# Modifying Two-Parameter Ridge Liu Estimator Based on Ridge Estimation

Tarek Omara
Department of statistics, Mathematics and Insurance,
Faculty of commerce, Kafrelsheikh University, Egypt
E-mail: Tarek_em@yahoo.com

## Abstract

In this paper, we introduce the new biased estimator to deal with the problem of multicollinearity. This estimator is considered a modification of Two-Parameter Ridge-Liu estimator based on ridge estimation. Furthermore, the superiority of the new estimator than Ridge, Liu and Two-Parameter Ridge-Liu estimator were discussed. We used the mean squared error matrix (MSEM) criterion to verify the superiority of the new estimate. In addition to, we illustrated the performance of the new estimator at several factors through the simulation study.

**Keywords**: Multicollinearity, Ridge estimator (RE), Liu Estimator (LE), Two-Parameter Ridge Liu Estimator (TLE), Modify Two-Parameter Ridge-Liu Estimator (TRLE), Mean Squared Error Matrix (MSEM).

## Introduction

The problem of multicollinearity is one of the problems that have preoccupied the statisticians for a long time. Many studies have been interested in how to overcome this problem in linear regression models, and was based primarily on ability to overcome the ill condition that appears in mean squared error method. In literature, a set of biased estimators has been proposed to overcome this problem. Horal and Kennard (1970) suggested a ridge estimator which depends on a small constant value known as ridge parameter which adding to the diagonal values of the matrix (X′X) to overcome the ill condition. In the same context, Liu (1993) introduce Liu estimator which it is a combination ridge and stein estimator which proposed by Stein (1956). Actually, the value of the ridge parameter may not be large enough to overcome the multicollinearity problem. Therefore, Liu (2003) suggested Liu-type estimator which has two parameters, so that the increase in one parameter can be limited by the other. There is a series of studies had been directed at improving Liu and Liu-type estimators. Yalian and Yang (2012) modified Liu estimator with prior information for the vector of parameters. Ozkale and Kaciranlar (2007) introduced two-parameter Ridge-Liu estimator that is superior to the Liu-type estimator through the mean square error matrix criteria. Sadullah and Selahattin, (2008), Yang H. and Chang, X. (2010) suggested a new biased estimator that makes Liu estimator based on ridge estimation. Jibo, (2014) proposed unbiased two parameter estimator based on prior information.

In this paper, we introduce a new biased estimator that make two parameter Ridge-Liu estimator based on ridge estimation and we show that the new biased estimator is superiority to ridge, Liu and two-parameter Ridge-Liu estimator and we use the simulation study to explain the theoretical results.

## 1. Background:

Consider the linear regression model

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

Where $Y$ represents an $n \times 1$ observation of response vector, $X$ represents an known $n \times p$ design matrix of rank p, $\beta$ represents an $p \times 1$ vector of unknown parameters and ε is $n \times 1$ of random error with $E(\varepsilon) = 0_{n\times 1}$ vector and $E(\varepsilon\varepsilon') = \Sigma = \sigma^2 I_n$ is n×n variance covariance matrix for errors. The ordinary least squares estimator (OLS) of model (1) is given by

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

This estimator is the best unbiased estimator. However, the existence of the problem of multicollinearity makes this estimator have large least squares error. To overcome the multicollinearity problem, Hoerl and Kennard (1970) introduced the ridge estimator (RE) that has a lower mean squares error than the (OLS) estimator and it is given by

$$\hat{\beta}_{RE}(\mathrm{k}) = (X'X + k\mathrm{I})^{-1}X'Y \qquad (3)$$

Where $k \geq 0$ is ridge biasing parameter.

Liu (1993) introduced the biased estimator which is known as Liu estimator (LE) and that has been obtained by combining the stein estimator which is introduced by Stein (1956) and the RE and it is defined by

$$\hat{\beta}_{LE}(d) = (X'X + \mathrm{I})^{-1}\left(X'Y + d\hat{\beta}_{OLS}\right) \qquad (4)$$

Where $0 < d < 1$ is Liu biasing parameter.

This estimator get by augmenting the equation $\mathrm{d}\hat{\beta}_{OLS} = \beta + \varepsilon$ to the model in (1) and then using the ordinary least squares method.

Liu (2003) introduced Liu-type estimator that improve the Liu estimator, since it has two parameters, by augmenting the equation $\left(-\mathrm{d}/\mathrm{k}^{1/2}\right)\hat{\beta}_{OLS} = \beta + \varepsilon$ to the model in (1) and then using the ordinary least squares method and is given by

$$\hat{\beta}_{LE}(k, \mathrm{d}) = (X'X + k\mathrm{I})^{-1}\left(X'Y + d\hat{\beta}_{OLS}\right) \qquad (5)$$

Sadullah and Selahattin, (2008) suggested a new biased estimator

$$\hat{\beta}_{LRE}(k, \mathrm{d}) = (X'X + \mathrm{I})^{-1}\left(X'Y + d\hat{\beta}_{RE}(\mathrm{k})\right) \qquad (6)$$

by augmenting the equation $\left(-\mathrm{d}/k^{1/2}\right)\hat{\beta}_{RE}(k) = \mathrm{k}^{1/2}\beta + \varepsilon$ to the model in (1) and then using the ordinary least squares method . This estimator has superior to ridge and Liu-estimator. Yang H. and Chang, X.(2010) proposed another form of the new Liu biased estimator which defined as

$$\hat{\beta}_{TE}(k, d) = (X'X + \mathrm{I})^{-1}(X'X + dI)(X'X + k\mathrm{I})^{-1}X'Y \qquad (7)$$

Ozkale and Kaciranlar (2007) introduced two parameter ridge-Liu estimator. This estimator is augmenting the equation $\left(\mathrm{d}\mathrm{k}^{1/2}\right)\hat{\beta}_{\mathrm{OLS}} = \mathrm{k}\beta + \varepsilon$ to the model in (1) and then using the ordinary least squares method and is given by

$$\hat{\beta}_{TLE}(\mathrm{k}, \mathrm{d}) = (X'X + k\mathrm{I})^{-1}\left(X'Y + kd\hat{\beta}_{OLS}\right) \qquad (8)$$

## 2. The new biased Two-Parameter Ridge-Liu Estimator:

We can improve the two parameter ridge-Liu estimator in (8) by augmenting the equation $\left(\mathrm{d}k^{1/2}\right)\hat{\beta}_{RE}(k) = k\beta + \varepsilon$ to the model in (1) and then using the ordinary least squares method, we can get the new estimator as

$$\hat{\beta}_{TRLE}(\mathrm{k}, \mathrm{d}) = (X'X + k\mathrm{I})^{-1}\left(X'Y + kd\hat{\beta}_{RE}(k)\right) \qquad (9)$$

Where $0 < d < 1$ , $\mathrm{k} \geq 0$ and $\hat{\beta}_{RE}(k) = (X'X + k\mathrm{I})^{-1}X'Y$.

The new biased estimator has more advantage than two-parameter Ridge-Liu estimator and at the same time includes the features in OLS estimator, ridge estimator and Liu estimator. We can illustrate the special cases of the new estimator as following:

$$\hat{\beta}_{TRLE}(k,0) = \hat{\beta}_{RE}(k)$$
$$\hat{\beta}_{TRLE}(1,\mathrm{d}) = \hat{\beta}_{LE}(d)$$
$$\hat{\beta}_{TRLE}(0,1) = \hat{\beta}_{OLS}$$

Let $Q$ *and* $\Lambda$ are eigenvector and eigenvalues of $X'X$ and $Z'Z = Q'X'XQ = \Lambda = diag(\lambda_1, \lambda_2, \ldots, \lambda_P)$ where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_P > 0$ then we can rewrite the model in (1) in canonical form

$$Y = Z\alpha + \varepsilon \qquad (10)$$

Where: $Z = XQ, \quad \alpha = Q'\beta$. Accordingly, the estimators are re-represented as

$$\hat{\alpha}_{OLS} = \Lambda^{-1}Z'Y \qquad (11)$$

$$\hat{\alpha}_{RE}(k) = (\Lambda + k\mathrm{I})^{-1}Z'Y = H_1Y \qquad (12)$$

$$\begin{aligned}\hat{\alpha}_{LRE}(\mathrm{k},\mathrm{d}) &= (\Lambda + \mathrm{I})^{-1}\left(Z'Y + d\hat{\alpha}_{RE}(\mathrm{k})\right) \qquad (13)\\ &= (\Lambda + \mathrm{I})^{-1}(\Lambda + (d + k)\mathrm{I})\hat{\alpha}_{RE}(\mathrm{k})\\ &= (\Lambda + \mathrm{I})^{-1}(\Lambda + (d + k)\mathrm{I})(\Lambda + k\mathrm{I})^{-1}Z'Y\\ &= H_2Y\end{aligned}$$

$$\begin{aligned}\hat{\alpha}_{TLE}(\mathrm{k},\mathrm{d}) &= (\Lambda + k\mathrm{I})^{-1}(Z'Y + kd\hat{\alpha}_{OLS}) \qquad (14)\\ &= (\Lambda + k\mathrm{I})^{-1}(\Lambda + kd)\hat{\alpha}_{OLS}\\ &= (\Lambda + k\mathrm{I})^{-1}(\Lambda + kd)\Lambda^{-1}Z'Y = H_3Y\end{aligned}$$

$$\begin{aligned}\hat{\alpha}_{TRLE}(\mathrm{k},\mathrm{d}) &= (\Lambda + k\mathrm{I})^{-1}\left(Z'Y + kd\hat{\alpha}_{RE}(\mathrm{k})\right) \qquad (15)\\ &= (\Lambda + k\mathrm{I})^{-1}\Lambda\,\hat{\alpha}_{OLS} + \mathrm{kd}(\Lambda + k\mathrm{I})^{-1}\hat{\alpha}_{RE}(\mathrm{k})\\ &= [(\Lambda + kI)^{-1} + kd(\Lambda + kI)^{-2}]Z'Y\\ &= [D_k^{-1} + kdD_k^{-2}]Z'Y \quad = H_4Y\end{aligned}$$

Where: $D_k^{-1} = (\Lambda + kI)^{-1}$

## 3. Superiority for the new biased Two-Parameter Ridge-Liu Estimator:

In this section, we use the mean squared error matrix (MSEM) criteria to illustrate the superiority of the new bias estimators to other estimators.

$$\mathrm{MSEM}(\hat{\beta}) = \mathrm{Var}(\hat{\beta}) + \mathrm{Bias}(\hat{\beta})\mathrm{Bias}(\hat{\beta})'$$

Where $\mathrm{Var}(\hat{\beta}) = E\left[\left(\hat{\beta} - E(\hat{\beta})\right)\left(\hat{\beta} - E(\hat{\beta})\right)'\right]$ is denote the dispersion matrix and $\mathrm{Bias}(\hat{\beta}) = E(\hat{\beta}) - \beta$ is bias vector. In fact, for any two estimator $\hat{\beta}_1, \hat{\beta}_2$, iff $\mathrm{MSEM}(\hat{\beta}_1) - \mathrm{MSEM}(\hat{\beta}_2) \geq 0$, we can say that the $\hat{\beta}_1$ is superior to $\hat{\beta}_2$ in MSEM criteria.

As follows, we illustrate the superiority of the new bias estimator to the $\hat{\alpha}_{RE}(\mathrm{k})$, $\hat{\alpha}_{LRE}(\mathrm{k},\mathrm{d})$ and $\hat{\beta}_{TLE}(\mathrm{k},\mathrm{d})$ estimators. The following lemma can be help.

**Lemma (1) :** (Trenkler,1980) Let $\hat{\beta}_j = A_j y, j = 1,2$ be two homogenous linear estimators of $\beta$, such that $D = \mathrm{Var}(\hat{\beta}_1) - \mathrm{Var}(\hat{\beta}_2)$ is p.d If $\mathrm{Bias}(\hat{\beta}_2)'D^{-1}\mathrm{Bias}(\hat{\beta}_2) < \sigma^2$, then $\mathrm{MSEM}(\hat{\beta}_1) - \mathrm{MSEM}(\hat{\beta}_2)$ is P.d.

**Lemma (2) :**( Rao et al, 2008) Let N>0, M>0, then N>M, iff $\lambda_{\max}(MN^{-1}) < 1$.

1. We illustrate the superiority of the new bias estimators $\hat{\alpha}_{TRLE}(k, \mathrm{d})$ to $\hat{\alpha}_{RE}(k)$

The expected value, the bias and $MSEM$ for $\hat{\alpha}_{TLE}(k, \mathrm{d})$ estimator and $\hat{\alpha}_{RE}(k)$ estimators are given at following:

$$E[\hat{\alpha}_{TRLE}(k, \mathrm{d})] = D_k^{-1}\Lambda\, \mathrm{E}[\hat{\alpha}_{OLS}] + k\mathrm{d}D_k^{-1}E[\hat{\alpha}_{RE}(k)]$$
$$= (\Lambda + (k + k\mathrm{d})\mathrm{I})\Lambda D_k^{-2}\alpha$$

$$Bias[\hat{\alpha}_{TRLE}(\mathrm{k}, \mathrm{d})] = (\Lambda + (k + k\mathrm{d})\mathrm{I})\Lambda D_k^{-2}\alpha - \alpha$$

$$= \Big((\Lambda^2 + 2k\Lambda + k^2 I) - k^2 I - D_k^2 + k\mathrm{d}\Lambda - k\Lambda\Big) D_k^{-2}\alpha$$

$$= -(k^2 I - (\mathrm{d}+1)k\Lambda) D_k^{-2}\alpha$$

$$\mathrm{Var}\big(\hat{\alpha}_{TRLE}(\mathrm{k}, \mathrm{d})\big) = \sigma^2[D_k^{-1} + kdD_k^{-2}]\Lambda[D_k^{-1} + kdD_k^{-2}]$$
$$= \sigma^2 D_k^{-1}(I - kdD_k^{-1})(I + kdD_k^{-1})^2 \qquad (16)$$

$$\mathrm{MSEM}\big(\hat{\alpha}_{TRLE}(\mathrm{k}, \mathrm{d})\big) = \sigma^2 D_k^{-1}(I - kdD_k^{-1})(I + kdD_k^{-1})^2 +$$
$$(k^2 I - (\mathrm{d}+1)k\Lambda)D_k^{-2}\alpha\alpha' D_k^{-2}(k^2 I - (\mathrm{d}+1)k\Lambda)'$$
$$= \sigma^2 D_k^{-1}(I - kdD_k^{-1})(I + kdD_k^{-1})^2 +$$
$$(k^2 I - (\mathrm{d}+1)k\Lambda)D_k^{-2}\alpha\alpha' D_k^{-2}(k^2 I - (\mathrm{d}+1)k\Lambda)' = H \qquad (17)$$

And $SMSEM\big(\hat{\alpha}_{TRLE}(\mathrm{k}, \mathrm{d})\big) = tr\Big(MSEM\big(\hat{\alpha}_{TRLE}(k, d)\big)\Big) = \sum_{i=1}^{p} H_i$

$$= \hat{\sigma}^2 \sum_{i=1}^{p}\left[\frac{\lambda_i[1 + dk(\lambda_i + k)^{-1}]^2}{(\lambda_i + k)^2} + \frac{\alpha_i{}^2(k^2 - dk\lambda_i - k\lambda_i)^2}{(\lambda_i + k)^4}\right]$$

$$= \hat{\sigma}^2 \sum_{i=1}^{p}\left[\frac{\lambda_i[(\lambda_i + k) + dk]^2}{(\lambda_i + k)^4} + \frac{\alpha_i{}^2(k^2 - dk\lambda_i - k\lambda_i)^2}{(\lambda_i + k)^4}\right]$$

$$= \hat{\sigma}^2 \sum_{i=1}^{p}\left[\frac{\lambda_i[\lambda_i + k(d+1)]^2}{(\lambda_i + k)^4} + \frac{\alpha_i{}^2(k^2 - dk\lambda_i - k\lambda_i)^2}{(\lambda_i + k)^4}\right]$$

$$= \hat{\sigma}^2 \sum_{i=1}^{p}\left[\frac{\lambda_i[\lambda_i + k(d+1)]^2 + \alpha_i{}^2(k^2 - dk\lambda_i - k\lambda_i)^2}{(\lambda_i + k)^4}\right]$$

Where $\alpha_i$ is *i* th element of $Q'\beta, i = 1,2,$ ...p.

We can write (17) as the following form

$$\mathrm{MSEM}\big(\hat{\alpha}_{TRLE}(\mathrm{k}, \mathrm{d})\big) = \sigma^2 D_k^{-1}(I - kdD_k^{-1})(I + kdD_k^{-1})^2 +$$
$$(k^2 I - (\mathrm{d}+1)k\Lambda)D_k^{-2}\alpha\alpha' D_k^{-2}(k^2 I - (\mathrm{d}+1)k\Lambda)' \qquad (18)$$

$$E[\hat{\alpha}_{RE}(k)] = D_k^{-1}\Lambda\alpha$$

$$Bias[\hat{\alpha}_{RE}(k)] = -kD_k^{-1}\alpha$$

$$\mathrm{Var}\big(\hat{\alpha}_{RE}(k)\big) = \sigma^2 D_k^{-1}\Lambda D_k^{-1\prime}$$

$$= \sigma^2 D_k^{-1}(\Lambda + kI - kI)D_k^{-1\prime}$$
$$= \sigma^2 (I - kD_k^{-1})D_k^{-1} \quad (19)$$

Then:

$$\text{MSEM}\big(\hat{\alpha}_{RE}(k)\big) = \sigma^2 D_k^{-1}\Lambda D_k^{-1} + k^2 D_k^{-1}\alpha\alpha' D_k^{-1\prime} \quad (20)$$

Using (16) and (19), we get the difference as

$$D_1 = Var\big(\hat{\alpha}_{RE}(k)\big) - Var\big(\hat{\alpha}_{TRLE}(k,\text{d})\big)$$
$$= \sigma^2[(I - kD_k^{-1})D_k^{-1} - D_k^{-1}(I - kdD_k^{-1})(I + kdD_k^{-1})^2]$$
$$= \sigma^2[(D_k^{-1} - kD_k^{-2}) - (D_k^{-1} - kdD_k^{-2})(I + kdD_k^{-1})^2]$$
$$= \sigma^2(D_k^{-1} - kD_k^{-2})[I - (I + kdD_k^{-1})^2]$$
$$= \sigma^2(D_k^{-1} - kD_k^{-2})[2kdD_k^{-1} + k^2d^2D_k^{-2}]$$
$$= \sigma^2 D_k^{-1}\Lambda D_k^{-1\prime}[2kdD_k^{-1} + k^2d^2D_k^{-2}] \quad (21)$$

Since $[2kdD_k^{-1} + k^2d^2D_k^{-2}] > 0$ and $D_1 > 0$ then for lemma 1 , $\text{MSEM}\big(\hat{\alpha}_{RE}(k)\big) - \text{MSEM}\big(\hat{\alpha}_{TRLE}(\text{k},\text{d})\big)$ is P.d if

$$\alpha' D_k^{-2}(k^2I - (\text{d}+1)k\Lambda)'\Big[D_k^{-1}\Lambda D_k^{-1\prime}[2kdD_k^{-1} + k^2d^2D_k^{-2}]\Big]^{-1}(k^2I - (\text{d}+1)k\Lambda)D_k^{-2}\alpha < \sigma^2$$

**Theorem (1):** Let $\hat{\beta}_j = A_j y, j = 1,2$ be two homogenous linear estimators of β, such that $D = \text{Var}\big(\hat{\alpha}_{RE}(k)\big) - \text{Var}\big(\hat{\alpha}_{TRLE}(k,\text{d})\big)$ is p.d If $\alpha' D_k^{-2}(k^2I - (\text{d}+1)k\Lambda)'\Big[D_k^{-1}\Lambda D_k^{-1\prime}[2kdD_k^{-1} + k^2d^2D_k^{-2}]\Big]^{-1}(k^2I - (\text{d}+1)k\Lambda)D_k^{-2}\alpha < \sigma^2$ , then $\Delta = \text{MSEM}\big(\hat{\alpha}_{RE}(k)\big) - \text{MSEM}\big(\hat{\alpha}_{TRLE}(k,\text{d})\big)$ is P.d.

We illustrate the superiority of the new bias estimators $\hat{\alpha}_{TRLE}(k,\text{d})$ to $\hat{\alpha}_{LRE}(k,\text{d})$

(Sadullah and Selahattin, (2008)) got the $E[\hat{\alpha}_{LRE}(k,\text{d})]$, $Bias[\hat{\alpha}_{LRE}(k,\text{d})]$, $\text{Var}\big(\hat{\alpha}_{LRE}(k,\text{d})\big)$ and $\text{MSEM}\big(\hat{\alpha}_{LRE}(k,\text{d})\big)$ as the following equation

$$E[\hat{\alpha}_{LRE}(\text{k},\text{d})] = (\Lambda + I)^{-1}\big((\Lambda + dI)(\Lambda + kI)^{-1} + k(\Lambda + kI)^{-1}\big)\Lambda\alpha$$

$$Bias[\hat{\alpha}_{LRE}(\text{k},\text{d})] = \big(((\Lambda + I)^{-1}(\Lambda + dI)(\Lambda + kI)^{-1} + k(\Lambda + I)^{-1}(\Lambda + kI)^{-1})\Lambda - I\big)\alpha$$
$$= (F_3\Lambda - I)\alpha$$

Where $F_3 = (\Lambda + I)^{-1}(\Lambda + dI)(\Lambda + kI)^{-1} + k(\Lambda + I)^{-1}(\Lambda + kI)^{-1}$

$$\text{Var}\big(\hat{\alpha}_{LRE}(\text{k},d)\big)$$
$$= \sigma^2((\Lambda + I)^{-1}(\Lambda + kI)^{-1}(\Lambda + dI) + k(\Lambda + dI)^{-1}(\Lambda + kI)^{-1})\,\Lambda((\Lambda + I)^{-1}(\Lambda + kI)^{-1}(\Lambda + dI) + k(\Lambda + dI)^{-1}(\Lambda + kI)^{-1})$$
$$= (\Lambda + I)^{-1}(I + dD_k^{-1})\,\Lambda(I + dD_k^{-1})\,(\Lambda + I)^{-1}$$
$$= \sigma^2 F_3\,\Lambda F_3{}' \quad (22)$$

Then:

$$\text{MSEM}\big(\hat{\alpha}_{LRE}(k,d)\big) = \sigma^2 F_3\,\Lambda F_3{}' + (F_3\Lambda - I)\alpha\alpha'(F_3\Lambda - I)'$$

Using (16) and (22), we get the difference as

$D_2 = \text{Var}\big(\hat{\alpha}_{LRE}(k,d)\big) - Var\big(\hat{\alpha}_{TRLE}(k,d)\big)$

$= \sigma^2 F_3\,\Lambda F_3{}' - \sigma^2 F_1 \Lambda F_1{}'$

$= \sigma^2\Big[(\Lambda + I)^{-1}(I + dD_k^{-1})\,\Lambda(I + dD_k^{-1})\,(\Lambda + I)^{-1} - [D_k^{-1} + kdD_k^{-2}]\Lambda[D_k^{-1} + kdD_k^{-2}]\Big]$

$= \sigma^2[(\Lambda + \text{I})^{-1}(I + dD_k^{-1})\,(\Lambda + kI - kI)(I + dD_k^{-1})\,(\Lambda + I)^{-1} - D_k^{-1}(I - kdD_k^{-2})(I + kdD_k^{-1})^2]$

$= \sigma^2[(\Lambda + I)^{-1}[(\Lambda + kI + dI)\ - k(I + dD_k^{-1})](I + dD_k^{-1})\,(\Lambda + I)^{-1} - D_k^{-1}(I - kdD_k^{-2})(I + kdD_k^{-1})^2]$

$= \sigma^2[(\Lambda + I)^{-1}[(\Lambda + dI)\ - dkD_k^{-1}](I + dD_k^{-1})\,(\Lambda + I)^{-1} - D_k^{-1}(I - kdD_k^{-2})(I + kdD_k^{-1})^2]$

$= \sigma^2[(\Lambda + I)^{-1}[\Lambda(I + dD_k^{-1}) + d(I - kD_k^{-1})(I + dD_k^{-1})]\,(\Lambda + I)^{-1} - D_k^{-1}(I - kdD_k^{-2})(I + kdD_k^{-1})^2]$

Since $[(\Lambda + I)^{-1}[\Lambda(I + dD_k^{-1}) + d(I - kD_k^{-1})(I + dD_k^{-1})]\,(\Lambda + I)^{-1}] > 0$ and $D_k^{-1}(I - kdD_k^{-2})(I + kdD_k^{-1})^2 > 0$ then by lemma (2) iff $\lambda_{\max}[\ D_k^{-1}(I - kdD_k^{-2})(I + kdD_k^{-1})^2[(\Lambda + \text{I})^{-1}[\Lambda(I + dD_k^{-1}) + d(I - kD_k^{-1})(I + dD_k^{-1})]\,(\Lambda + I)^{-1}]^{-1}] < 1$ then $(\Lambda + I)^{-1}[\Lambda(I + dD_k^{-1}) + d(I - kD_k^{-1})(I + dD_k^{-1})]\,(\Lambda + I)^{-1} - D_k^{-1}(I - kdD_k^{-2})(I + kdD_k^{-1})^2 \geq 0$ ,then by lemma (1) $\Delta_2 = \text{MSEM}\big(\hat{\alpha}_{LRE}(k,d)\big) - MSEM\big(\hat{\alpha}_{TRLE}(k,d)\big)$ is P.d.

**Theorem (2):** Iff $\lambda_{\max}[\ D_k^{-1}(I - kdD_k^{-2})(I + kdD_k^{-1})^2[(\Lambda + \text{I})^{-1}[\Lambda(I + dD_k^{-1}) + d(I - kD_k^{-1})(I + dD_k^{-1})]\,(\Lambda + I)^{-1}]^{-1}] < 1$ , the $\hat{\alpha}_{TRLE}(k,d)$ estimator is superior to the $\hat{\alpha}_{LRE}(k,d)$ estimator.

We illustrate the superiority of the new bias estimators $\hat{\alpha}_{TRLE}(k,d)$ to $\hat{\alpha}_{TLE}(k,d)$

$$E[\hat{\alpha}_{TLE}(k,d)] = (\Lambda + k\text{I})^{-1}(\Lambda + dk\text{I})\alpha$$

$$Bias[\hat{\alpha}_{TLE}(k,d)] = D_k^{-1}[(d-1)k\text{I}]\alpha$$

$$\text{Var}\big(\hat{\alpha}_{TLE}(k,d)\big) = \sigma^2 D_k^{-1}(\Lambda + kd)\Lambda^{-1}(\Lambda + kd)D_k^{-1} \qquad (23)$$

Then:

$$\begin{aligned}\text{MSEM}\big(\hat{\alpha}_{TLE}(\text{k},\text{d})\big) &= \sigma^2 D_k^{-1}(\Lambda + kdI)\Lambda^{-1}(\Lambda + kdI)D_k^{-1} \\ &\quad + D_k^{-1}[(d-1)k\text{I}]\alpha\alpha'[(d-1)k\text{I}]' D_k^{-1}\end{aligned}$$

Using (16) and (23), we get the difference as

$D_3 = \text{Var}\big(\hat{\alpha}_{TLE}(k,d)\big) - Var\big(\hat{\alpha}_{TRLE}(k,d)\big)$

$$= \sigma^2 D_k^{-1}(\Lambda + kdI)\Lambda^{-1}(\Lambda + kdI)D_k^{-1} - [D_k^{-1} + kdD_k^{-2}]\Lambda[D_k^{-1} + kdD_k^{-2}]$$

$$= \sigma^2 D_k^{-1}\Big[(\Lambda + kdI)\Lambda^{-1}(\Lambda + kdI) - [I + kdD_k^{-1}]\Lambda[I + kdD_k^{-1}]\Big]D_k^{-1}$$

$$= \sigma^2 D_k^{-2}\big[(\Lambda + k\mathrm{I})(\mathrm{I} + kd\Lambda^{-1})(\Lambda + kdI)(\Lambda + k\mathrm{I}) - [(\Lambda + k\mathrm{I}) + kdI]\Lambda[(\Lambda + k\mathrm{I}) + kdI]\big]D_k^{-2}$$

$$= \sigma^2 D_k^{-2}\Big[(\Lambda + kI)\big((\Lambda + kdI) + kd\Lambda^{-1}(\Lambda + kdI)\big)(\Lambda + kI) - [(\Lambda + kI) + kdI]\Lambda[(\Lambda + kI) + kdI]\Big]D_k^{-2}$$

$$= \sigma^2 D_k^{-2}\Big[(\Lambda + kI)\big((\Lambda + kdI) + k^2d^2\Lambda^{-1} + kdI\big)(\Lambda + kI) - [(\Lambda + kI) + kdI]\Lambda[(\Lambda + kI) + kdI]\Big]D_k^{-2}$$

$$= \sigma^2 D_k^{-2}\Big[(\Lambda + kI)\big((\Lambda + kdI) + kd(I + kd\Lambda^{-1})\big)(\Lambda + kI) - [(\Lambda + kI) + kdI]\Lambda[(\Lambda + kI) + kdI]\Big]D_k^{-2}$$

$$= \sigma^2 D_k^{-2}\big[(\Lambda + kI)(\Lambda + kdI)(\Lambda + kI) + kd(\Lambda + kI)(I + kd\Lambda^{-1})(\Lambda + kI) - [(\Lambda + kI)\Lambda + kd\Lambda][(\Lambda + kI) + kdI]\big]D_k^{-2}$$

$$= \sigma^2 D_k^{-2}[(\Lambda + kI)(\Lambda + kdI)(\Lambda + kI) + kd(\Lambda + kI)(I + kd\Lambda^{-1})(\Lambda + kI) - (\Lambda + kI)\Lambda(\Lambda + kI) - kd\Lambda(\Lambda + kI) - kd\Lambda(\Lambda + kI) - k^2d^2\Lambda]D_k^{-2}$$

$$= \sigma^2 D_k^{-2}[(\Lambda + kI)(\Lambda + kdI)(\Lambda + kI) + kd(\Lambda + kI)(I + kd\Lambda^{-1})(\Lambda + kI) - (\Lambda + kI)\Lambda(\Lambda + kI) - 2kd\Lambda(\Lambda + kI) - k^2d^2\Lambda]D_k^{-2}$$

$$= \sigma^2 D_k^{-2}[(\Lambda + kI)(\Lambda + kdI)(\Lambda + kI) + k^2d^2(\Lambda + kI)\Lambda^{-1}(\Lambda + \mathrm{k}) + kd(\Lambda + kI)(\Lambda + kI) - (\Lambda + kI)\Lambda(\Lambda + kI) - 2kd(\Lambda + kI)\Lambda - k^2d^2\Lambda]D_k^{-2}$$

$$=\sigma^2 D_k^{-2}[(\Lambda + kI)(\Lambda + kdI)(\Lambda + kI) + k^2d^2(\mathrm{I} + \mathrm{k}\Lambda^{-1})(\Lambda + kI) + kd(\Lambda + kI)(\Lambda + kI) - (\Lambda + kI)\Lambda(\Lambda + kI) - 2kd(\Lambda + kI)\Lambda - k^2d^2\Lambda]D_k^{-2}$$

$$= \sigma^2 D_k^{-2}[(\Lambda + kI)(\Lambda + kdI)(\Lambda + kI) + k^2d^2(\Lambda + kI) + k^3d^2\Lambda^{-1}(\Lambda + kI) + kd(\Lambda + kI)(\Lambda + kI) - (\Lambda + kI)\Lambda(\Lambda + kI) - 2kd(\Lambda + kI)\Lambda - k^2d^2\Lambda]D_k^{-2}$$

$$= \sigma^2 D_k^{-2}[(\Lambda + kI)\Lambda(\Lambda + kI) + kd(\Lambda + kI)(\Lambda + kI) + k^2d^2(\Lambda + kI) + k^3d^2\Lambda^{-1}(\Lambda + kI) + kd(\Lambda + kI)(\Lambda + kI) - (\Lambda + kI)\Lambda(\Lambda + kI) - 2kd(\Lambda + kI)\Lambda - k^2d^2\Lambda]D_k^{-2}$$

$$= \sigma^2 D_k^{-2}[kd(\Lambda + kI)\Lambda + k^2d(\Lambda + kI) + k^2d^2(\Lambda + kI) + k^3d^2\Lambda^{-1}(\Lambda + kI) + kd(\Lambda + kI)\Lambda + k^2d(\Lambda + kI) - 2kd(\Lambda + kI)\Lambda - k^2d^2\Lambda]D_k^{-2}$$

$$= \sigma^2 D_k^{-2}[k^2d\Lambda + k^3dI + k^2d^2\Lambda + k^3d^2I + k^3d^2I + k^4d^2\Lambda^{-1} + k^2d\Lambda + k^3dI - k^2d^2\Lambda]D_k^{-2}$$

$$= \sigma^2 D_k^{-2}[2k^2d\Lambda + 2k^3dI + 2k^3d^2I + k^4d^2\Lambda^{-1}]D_k^{-2}$$

Since $[2k^2d\Lambda + 2k^3dI + 2k^3d^2I + k^4d^2\Lambda^{-1}] > 0$ then $D_3 > 0$ and for lemma 1,
$\mathrm{MSEM}\big(\boldsymbol{\hat{\alpha}_{LRE}}(\mathbf{k},\mathbf{d})\big)$-$\mathrm{MSEM}\big(\hat{\alpha}_{TRLE}(\mathrm{k},\mathrm{d})\big) > 0$

**Theorem (3):** Let $\hat{\beta}_j = A_j y, j = 1,2$ be two homogenous linear estimators of β, such that $D = \mathrm{Var}\big(\boldsymbol{\hat{\alpha}_{LRE}}(\mathbf{k},\mathbf{d})\big) - \mathrm{Var}\big(\hat{\alpha}_{TRLE}(\mathrm{k},\mathrm{d})\big)$ is p.d If

$\text{Bias}(\hat{\alpha}_{TRLE}(\text{k,d}))'\text{D}^{-1}\text{Bias}(\hat{\alpha}_{TRLE}(\text{k,d})) < \sigma^2$, then $\Delta = \text{MSEM}(\hat{\boldsymbol{\alpha}}_{\boldsymbol{LRE}}(\mathbf{k},\mathbf{d}))$-$\text{MSEM}(\hat{\alpha}_{TRLE}(\text{k,d}))$ is P.d.

## 2. Choice for d and k :

For chose the optimal shrinking parameter (d), we differentiating the trace mean squared error matrix $\text{TMSEM}(\hat{\alpha}_{TRLE}(\text{k,d}))$ with respect to d and equating the result to zero and then we can get the optimal estimators for shrinking parameter (d) as the following:

$$\frac{\partial \text{TMSEM}(\hat{\alpha}_{TRLE}(\text{k,d}))}{\partial d} = 2\sigma^2\left[\sum_{i=1}^{p}[k\lambda_i[\lambda_i + k(d+1)]] - \sum_{i=1}^{p} k{\alpha_i}^2\lambda_i(k^2 - dk\lambda_i - k\lambda_i)\right] = 0$$

$$\sum_{i=1}^{p}[k\lambda_i[\lambda_i + k(d+1)]] - \sum_{i=1}^{p}\left[k^3{\alpha_i}^2\lambda_i - k^2 d{\alpha_i}^2{\lambda_i}^2 - k^2{\alpha_i}^2{\lambda_i}^2\right] = 0$$

$$\sum_{i=1}^{p}\left[dk^2\lambda_i + k{\lambda_i}^2 + k^2\lambda_i\right] - \sum_{i=1}^{p}\left[k^3{\alpha_i}^2\lambda_i - dk^2{\alpha_i}^2{\lambda_i}^2 - k^2{\alpha_i}^2{\lambda_i}^2\right] = 0$$

$$d\sum_{i=1}^{p}\left[k^2\lambda_i + k^2{\alpha_i}^2{\lambda_i}^2\right] - \sum_{i=1}^{p}\left[k^3{\alpha_i}^2\lambda_i - k^2{\alpha_i}^2{\lambda_i}^2 - k{\lambda_i}^2 - k^2\lambda_i\right] = 0$$

$$\hat{d}_{opt} = \frac{\sum_{i=1}^{p}\left[k^3{\alpha_i}^2\lambda_i - k^2{\alpha_i}^2{\lambda_i}^2 - k{\lambda_i}^2 - k^2\lambda_i\right]}{\sum_{i=1}^{p}\left[k^2\lambda_i + k^2{\alpha_i}^2{\lambda_i}^2\right]}$$

$$\hat{d}_{opt} = \frac{\sum_{i=1}^{p}\left[k^2{\alpha_i}^2\lambda_i - k{\alpha_i}^2{\lambda_i}^2 - {\lambda_i}^2 - k\lambda_i\right]}{\sum_{i=1}^{p}\left[k[1 + {\alpha_i}^2\lambda_i]\right]} \qquad (24)$$

We chose the k parameter which minimize the Generalized Cross Validation (GCV):

$$GCV(k) = \frac{\sum_{i=1}^{n}(Y_i - Z_i\hat{\alpha})^2}{\left(1 - n^{-1}tr(h_i(k))\right)^2} \qquad (25)$$

Where $tr(h(k))$ is trace for hat matrix $h(k) = Z'(\Lambda + kI)Z$.

## 3. The simulation study:

This section conducts a simulation study to compare the performance of the two-parameter ridge-Liu estimator $(\hat{\alpha}_{\text{TRLE}}(\text{k,d}))$ with other estimators. To generate the explanatory variable with deferent degrees of collinearity, we follow (Liu, 2003) who use the following equation

$$x_{ij} = (1-\gamma^2)^{1/2}\mathcal{Z}_{ij} + \gamma\mathcal{Z}_{ip}, \text{i} = 1,2,\dots,\text{n}, \text{j} = 1,2,\dots,\text{p} - 1$$

Where $\mathcal{Z}_{\text{ij}}$ and $\mathcal{Z}_{\text{ip}}$ are the independent standard normal pseudo-random numbers and it they are generated independently from $\text{N}(0,5)$ and $\gamma$ is specified so $\gamma^2$ is the correlation between any two explanatory variables. We use the three sets of correlations $\gamma = 0.65, 0.80, 0.95$ to show the effect of the week and strong correlation between the explanatory variables. The observations on dependent variable are generate by the following equation

$$y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + e_i \quad , e_i \sim N(0, \sigma^2 I_n) , i = 1,2, \dots, n, j = 1, 2, \dots, p$$

We use sample size n=150,50 and we select $\sigma^2 = 0.01, 0.25$. The value of d and k are calculate by the equations (24), (25). The parameters $\beta_{0i}$ were set to be $(1,2,\dots,5)$ *and* $(1,2,\dots,10)$.We repeated the simulation 2000 times and we use the standard mean squares error MSE to illustrate the superior for the new estimator which is defined by

$$MSE(\hat{\beta}) = \frac{1}{2000} \sum_{i=1}^{2000} (\hat{\beta}_i - \beta)'(\hat{\beta}_i - \beta)$$

Where $\hat{\beta}_i$ is the estimator in ith replication and β is the true parameter values.

The result of the simulation was summarized at table (1- 4). We chose the number of independent variable p , the degree of correlation $\boldsymbol{\gamma}$ , the number of observation n and the variance of the disturbance term $\boldsymbol{\sigma^2}$. The result of the simulation study showed the OLS estimator had a worst for other estimator in all case. They illustrate that the RE, LRE, TLE, TRLE estimators work will at the several degrees of multicollinearity. The new estimator performs well especially when the degrees of multicollinearity is decreases and also it is not affected by the multicollinearity like the other estimator. Moreover, when n increases and at the same time $\sigma^2$ decreases, the MSE value for our new estimator is decreases. It is clear that, increase in the number of observation n and decreases in the number of independent variable p at the several degrees of multicollinearity had a good effect on the work of all estimators especially on the new estimator.

**Table (1) The value for MSE for different estimators p=5,** $\boldsymbol{\sigma^2 = 0.01}$ **,** $\boldsymbol{\hat{d}_{opt}.\hat{k}}$

| Estimators | n=150 | | | n=50 | | |
|---|---|---|---|---|---|---|
| | $\gamma = 0.65$ | $\gamma = 0.80$ | $\gamma = 0.95$ | $\gamma = 0.65$ | $\gamma = 0.80$ | $\gamma = 0.95$ |
| **OLS** | 12.251 | 20.195 | 25.112 | 13.892 | 23.451 | 28.602 |
| **RE** | 9.185 | 9.194 | 9.754 | 11.163 | 10.229 | 13.35 |
| **LRE** | 5.951 | 5.559 | 7.054 | 8.171 | 8.252 | 9.792 |
| **TLE** | 4.024 | 4.081 | 5.011 | 6.419 | 5.544 | 8.884 |
| **TRLE** | 0.253 | 0.893 | 1.085 | 1.981 | 2.025 | 4.128 |

**Table (2)** The value for **MSE** for different estimators $\mathbf{p = 5}, \boldsymbol{\sigma^2 = 0.25}, \boldsymbol{\hat{d}_{opt}.\hat{k}}$

| Estimators | n=150 | | | n=50 | | |
|---|---|---|---|---|---|---|
| | $\gamma = 0.65$ | $\gamma = 0.80$ | $\gamma = 0.95$ | $\gamma = 0.65$ | $\gamma = 0.80$ | $\gamma = 0.95$ |
| **OLS** | 14.082 | 24.051 | 29.932 | 15.081 | 27.191 | 33.219 |
| **RE** | 3.571 | 4.572 | 10.121 | 3.936 | 5.982 | 14.575 |
| **LRE** | 1.795 | 4.042 | 5.255 | 2.025 | 5.215 | 10.682 |
| **TLE** | 4.192 | 4.215 | 5.527 | 6.917 | 6.644 | 9.928 |
| **TRLE** | 1.029 | 1.216 | 1.583 | 2.045 | 2.141 | 6.376 |

**Table (3)** The value for **MSE** for different estimators p=**10**, $\boldsymbol{\sigma^2 = 0.01}$, $\boldsymbol{\hat{d}_{opt}.\hat{k}}$

| Estimators | n=150 | | | n=50 | | |
|---|---|---|---|---|---|---|
| | $\gamma = 0.65$ | $\gamma = 0.80$ | $\gamma = 0.95$ | $\gamma = 0.65$ | $\gamma = 0.80$ | $\gamma = 0.95$ |
| **OLS** | 35.231 | 42.321 | 79.654 | 43.564 | 62.545 | 93.654 |
| **RE** | 15.324 | 19.672 | 26.902 | 19.985 | 25.743 | 33.215 |
| **LRE** | 10.325 | 14.021 | 20.654 | 16.262 | 18.321 | 24.109 |
| **TLE** | 8.658 | 9.945 | 12.358 | 12.252 | 14.068 | 18.325 |
| **TRLE** | 1.927 | 2.059 | 3.325 | 4.024 | 4.921 | 6.179 |

**Table (4)** The value for **MSE** for different estimators $\mathbf{p = 10}, \boldsymbol{\sigma^2 = 0.25}$ , $\boldsymbol{\hat{d}_{opt}.\hat{k}}$

| **Estimators** | **n=150** | | | **n=50** | | |
|---|---|---|---|---|---|---|
| | $\gamma = 0.65$ | $\gamma = 0.80$ | $\gamma = 0.95$ | $\gamma = 0.65$ | $\gamma = 0.80$ | $\gamma = 0.95$ |
| **OLS** | 44.489 | 69.052 | 85.065 | 51.360 | 68.901 | 98.654 |
| **RE** | 19.984 | 26.325 | 30.032 | 35.212 | 43.021 | 50.193 |
| **LRE** | 15.023 | 16.215 | 23.065 | 17.029 | 20.093 | 28.097 |
| **TLE** | 13.335 | 16.685 | 14.594 | 18.094 | 16.009 | 20.095 |
| **TRLE** | 2.304 | 2.905 | 3.531 | 4.906 | 5.302 | 7.932 |

## 4. Conclusions:

In these paper, we introduce the new biased estimator that modifies the two-parameter Ridge-Liu estimator. Moreover, we checked the superiority for the new estimator over the ridge estimator, Liu estimator, and two-parameter Ridge-Liu estimator. The theoretical study was supported by a simulated study which depended on (MSE) criterion to verify the advantage of the new biased estimator.  The result of the simulation study showed the new biased estimator had a superiority over other estimators.

## References

1. Hoerl, A and Kennard.W (1970),” Ridge regression: biased estimation for nonorthogonal Problems”, Technometrics, vol.12: 55-77.
2. Jibo W. (2014), “An unbiased two-parameter estimation with prior information in linear regression model “, The Scientific World Journal, 2014, Article ID 679835.
3. Liu K. (1993),” A new class of biased estimate in linear regression”, Communications in Statistics-Theory and Methods, vol.22:393–402.
4. Liu K. (2003), “Using Liu-type estimator to combat collinearity”, Communications in Statistics-Theory and Methods, vol. 32(5), 1009-1020.
5. Ozkale, M. R. and Kaciranlar S. (2007), " The restricted and unrestricted two-parameter estimators. Communications in Statistics-Theory and Methods, Vol.36:2707-2725.
6. Rao C., Toutenburg H., Shalabh, Heumann C. (2008),” Linear models and generalizations”, Springer,Berlin
7. Sadullah S. and Selahattin K. (2008),” A new biased estimator based on ridge estimation” Stat Papers vol. 49: PP669–689.
8. Stein C. (1956),” Inadmissibility of the usual estimator for mean of multivariate normal distribution”, Neyman J (ed) Proceedings of the third Berkley symposium on mathematical and statistics probability, vol 1, pp 197–206.
9. Swindel B. (1976), “Good ridge estimators based on prior information” , Communications in Statistics-Theory and Methods, vol.5, no. 11, pp. 1065–1075,
10. Trenkler G. (1980),” Generalized mean squared error comparisons of biased regression estimators”, Communication in Statistics Theory and Methods, A9 (Vol.12), 1247-1259.
11. Yalian Li · Yang H. (2012),” A new Liu-type estimator in linear regression model”, Stat Papers, vol. 53:427–437.
12. Yang H. and Chang, X. (2010), “A new two-parameter estimator in linear regression,” Communications in Statistics—Theory and Methods, vol. 39, no. 6, pp. 923–934.