysis of the widely-analysed dataset on portland cement. *Sankhya: The Indian Journal of Statistics, Series B* 61: 443–459.
18. Khalaf, G., and Shukur, G. 2005. Choosing ridge parameter for regression problems. *Communications in Statistics Theory and Methods* 34(5): 1177–1182.
19. Kibria, B.M.G. 2003. Performance of some new ridge regression estimators. *Communications in Statistics-Simulation and Computation* 32(1): 419–435.
20. Kibria, B.M.G., and Lukman, A.F. 2020. A New Ridge-Type Estimator for the Linear Regression Model: Simulations and Applications. *Hindawi, Scientifica* Article ID 9758378: 16 pages.
21. Lawless, J.F., and Wang, P. 1976. A simulation study of ridge and other regression estimators. *Communications in Statistics Theory and Methods* 5(4): 307–323.
22. Li, Y., and Yang, H. 2012. Anew Liu-type estimator in linear regression model. Statistical Papers 53(2): 427–437.
23. Lukman, A.F., Aladeitan, B., Ayinde, K., and Abonazel, M.R. 2021. Modified ridge-type for the Poisson regression model: simulation and application. *Journal of Applied Statistics*, 1-13.
24. Lukman, A.F., Ayinde, K., Binuomote, S., and Clement O.A. 2019. Modified ridge-type estimator to combat multicollinearity: application to chemical data. *Journal of Chemometrics* 33(5):p. e3125.
25. Muniz, G., and Kibria, B.M.G. 2009. On some ridge regression estimators: An empirical comparisons. *Communications in Statistics-Simulation and Computation* 38 (3): 621–630.
26. Muniz, G., Kibria, B.M.G., Shukur, G., and Mansson, K. 2012. On developing ridge regression parameters: a graphical investigation. *SORT* 36(2): 115–138.
27. Suhail, M., Chand, S., and Kibria, B.M.G. 2020. Quantile based estimation of biasing parameters in ridge regression model. *Communications in Statistics - Simulation and Computation* 49(10): 2732-2744
28. Ullah, M.I., Aslam, M., Altaf, S., and Ahmed, M. 2019. Some New Diagnostics of Multicollinearity in Linear Regression Model. *Sains Malaysiana* 48(9): 2051–2060.
29. Woods, H., Steinour, H.H., and Starke, H.R. 1932. Effect of composition of Portland cement on heat evolved during hardening. *Industrial & Engineering Chemistry* 24(11): 1207–1214.