Predictive Performance Evaluation of the Kibria-Lukman Estimator

ISSAM DAWOUD[1], MOHAMED R. ABONAZEL[2], ELSAYED TAG ELDIN[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]Electrical Engineering Department, Faculty of Engineering & Technology, Future University in Egypt, New Cairo, EGYPT

Abstract: - Regression models are commonly used in prediction, but their predictive performances may be affected by the problem called the multicollinearity. To reduce the effect of the multicollinearity, different biased estimators have been proposed as alternatives to the ordinary least squares estimator. But there are still little analyses of the different proposed biased estimators' predictive performances. Therefore, this paper focuses on discussing the predictive performance of the recently proposed "new ridge-type estimator", namely the Kibria-Lukman (KL) estimator. The theoretical comparisons among the predictors of these estimators are done according to the prediction mean squared error criterion in the two-dimensional space and the results are explained by a numerical example. The regions are determined where the KL estimator gives better results than the other estimators.

Key-Words: - Biased Estimator, Ridge Estimator, Liu Estimator, Kibria-Lukman estimator, Prediction Mean Square Error, Multicollinearity.

Received: July 28, 2021. Revised: June 29, 2022. Accepted: August 11, 2022. Published: September 20, 2022.

1 Introduction

The multiple linear regression model is given by

$$m = X\beta + \varepsilon, \qquad (1)$$

where m is an $n \times 1$ vector of dependent variable, β is a $p\times 1$ vector of unknown parameters, X is an $n\times p$ full column rank matrix of non-stochastic predetermined regressors, and ε is an $n\times 1$ vector of $i.i.d.(0,\sigma^2)$ random errors.
The Ordinary Least Squares (OLS) estimator of the unknown parameters in (1) is given by

$$\hat{\beta}_{OLS} = (X'X)^{-1}X'm. \qquad (2)$$

To reduce the effect of multicollinearity problem, Hoerl and Kennard [1] proposed the most common estimator which is called the ordinary ridge regression (ORR) estimator and is defined as follows:

$$\hat{\beta}_k = (X'X + k\mathrm{I})^{-1}X'm, \quad k > 0 \qquad (3)$$

where k is the biasing parameter.

Then, Liu [2] proposed another alternative biased estimator called the Liu estimator and is defined as follows:

$$\hat{\beta}_d = (X'X + \mathrm{I})^{-1}(X'X + d\mathrm{I})\hat{\beta}_{OLS}, \quad 0 < d < 1 \qquad (4)$$

where d is the biasing parameter.
Recently, Kibria and Lukman [3] proposed a new one parameter ridge-type estimator called the Kibria-Lukman (KL) estimator and is defined as

$$\hat{\beta}_{KL} = (X'X + k\mathrm{I})^{-1}(X'X - k\mathrm{I})\hat{\beta}_{OLS}. \quad k > 0 \qquad (5)$$

Since the predictive performance of the regression models which are commonly used in prediction is affected by the multicollinearity, different biased estimators have been proposed as an alternative to the ordinary least squares estimator to reduce its effect. But unfortunately there are few studies about the predictive performances of the biased estimators, as [4, 5, 6, 7, 8, 9, 10, 11].
As a consequence, it appears reasonable to evaluate the predictive performance of the recently proposed KL estimator compared with the OLS, ORR and Liu estimators. The rest of this article is organized as follows: In section 2, we present the evaluations of

the prediction mean squared error (PMSE). In section 3, the theoretical comparison of the PMSEs in the two dimensional space among the above mentioned estimators are given. A numerical example (an application) is given to demonstrate the theoretical results in section 4. Finally, some concluding remarks are given in section 5.

2 Evaluation of the Prediction Mean Squared Errors

We recall the developed PMSEs of Friedman and Montgomery [4] for the OLS and the ORR estimators and the developed PMSE of the Liu estimator given by [5] and then obtain the PMSE of the recently proposed KL estimator.
The PMSE is defined as:

$$J = E(m_0 - \hat{m}_0)^2 = Var + Bias^2, \tag{6}$$

where J is the PMSE, m_0 is the value to be predicted, $\hat{m}_0$ is the prediction of that value, (Var) is the variance and $(Bias^2)$ is the squared bias.
Now, the prediction error variance and bias are given as follows:

$$Var(m_0 - \hat{m}_0) = Var(m_0) + Var(\hat{m}_0) \tag{7}$$

and

$$Bias = E(m_0 - \hat{m}_0). \tag{8}$$

For convenience, the canonical form of model (1) is given by

$$m = Z\alpha + \varepsilon, \tag{9}$$

where $Z = XD$, $\alpha = D'\beta$. Here, D is an orthogonal matrix such that $Z'Z = D'X'XD = \Gamma = diag(\gamma_1, \gamma_2, ..., \gamma_p)$. Then the OLS estimator of α in model (9) is

$$\hat{\alpha}_{OLS} = \Gamma^{-1} Z' m. \tag{10}$$

The PMSE of the OLS estimator is given by

$$J_{OLS} = Var_{OLS} = \sigma^2 \left(1 + \sum_{i=1}^{p} \frac{z_{0i}^2}{\gamma_i}\right), \tag{11}$$

where z_0 is the orthonormalized point of the prediction $\hat{m}_0$.
The ORR estimator of α is defined by Hoerl and Kennard (1970) as follows:

$$\hat{\alpha}_k = (\Gamma + k\mathrm{I})^{-1} Z' m, \quad k > 0 \tag{12}$$

and then Friedman and Montgomery [4] found the PMSE of the ORR estimator as follows

$$J_k = \sigma^2 \left(1 + \sum_{i=1}^{p} \frac{z_{0i}^2 \gamma_i}{(\gamma_i + k)^2}\right) + k^2 \left(\sum_{i=1}^{p} \frac{z_{oi}\alpha_i}{(\gamma_i + k)}\right)^2. \tag{13}$$

The Liu estimator of α is defined by Liu [2] as follows:

$$\hat{\alpha}_d = (\Gamma + \mathrm{I})^{-1} (\Gamma + d\,\mathrm{I}) \hat{\alpha}_{OLS}, \; 0 < d < 1 \tag{14}$$

and then Özbey and Kaçıranlar [5] found the PMSE of the Liu estimator as follows:

$$J_d = \sigma^2 \left(1 + \sum_{i=1}^{p} \frac{z_{0i}^2 (\gamma_i + d)^2}{\gamma_i (\gamma_i + 1)^2}\right) + (1-d)^2 \left(\sum_{i=1}^{p} \frac{z_{oi}\alpha_i}{(\gamma_i + 1)}\right)^2. \tag{15}$$

The recently proposed NRT estimator of α is defined by Kibria and Lukman [3] as follows:

$$\hat{\alpha}_{NRT} = (\Gamma + k\mathrm{I})^{-1} (\Gamma - k\mathrm{I}) \hat{\alpha}_{OLS}, \; k > 0. \tag{16}$$

The NRT estimator has been extended in different regression models, such as [12, 13, 14, 15, 16].
The variance of the prediction error of the NRT estimator is

$$\begin{aligned} Var_{\mathrm{NRT}}(m_0 - \hat{m}_0) &= Var(m_0) + Var_{\mathrm{NRT}}(\hat{m}_0) \\ &= \sigma^2 + Var(z_0' \hat{\alpha}_{\mathrm{NRT}}) \\ &= \sigma^2 \left(1 + \sum_{i=1}^{p} \frac{(\gamma_i - k)^2 z_{0i}^2}{\gamma_i (\gamma_i + k)^2}\right). \end{aligned} \tag{17}$$

The bias of the prediction error of the NRT estimator is

$$Bias_{\mathrm{NRT}} = E(m_0 - \hat{m}_0) = z_0'\,\alpha - z_0'\,E(\hat{\alpha}_{\mathrm{NRT}}) = 2k\sum_{i=1}^{p}\frac{z_{oi}\alpha_i}{(\gamma_i + k)} \quad (18)$$

so, the squared bias is

$$Bias_{\mathrm{NRT}}^2 = 4k^2\left(\sum_{i=1}^{p}\frac{z_{oi}\alpha_i}{(\gamma_i + k)}\right)^2. \quad (19)$$

By summing up the variance and the squared bias of the NRT estimator we obtain

$$J_{\mathrm{NRT}} = Var_{\mathrm{NRT}} + Bias_{NRT}^2 = \sigma^2\left(1 + \sum_{i=1}^{p}\frac{(\gamma_i - k)^2 z_{0i}^2}{\gamma_i(\gamma_i + k)^2}\right) + 4k^2\left(\sum_{i=1}^{p}\frac{z_{oi}\alpha_i}{(\gamma_i + k)}\right)^2. \quad (20)$$

3 Comparisons of Prediction Mean Squared Errors in the Two Dimensional Space

We discuss here the prediction performance of the recently proposed KL estimator by following the method of [4, 5, 6, 7, 8, 9, 10, 11] such that our inferences are based on the predicted observations subspace (i.e., the ratio z_{02}^2 / z_{01}^2) and since the non-zero values of α_1^2 only increase the intercept values for J_k, J_d and J_{NRT} and leave the curve for J_{OLS} unchanged, we set α_1^2 to zero. So, the comparisons of J_{NRT} with J_{OLS}, J_k and J_d will be done and written in the following three theorems.

Theorem 1:

a. If $\alpha_2^2 > \dfrac{\sigma^2\left((\gamma_2 + k)^2 - (\gamma_2 - k)^2\right)}{4\gamma_2 k^2}$, then $J_{NRT} < J_{OLS} \Leftrightarrow \dfrac{z_{02}^2}{z_{01}^2} < f_1(\alpha_2^2)$.

b. If $\alpha_2^2 < \dfrac{\sigma^2\left((\gamma_2 + k)^2 - (\gamma_2 - k)^2\right)}{4\gamma_2 k^2}$, then $J_{NRT} < J_{OLS}$.

Where

$$f_1(\alpha_2^2) = \frac{\sigma^2\left(\dfrac{1}{\gamma_1} - \dfrac{(\gamma_1 - k)^2}{\gamma_1(\gamma_1 + k)^2}\right)}{\left(\dfrac{\sigma^2(\gamma_2 - k)^2}{\gamma_2(\gamma_2 + k)^2} + \dfrac{4k^2\alpha_2^2}{(\gamma_2 + k)^2} - \dfrac{\sigma^2}{\gamma_2}\right)}. \quad (21)$$

Proof:
The NRT estimator gives better results than the OLS estimator due to the PMSE criterion, when $J_{NRT} < J_{OLS}$. That means,

$$\sigma^2 + \sigma^2\left[\frac{(\gamma_1 - k)^2 z_{01}^2}{\gamma_1(\gamma_1 + k)^2} + \frac{(\gamma_2 - k)^2 z_{02}^2}{\gamma_2(\gamma_2 + k)^2}\right] + \frac{4k^2\alpha_2^2 z_{02}^2}{(\gamma_2 + k)^2} < \sigma^2 + \sigma^2\left(\frac{z_{01}^2}{\gamma_1} + \frac{z_{02}^2}{\gamma_2}\right). \quad (22)$$

After rearranging the inequality in (22), we get

$$z_{02}^2\left(\frac{\sigma^2(\gamma_2 - k)^2}{\gamma_2(\gamma_2 + k)^2} + \frac{4k^2\alpha_2^2}{(\gamma_2 + k)^2} - \frac{\sigma^2}{\gamma_2}\right) < z_{01}^2\sigma^2\left(\frac{1}{\gamma_1} - \frac{(\gamma_1 - k)^2}{\gamma_1(\gamma_1 + k)^2}\right). \quad (23)$$

If both

$$\frac{\sigma^2(\gamma_2 - k)^2}{\gamma_2(\gamma_2 + k)^2} + \frac{4k^2\alpha_2^2}{(\gamma_2 + k)^2} - \frac{\sigma^2}{\gamma_2} \quad (24)$$

and

$$\sigma^2\left(\frac{1}{\gamma_1} - \frac{(\gamma_1 - k)^2}{\gamma_1(\gamma_1 + k)^2}\right) \quad (25)$$

have the same signs, the NRT estimator gives better results than the OLS estimator when $z_{02}^2 / z_{01}^2 < f_1(\alpha_2^2)$ holds where

$$f_1(\alpha_2^2) = \frac{\sigma^2\left(\dfrac{1}{\gamma_1} - \dfrac{(\gamma_1 - k)^2}{\gamma_1(\gamma_1 + k)^2}\right)}{\left(\dfrac{\sigma^2(\gamma_2 - k)^2}{\gamma_2(\gamma_2 + k)^2} + \dfrac{4k^2\alpha_2^2}{(\gamma_2 + k)^2} - \dfrac{\sigma^2}{\gamma_2}\right)}. \quad (26)$$

Also, if (24) and (25) have opposite signs, the NRT estimator always gives better results than the OLS estimator where $f_1(\alpha_2^2)$ is negative and

$z_{02}^2 / z_{01}^2 > f_1(\alpha_2^2)$ always holds. Consequently, at that region the NRT estimator is better than the OLS estimator.
The positiveness condition of (24) is given by

$$\alpha_2^2 > \frac{\sigma^2\left((\gamma_2+k)^2-(\gamma_2-k)^2\right)}{4\gamma_2 k^2} \tag{27}$$

and equation (25) is always positive.
The hyperbola $f_1(\alpha_2^2)$ vertical asymptote is at the point

$$\alpha_2^2 = \frac{\sigma^2\left((\gamma_2+k)^2-(\gamma_2-k)^2\right)}{4\gamma_2 k^2}. \tag{28}$$

Theorem 2:

a. If $\alpha_2^2 > \frac{\sigma^2\left(\gamma_2^2-(\gamma_2-k)^2\right)}{3\gamma_2 k^2}$, then $J_{NRT} < J_k \Leftrightarrow \frac{z_{02}^2}{z_{01}^2} < f_2(\alpha_2^2)$.

b. If $\alpha_2^2 < \frac{\sigma^2\left(\gamma_2^2-(\gamma_2-k)^2\right)}{3\gamma_2 k^2}$, then $J_{NRT} < J_k$.

Where

$$f_2(\alpha_2^2) = \frac{\sigma^2\left(\frac{\gamma_1}{(\gamma_1+k)^2} - \frac{(\gamma_1-k)^2}{\gamma_1(\gamma_1+k)^2}\right)}{\left(\frac{\sigma^2(\gamma_2-k)^2}{\gamma_2(\gamma_2+k)^2} + \frac{4k^2\alpha_2^2}{(\gamma_2+k)^2} - \frac{\sigma^2\gamma_2}{(\gamma_2+k)^2} - \frac{k^2\alpha_2^2}{(\gamma_2+k)^2}\right)}. \tag{29}$$

Proof:
The NRT estimator gives better results than the ORR estimator due to the PMSE criterion, when $J_{NRT} < J_k$. That means,

$$\sigma^2 + \sigma^2\left[\frac{(\gamma_1-k)^2 z_{01}^2}{\gamma_1(\gamma_1+k)^2} + \frac{(\gamma_2-k)^2 z_{02}^2}{\gamma_2(\gamma_2+k)^2}\right] + \frac{4k^2\alpha_2^2 z_{02}^2}{(\gamma_2+k)^2} < \sigma^2 + \sigma^2\left(\frac{\gamma_1 z_{01}^2}{(\gamma_1+k)^2} + \frac{\gamma_2 z_{02}^2}{(\gamma_2+k)^2}\right) + \frac{k^2\alpha_2^2 z_{02}^2}{(\gamma_2+k)^2}. \tag{30}$$

After rearranging the inequality in (30), we get

$$z_{02}^2\left(\frac{\sigma^2(\gamma_2-k)^2}{\gamma_2(\gamma_2+k)^2} + \frac{4k^2\alpha_2^2}{(\gamma_2+k)^2} - \frac{\sigma^2\gamma_2}{(\gamma_2+k)^2} - \frac{k^2\alpha_2^2}{(\gamma_2+k)^2}\right) < z_{01}^2\sigma^2\left(\frac{\gamma_1}{(\gamma_1+k)^2} - \frac{(\gamma_1-k)^2}{\gamma_1(\gamma_1+k)^2}\right). \tag{31}$$

If both

$$\left(\frac{\sigma^2(\gamma_2-k)^2}{\gamma_2(\gamma_2+k)^2} + \frac{4k^2\alpha_2^2}{(\gamma_2+k)^2} - \frac{\sigma^2\gamma_2}{(\gamma_2+k)^2} - \frac{k^2\alpha_2^2}{(\gamma_2+k)^2}\right) \tag{32}$$

and

$$\sigma^2\left(\frac{\gamma_1}{(\gamma_1+k)^2} - \frac{(\gamma_1-k)^2}{\gamma_1(\gamma_1+k)^2}\right) \tag{33}$$

have the same signs, the NRT estimator gives better results than the ORR estimator when $z_{02}^2 / z_{01}^2 < f_2(\alpha_2^2)$ holds where

$$f_2(\alpha_2^2) = \frac{\sigma^2\left(\frac{\gamma_1}{(\gamma_1+k)^2} - \frac{(\gamma_1-k)^2}{\gamma_1(\gamma_1+k)^2}\right)}{\left(\frac{\sigma^2(\gamma_2-k)^2}{\gamma_2(\gamma_2+k)^2} + \frac{4k^2\alpha_2^2}{(\gamma_2+k)^2} - \frac{\sigma^2\gamma_2}{(\gamma_2+k)^2} - \frac{k^2\alpha_2^2}{(\gamma_2+k)^2}\right)}. \tag{34}$$

Also, if (32) and (33) have opposite signs, the NRT estimator always gives better results than the ORR estimator where $f_2(\alpha_2^2)$ is negative and

$z_{02}^2 / z_{01}^2 > f_2(\alpha_2^2)$ always holds. Consequently, at that region the NRT estimator is better than the ORR estimator.
The positiveness condition of (32) is given by

$$\alpha_2^2 > \frac{\sigma^2\left(\gamma_2^2 - (\gamma_2 - k)^2\right)}{3\gamma_2 k^2} \tag{35}$$

and the equation (33) is always positive.

The hyperbola $f_2(\alpha_2^2)$ vertical asymptote is at the point

$$\alpha_2^2 = \frac{\sigma^2\left(\gamma_2^2 - (\gamma_2 - k)^2\right)}{3\gamma_2 k^2}. \tag{36}$$

Theorem 3:
a. If

$$\alpha_2^2 > \frac{\sigma^2\left((\gamma_2 + d)^2(\gamma_2 + k)^2 - (\gamma_2 + 1)^2(\gamma_2 - k)^2\right)}{4k^2\gamma_2(\gamma_2 + 1)^2 - (1-d)^2\gamma_2(\gamma_2 + k)^2}$$

then

- $J_{NRT} < J_d$ for $(\gamma_1 + d)^2(\gamma_1 + k)^2 < (\gamma_1 + 1)^2(\gamma_1 - k)^2$,
- $J_{NRT} < J_d \Leftrightarrow \dfrac{z_{02}^2}{z_{01}^2} < f_3(\alpha_2^2)$ for $(\gamma_1 + d)^2(\gamma_1 + k)^2 > (\gamma_1 + 1)^2(\gamma_1 - k)^2$.

b. If

$$\alpha_2^2 < \frac{\sigma^2\left((\gamma_2 + d)^2(\gamma_2 + k)^2 - (\gamma_2 + 1)^2(\gamma_2 - k)^2\right)}{4k^2\gamma_2(\gamma_2 + 1)^2 - (1-d)^2\gamma_2(\gamma_2 + k)^2}$$

then

- $J_{NRT} < J_d$ for $(\gamma_1 + d)^2(\gamma_1 + k)^2 > (\gamma_1 + 1)^2(\gamma_1 - k)^2$,
- $J_{NRT} < J_d \Leftrightarrow \dfrac{z_{02}^2}{z_{01}^2} < f_3(\alpha_2^2)$ for $(\gamma_1 + d)^2(\gamma_1 + k)^2 < (\gamma_1 + 1)^2(\gamma_1 - k)^2$.

Where
$$f_3(\alpha_2^2) = \frac{\sigma^2\left(\dfrac{(\gamma_1 + d)^2}{\gamma_1(\gamma_1 + 1)^2} - \dfrac{(\gamma_1 - k)^2}{\gamma_1(\gamma_1 + k)^2}\right)}{\left(\dfrac{(\gamma_2 - k)^2\sigma^2}{\gamma_2(\gamma_2 + k)^2} + \dfrac{4k^2\alpha_2^2}{(\gamma_2 + k)^2} - \dfrac{(\gamma_2 + d)^2\sigma^2}{\gamma_2(\gamma_2 + 1)^2} - \dfrac{(1-d)^2\alpha_2^2}{(\gamma_2 + 1)^2}\right)}. \tag{37}$$

Proof:
The NRT estimator gives better results than the Liu estimator due to the PMSE criterion, when $J_{NRT} < J_d$. That means,

$$\sigma^2 + \sigma^2\left[\frac{(\gamma_1 - k)^2 z_{01}^2}{\gamma_1(\gamma_1 + k)^2} + \frac{(\gamma_2 - k)^2 z_{02}^2}{\gamma_2(\gamma_2 + k)^2}\right] + \frac{4k^2\alpha_2^2 z_{02}^2}{(\gamma_2 + k)^2} < \sigma^2 + \sigma^2\left(\frac{(\gamma_1 + d)^2 z_{01}^2}{\gamma_1(\gamma_1 + 1)^2} + \frac{(\gamma_2 + d)^2 z_{02}^2}{\gamma_2(\gamma_2 + 1)^2}\right) + \frac{(1-d)^2\alpha_2^2 z_{02}^2}{(\gamma_2 + 1)^2}. \tag{38}$$

After rearranging the inequality in (38), we get

$$z_{02}^2\left(\frac{(\gamma_2 - k)^2\sigma^2}{\gamma_2(\gamma_2 + k)^2} + \frac{4k^2\alpha_2^2}{(\gamma_2 + k)^2} - \frac{(\gamma_2 + d)^2\sigma^2}{\gamma_2(\gamma_2 + 1)^2} - \frac{(1-d)^2\alpha_2^2}{(\gamma_2 + 1)^2}\right) < z_{01}^2\sigma^2\left(\frac{(\gamma_1 + d)^2}{\gamma_1(\gamma_1 + 1)^2} - \frac{(\gamma_1 - k)^2}{\gamma_1(\gamma_1 + k)^2}\right). \tag{39}$$

If both

$$\left(\frac{(\gamma_2 - k)^2\sigma^2}{\gamma_2(\gamma_2 + k)^2} + \frac{4k^2\alpha_2^2}{(\gamma_2 + k)^2} - \frac{(\gamma_2 + d)^2\sigma^2}{\gamma_2(\gamma_2 + 1)^2} - \frac{(1-d)^2\alpha_2^2}{(\gamma_2 + 1)^2}\right) \tag{40}$$

and

$$\sigma^2\left(\frac{(\gamma_1+d)^2}{\gamma_1(\gamma_1+1)^2}-\frac{(\gamma_1-k)^2}{\gamma_1(\gamma_1+k)^2}\right) \tag{41}$$

have the same signs; so the NRT estimator gives better results than the Liu estimator when $z_{02}^2 / z_{01}^2 < f_3(\alpha_2^2)$ holds where

$$f_3(\alpha_2^2)=\frac{\sigma^2\left(\frac{(\gamma_1+d)^2}{\gamma_1(\gamma_1+1)^2}-\frac{(\gamma_1-k)^2}{\gamma_1(\gamma_1+k)^2}\right)}{\left(\frac{(\gamma_2-k)^2\sigma^2}{\gamma_2(\gamma_2+k)^2}+\frac{4k^2\alpha_2^2}{(\gamma_2+k)^2}-\frac{(\gamma_2+d)^2\sigma^2}{\gamma_2(\gamma_2+1)^2}-\frac{(1-d)^2\alpha_2^2}{(\gamma_2+1)^2}\right)}. \tag{42}$$

Also, if (40) and (41) have opposite signs, the NRT estimator always gives better results than the Liu estimator where $f_3(\alpha_2^2)$ is negative and $z_{02}^2 / z_{01}^2 > f_3(\alpha_2^2)$ always holds. Consequently, at that region the NRT estimator is better than the Liu estimator.

The positiveness condition of (40) is given by

$$\alpha_2^2 > \frac{\sigma^2\left((\gamma_2+d)^2(\gamma_2+k)^2-(\gamma_2+1)^2(\gamma_2-k)^2\right)}{4k^2\gamma_2(\gamma_2+1)^2-(1-d)^2\gamma_2(\gamma_2+k)^2} \tag{43}$$

and the positiveness condition of (41) is given by

$$(\gamma_1+d)^2(\gamma_1+k)^2 > (\gamma_1+1)^2(\gamma_1-k)^2. \tag{44}$$

Of course, the opposite conditions are needed for the negativeness of (40) and (41). The hyperbola $f_3(\alpha_2^2)$ vertical asymptote is at the point

$$\alpha_2^2 = \frac{\sigma^2\left((\gamma_2+d)^2(\gamma_2+k)^2-(\gamma_2+1)^2(\gamma_2-k)^2\right)}{4k^2\gamma_2(\gamma_2+1)^2-(1-d)^2\gamma_2(\gamma_2+k)^2}. \tag{45}$$

The biasing parameters (k, d) estimation is significant for the multiple regression model suffers from the multicollinearity problem. So, we have not here made any attempt to estimate them. However, we refer the readers to some of these studies, for example [1, 2, 3, 17, 18, 19].

Several biased estimators are developed in different regression models for solving the multicollinearity, such as [20, 21, 22, 23, 24, 25, 26].

4 Application

In this section, we explain the theoretical results of this study using the example given by [4] (i.e., $\sigma^2=1$, $k=0.1$, $\lambda_1=1.95$ and $\lambda_2=0.05$) and [5] (i.e., $d=0.9$).

Firstly, considering the NRT and the OLS estimators' predictive performances. From (21), we get

$$f_1(\alpha_2^2)=\frac{0.05354}{\alpha_2^2-10}, \tag{46}$$

which is a hyperbola with the vertical asymptote at

$$\alpha_2^2 = 10. \tag{47}$$

We are here interested only in the points lie in the first quadrant because of both z_{02}^2 / z_{01}^2 and α_2^2 are positive.

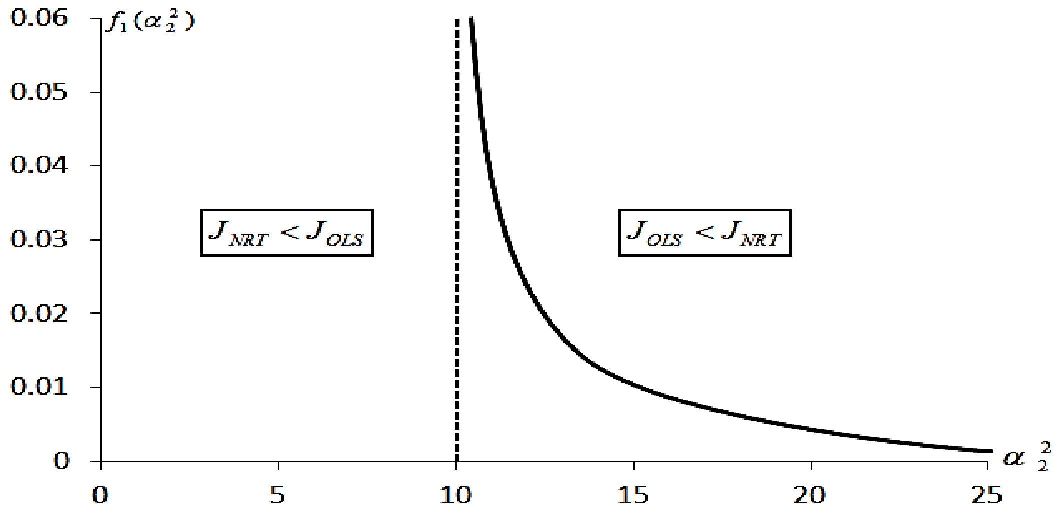

Fig. 1: $f_1(\alpha_2^2)$ versus α_2^2 values

Figure 1 shows that when α_2^2 values are smaller than 10, the NRT estimator gives better results than the OLS estimator and when α_2^2 values are greater than 10, there is a trade-off between the NRT and the OLS estimators such that if the ratio value z_{02}^2 / z_{01}^2 is smaller than the $f_1(\alpha_2^2)$ value, then the NRT estimator gives better results than the OLS estimator, otherwise the OLS estimator is better.

Secondly, considering the NRT and the ORR estimators' predictive performances. From (29), we get

$$f_2(\alpha_2^2) = \frac{0.03478}{\alpha_2^2}, \tag{48}$$

which is a hyperbola with a vertical asymptote at

$$\alpha_2^2 = 0. \tag{49}$$

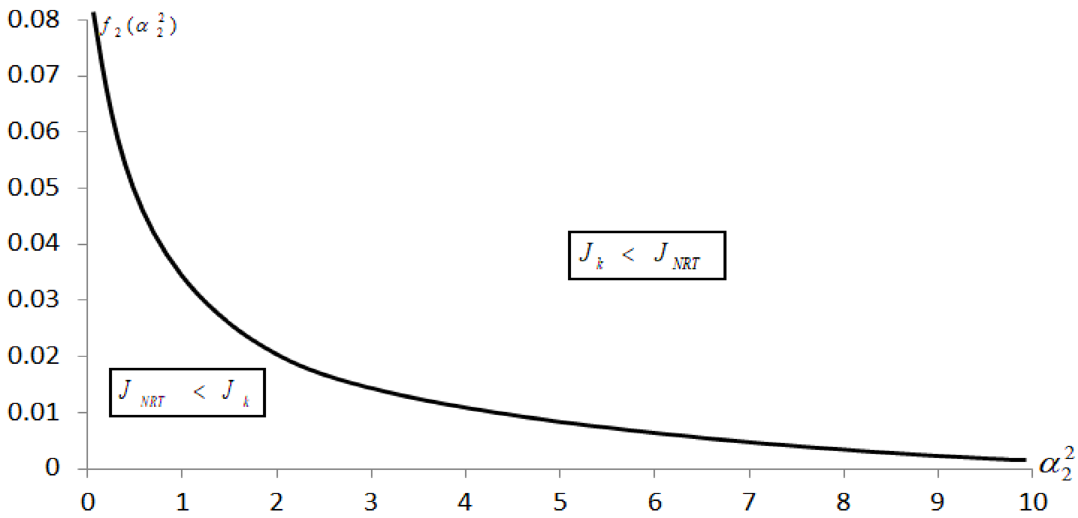

Fig. 2: $f_2(\alpha_2^2)$ versus α_2^2 values

Figure 2 shows that when α_2^2 values are greater than zero, there is a trade-off between the NRT and the ORR estimators such that if the ratio value z_{02}^2 / z_{01}^2 is smaller than the $f_2(\alpha_2^2)$ value, then the NRT estimator gives better results than the ORR estimator, otherwise the ORR estimator is better.

Finally, considering the NRT and Liu estimators predictive performances. From (37), we get

$$f_3(\alpha_2^2) = \frac{0.03449}{\alpha_2^2 - 8}, \quad (50)$$

which is a hyperbola with a vertical asymptote at

$$\alpha_2^2 = 8. \quad (51)$$

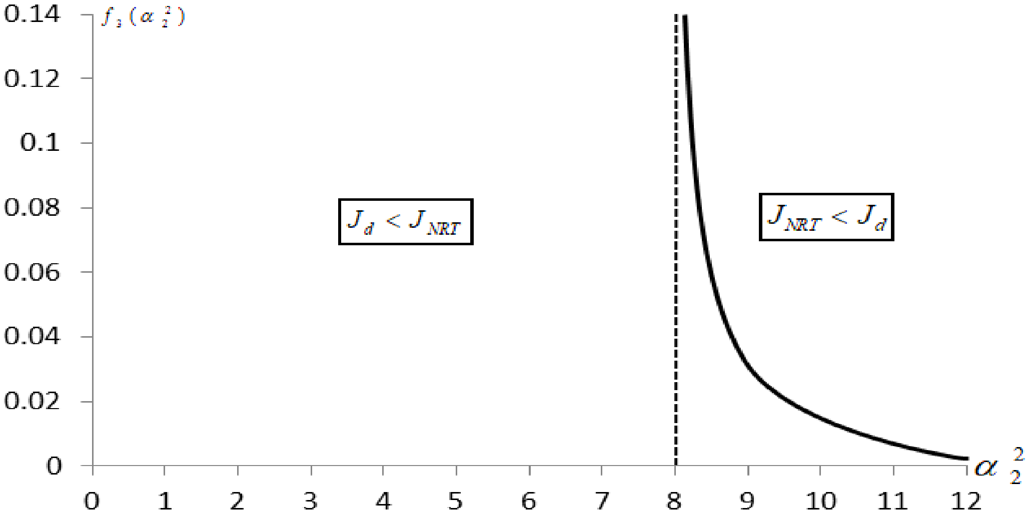

Fig. 3: $f_3(\alpha_2^2)$ versus α_2^2 values

Figure 1 shows that when α_2^2 values are smaller than 8, the NRT estimator gives better results than the Liu estimator and when α_2^2 values are greater than 8, there is a trade-off between the NRT and the Liu estimators such that if the ratio value z_{02}^2 / z_{01}^2 is smaller than the $f_3(\alpha_2^2)$ value, then the NRT estimator gives better results than the Liu estimator, otherwise the Liu estimator is better.

5 Conclusion

We consider and examine the predictive performance of the recently proposed NRT estimator and is compared with the OLS, the ORR and the Liu estimators according to the PMSE criterion at a specific point in the two-dimensional space. The PMSE of the NRT estimator is obtained and three theorems are given. The theoretical results are explained by a numerical example and the regions are assigned where the NRT estimator gives better results than the other mentioned estimators. For some α_2^2 values, there are trade-offs among the above mentioned estimators. The OLS estimator is good only when the value of α_2^2 is very large compared to the NRT estimator. These techniques effectiveness is also affected by the prediction point location. In the numerical example, a region is established where the NRT estimator gives better results than the other mentioned estimators. So, it is theoretically possible to determine such a region.

References:

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

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

[3] B.M. Kibria, A.F. Lukman, A New Ridge-Type Estimator for the Linear Regression Model: Simulations and Applications. *Scientifica* Article ID 9758378, 1-16 (2020).

[4] D.J. Friedman, D.C. Montgomery, Evaluation of the predictive performance of biased regression estimators. *Journal of Forecasting* 4: 153-163 (1985).

[5] F. Özbey, S. Kaçıranlar, Evaluation of the Predictive Performance of the Liu Estimator. *Communications in Statistics Theory Methods* 44: 1981- 1993 (2015).

[6] I. Dawoud, S. Kaciranlar, The Predictive Performance Evaluation of Biased Regression

Predictors With Correlated Errors. *Journal of Forecasting* 34: 364–378 (2015).

[7] I. Dawoud, S. Kaciranlar, The Prediction of the Two Parameter Ridge Estimator. *Istatistik: Journal of the Turkish Statistical Association* 9: 56–66 (2016).

[8] I. Dawoud, S. Kaciranlar, Evaluation of the predictive performance of the Liu type estimator. *Communications in Statistics - Simulation and Computation* 46: 2800-2820 (2017).

[9] I. Dawoud, S. Kaciranlar, Evaluation of the predictive performance of the r-k and r-d class estimators. *Communications in Statistics-Theory and Methods* 46: 4031-4050 (2017).

[10] R. Li, F. Li, J. Huang, Evaluation of the predictive performance of the principal component two-parameter estimator. *Concurrency and Computation: Practice and Experience* e4710 (2018).

[11] I. Dawoud, S. Kaciranlar, The Prediction Performance of the Alternative Biased Estimators for the Distributed Lag Models. *Iranian Journal of Science and Technology, Transactions A: Science* 44: 85–98 (2020).

[12] A.F. Lukman, Z. Y. Algamal, B.G. Kibria, K. Ayinde, The KL estimator for the inverse Gaussian regression model. *Concurrency and Computation: Practice and Experience*, *33*(13), e6222 (2021).

[13] A.F. Lukman, I. Dawoud, B.M. Kibria, Z.Y., Algamal, B. Aladeitan, A new ridge-type estimator for the gamma regression model. *Scientifica*, (2021).

[14] M.N. Akram, B.G. Kibria, M.R. Abonazel, N. Afzal, On the performance of some biased estimators in the gamma regression model: simulation and applications. Journal of Statistical Computation and Simulation, 1-23. DOI: 10.1080/00949655.2022.2032059.

[15] M.R. Abonazel, I. Dawoud, F.A. Awwad, A.F. Lukman, Dawoud–Kibria estimator for beta regression model: simulation and application. *Frontiers in Applied Mathematics and Statistics*, 8:775068 (2022).

[16] I. Dawoud, M.R. Abonazel, Generalized Kibria-Lukman Estimator: Method, Simulation, and Application. *Frontiers in Applied Mathematics and Statistics*, 8:880086 (2022).

[17] B.M. Kibria, Performance of some new ridge regression estimators. *Communications in Statistics Simulation and Computation* 32: 419-435 (2003).

[18] G. Khalaf, G. Shukur, Choosing ridge parameters for regression problems. *Communications in Statistics Theory Methods* 34: 1177-1182 (2005).

[19] G. Muniz, B.M. Kibria, On some ridge regression estimators: An empirical comparison. *Communications in Statistics Simulation and Computation* 38: 621-630 (2009).

[20] M.R. Abonazel, I. Dawoud, Developing robust ridge estimators for Poisson regression model. *Concurrency and Computation: Practice and Experience*, 34: e6979 (2022). https://doi.org/10.1002/cpe.6979.

[21] I. Dawoud, M.R. Abonazel, Robust Dawoud–Kibria estimator for handling multicollinearity and outliers in the linear regression model. *J Stat Comput Simul*. 91:3678–92 (2021).

[22] M.N. Akram, M.R. Abonazel, M. Amin, B.M. Kibria, N. Afzal, A new Stein estimator for the zero-inflated negative binomial regression model. Concurrency Computat Pract Exper. 34:e7045 (2022).

[23] M.R. Abonazel, Z.Y. Algamal, F.A. Awwad, I.M. Taha, A New Two-Parameter Estimator for Beta Regression Model: Method, Simulation, and Application. *Front. Appl. Math. Stat. 7: 780322* (2022).

[24] I. Dawoud, M.R. Abonazel, F.A. Awwad, E. Tag Eldin, A New Tobit Ridge-Type Estimator of the Censored Regression Model With Multicollinearity Problem. *Front. Appl. Math. Stat. 8:952142* (2022).

[25] F.A. Awwad, K.A. Odeniyi, I. Dawoud, Z.Y. Algamal, M.R. Abonazel, B.M. Kibria, E. Tag Eldin, New Two-Parameter Estimators for the Logistic Regression Model with Multicollinearity. WSEAS TRANSACTIONS on MATHEMATICS 21:403-414.

[26] Z.Y. Algamal, M.R. Abonazel, Developing a Liu- type estimator in beta regression model. *Concurrency and Computation: Practice and Experience*, *34*(5):e6685 (2022).

Creative Commons Attribution License 4.0 (Attribution 4.0 International, CC BY 4.0)
This article is published under the terms of the Creative Commons Attribution License 4.0
https://creativecommons.org/licenses/by/4.0/deed.en_US