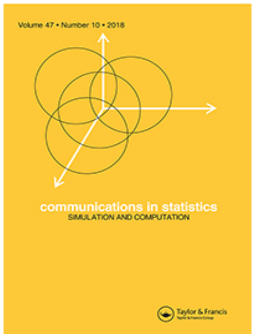

Communications in Statistics - Simulation and Computation

ISSN: 0361-0918 (Print) 1532-4141 (Online) Journal homepage: https://www.tandfonline.com/loi/lssp20

A new version of unbiased ridge regression estimator under the stochastic restricted linear regression model

Mustafa I. Alheety & B. M. Golam Kibria

COMMUNICATIONS IN STATISTICS - SIMULATION AND COMPUTATION®
https://doi.org/10.1080/03610918.2019.1586925

Taylor & Francis
Taylor & Francis Group

A new version of unbiased ridge regression estimator under the stochastic restricted linear regression model

Mustafa I. Alheety[a] and B. M. Golam Kibria[b]

[a]Department of Mathematics, College of Education for Pure Sciences, University of Anbar FIU, Ramadi, Iraq; [b]Department of Mathematics and Statistics, Florida International University, FIU, Miami, FL, USA

ABSTRACT

Crouse et al. (Commun Stat Theory Methods 24:2341–2354, 1995) proposed the unbiased ridge regression estimator for the multicollinear regression model. Jibo Wu (The Scientific World Journal; Volume 2014, Article ID 206943, 1–8) introduced an unbiased two-parameter estimator based on prior information and two-parameter estimator proposed by Özkale and Kacıranlar, 2007. A new version of unbiased two-parameter estimator for the stochastic restricted linear regression model is proposed in this paper. The properties and the performance of the proposed estimator compared to other common estimators using the mean squares error criterion for the goodness of fit have been studied. Finally; A numerical example and a simulation study has been given to illustrate the performance of the proposed estimator.

ARTICLE HISTORY
Received 9 July 2018
Accepted 18 February 2019

KEYWORDS
Mean squared error matrix; Mixed estimator; Ridge regression estimator; Unbiased Two-Parameter Estimation

MATHEMATICS SUBJECT CLASSIFICATION
Primary 62J05; Secondary 62J07

1. Introduction

We consider the following linear regression model,

$$Y = X\beta + \epsilon, \tag{1}$$

where Y is an $n \times 1$ vector of observations that explain the dependent variable, X is an $n \times p$ matrix of observations on p independent variables, β is an $p \times 1$ vector of unknown parameters and ϵ is an $n \times 1$ vector of residuals with zero expectation and variance – covariance matrix, $\sigma^2 I_n$.

When all assumptions of the linear model (1) are satisfied, the least squares estimator which is referred to (OLS) will be the best linear unbiased estimator of β and defined as

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

Unexpectedly, the quality of OLS estimator might be lost, when the multicollinearity problem exists in the model. Therefore, researchers developed different biased estimation techniques which helps to reduce the mean squared error (MSE) of the estimator, such as, the ridge regression estimator (RRE) (Hoerl and Kennard 1970) which is defined as follows:

CONTACT Mustafa I. Alheety alheety@yahoo.com Department of Mathematics, College of Education for Pure Sciences, University of Anbar FIU, Ramadi, Iraq.

$$\hat{\beta}(k) = (X'X + kI)^{-1}X'Y \tag{3}$$

Crouse et al. (1995) presented the unbiased ridge estimator (URE) as a linear combination of prior information with the OLS estimator as follows:

$$\hat{\beta}(k, J) = (X'X + kI)^{-1}(X'Y + kJ), \tag{3}$$

where J is a random vector with $J \sim N(\beta, \sigma^2/k\ I)$ for $k > 0$ and independent of $\hat{\beta}$. Özkale and Kaçıranlar (2007) compared the URE with the OLS estimator and showed that the URE is better than the OLS estimator in the mean squares error matrix (MMSE) sense. Özkale and Kaçıranlar (2007) proposed the following two-parameter estimator as follows:

$$\hat{\beta}(k, d) = (X'X + kI)^{-1}\left(X'Y + kd\ \hat{\beta}\right) = F_{kd}\ \hat{\beta}, \tag{4}$$

where $F_{kd} = (X'X + kI)^{-1}(X'X + kd\ I)$ and $k > 0,\ \ 0 < d < 1$.

Jibo Wu (2014) introduced an unbiased two-parameter estimator based on prior information J and two-parameter estimator UTP, $\hat{\beta}(k, d)$ as follows:

$$\hat{\beta}(F_{kd}, J) = \hat{\beta}(k, d) + (I - F_{kd})J. \tag{5}$$

Suppose we have some prior information about β in the form of stochastic linear restrictions as:

$$h = H\beta + e \sim (0, \sigma^2 V) \tag{6}$$

where h is an $j \times 1$ random vector, H is an $j \times p$ known matrix and V is assumed to be known and positive definite. Also, it is assumed that ϵ is stochastically independent of e.

Theil and Goldberger (1961 and Theil 1963) developed the following ordinary mixed estimator (OME) of β:

$$\hat{\beta}_m = (S + H'V^{-1}H)^{-1}(X'Y + H'V^{-1}h), \tag{7}$$

where $S = X'X$. In case when the prior information and sample information are not equally important, Schaffrin and Toutenburg (1990) introduced the weighted mixed estimator (WME) as follows:

$$\hat{\beta}_w = (S + wH'V^{-1}H)^{-1}(X'Y + wH'V^{-1}h), \tag{8}$$

where $0 \leq w \leq 1$ is a scalar weight.

Li and Yang (2011) proposed a new ridge-type estimator called the weighted mixed ridge estimator (WMRR) by unifying the sample and prior information in linear model with additional stochastic linear restrictions. Their new estimator is a generalization of the weighted mixed estimator and ridge estimator (RE) which is given as follows:

$$\begin{aligned}\hat{\beta}_w(k) &= (S + wH'V^{-1}H)^{-1}\left(T_k X'Y + wH'V^{-1}h\right) \\ &= \hat{\beta}(k) + wS^{-1}H(W + wH'S^{-1}h)\ \left(h - H\hat{\beta}(k)\right)\end{aligned} \tag{9}$$

where $T_k = (I + kS^{-1})^{-1}$.

The goal of this paper is to propose a new version of unbiased ridge regression estimation based on the unbiased two-parameter estimator and stochastic restriction in the model parameters.

This paper is organized as follows: In Sec. 2, the new weighted mixed unbiased ridge regression estimator is introduced. In Sec. 3, a list of some lemmas required for theoretical discussions is presented and the superiority of the proposed estimator compared with some related estimators is given. A numerical example that illustrates some of theoretical results is provided in Sec. 4. Some concluding remarks are given in Sec. 5.

2. The proposed estimator

According to the model in (1), using the prior information J and the stochastic restrictions in (6), we propose the following estimator depending on UTP estimator as follows:

$$\hat{\beta}_w(F_{kd}, J) = \hat{\beta}(F_{kd}, J) + wS^{-1}H(W + wH'S^{-1}h)\ \left(h - H\hat{\beta}(F_{kd}, J)\right) \tag{10}$$

which can be rewritten as follows:

$$\hat{\beta}_w(F_{kd}, J) = (S + wH'W^{-1}H)^{-1}\left(F_{kd}X'Y + wH'^{W^{-1}}h + (I - F_{kd})SJ\right) \tag{11}$$

We call the estimator in (10) as weighted mixed unbiased two- parameter estimator (WMUTP).

From (11), it can be seen that the proposed estimator $\hat{\beta}_w(F_{kd}, J)$ is a general estimator and as special cases of it, the OLS; WME and UTP estimators:

$$\hat{\beta}_w(F_{00}, J) = \hat{\beta}_w;\ \ \hat{\beta}_w(F_{k1}, J) = \hat{\beta}_w,\ \ \forall\ k \in R \text{ and } k \neq 0, \text{ and}$$
$$\hat{\beta}_0(F_{00}, J) = \hat{\beta}_0(F_{01}, J) = \hat{\beta} \text{ and if } \mathrm{H} = 0, \text{ then } \hat{\beta}_w(F_{kd}, J) = \hat{\beta}(F_{kd}, J);\ \text{ where}$$
$$J \sim N\left[\beta, \left(\frac{\sigma^2}{k(1-d)}\right)(S + kdI)S^{-1}\right].$$

Now we will provide the expected value and the covariance matrices of OLS, WME, UPT and WMUTP estimators as follows:

$$E(\hat{\beta}) = \beta;\ \ Cov(\hat{\beta}) = \sigma^2 S^{-1};$$
$$E\left(\hat{\beta}_w\right) = \beta;\ \ Cov\left(\hat{\beta}_w\right) = \sigma^2 A(S + w^2H'W^{-1}H)A;\ \text{ where } A = (S + wH'W^{-1}H)^{-1}.$$
$$E\left(\hat{\beta}(F_{kd}, J)\right) = \beta;\ \ Cov\left(\hat{\beta}(F_{kd}, J)\right) = \sigma^2(S + kI)^{-1}(S + kdI)S^{-1}$$

The properties of the proposed estimator $\hat{\beta}_w(F_{kd}, J)$ is given as follows:

$$E\left(\hat{\beta}_w(F_{kd}, J)\right) = \beta; \tag{12}$$

$$Cov\left(\hat{\beta}_w(F_{kd}, J)\right) = \sigma^2 A\left(F_{kd}S + w^2H'W^{-1}H\right)A. \tag{13}$$

From (12), $\hat{\beta}_w(F_{kd}, J)$ is unbiased estimator.

3. Superiority of the proposed estimator

In this suction we study the performance of the WUMTP estimator in the smaller mean square error matrix MMSE criterion. Firstly, we give the concept of MMSE as follows:

Definition 3.1: Let β^* be any estimator for β, then the MMSE of β^* is given by:

$$MMSE\left(\hat{\beta}^*\right) = \mathrm{Cov}(\hat{\beta}^*) + \mathrm{Bias}(\hat{\beta}^*)\mathrm{Bias}(\hat{\beta}^*)'.$$

Also, the concept of scalar mean squared error is the trace of MMSE; i.e.

$$mse\left(\hat{\beta}^*\right) = \mathrm{tr}\left(MMSE\left(\hat{\beta}^*\right)\right).$$

In order to know the superiority of any estimator $\hat{\beta}^1$ compared to other estimator $\hat{\beta}^2$ under the MMSE criterion, the estimator $\hat{\beta}^2$ is better than $\hat{\beta}^1$ with respect to MMSE sense if and only if:

$MMSE\left(\hat{\beta}^1\right) - MMSE\left(\hat{\beta}^2\right) \geq 0$, i.e. the difference will be positive definite (pd) or positive semi definite (psd) matrix.

3.1. The comparison between the WME and WMUTP estimators

Since the WME and WMUTP estimators are unbiased, the difference of MMSE values between these two estimators can be written as:

$$\begin{aligned}\Delta_1 &= MMSE\left(\hat{\beta}_w\right) - MMSE\left(\hat{B}_w(F_{kd}, J)\right)\\ &= Cov\left(\hat{\beta}_w\right) - Cov\left(\hat{B}_w(F_{kd}, J)\right)\\ &= \sigma^2 A(S + w^2 H' \; W^{-1} H)A - \sigma^2 A\left(F_{kd} S + w^2 H' \; W^{-1} H\right)A\\ &= \sigma^2 A(I - F_{kd})SA\end{aligned}$$

Hence, Δ_1 is pd if $(I - F_{kd}) > 0$. To find the condition that makes $(I - F_{kd}) > 0$, we can use the following canonical form:

Let $S = P\Lambda P'$, $\Lambda = diag\ (\lambda_1, \ldots, \lambda_p)$, where $\lambda_1, \ldots, \lambda_p$ are the eigenvalues of S s.t. $\lambda_1 > \lambda_2 > \ldots > \lambda_p$ and P is the eigenvectors of S. Then,

$$\begin{aligned}I - F_{kd} &= I - (S + kI)^{-1}(S + kdI)\\ &= P\Gamma P' = Pdiag_{(}\gamma_1, \gamma_2, \ldots, \gamma_p)P',\end{aligned}$$

where $\gamma_i = 1 - \frac{\lambda_i + kd}{\lambda_i + k}$. Since $0 < d < 1$ and $k > 0$, it is clear that $\lambda_i + kd < \lambda_i + k$ and consequently $1 - \frac{\lambda_i + kd}{\lambda_i + k} > 0$. Therefore, we can state the following theorem:

Theorem 3.1: For $0 < d < 1$ and $k > 0$, the WMUTP estimator $\hat{\beta}_w(F_{kd}, J)$ is always better than the WME in the sense of MMSE.

3.2. The comparison between the UTP and WMUTP estimators

Since $\hat{\beta}_w(F_{kd}, J)$ *is unbiased estimator as well as* $\hat{\beta}(F_{kd}, J)$, then:

$$\begin{aligned}\Delta_2 &= MMSE\left(\hat{\beta}(F_{kd}, J)\right) - MMSE\left(\hat{\beta}_w(F_{kd}, J)\right)\\ &= Cov\left(\hat{\beta}(F_{kd}, J)\right) - Cov\left(\hat{\beta}_w(F_{kd}, J)\right)\\ &= \sigma^2(S+kI)^{-1}(S+kdI)S^{-1} - \sigma^2 A\left(F_{kd}S + w^2H'W^{-1}H\right)A\\ &= \ \sigma^2 D_1\end{aligned}$$

where $D_1 = F_{kd}S^{-1} - A(F_{kd}S + w^2H'W^{-1}H)A$. The following Lemma is needed to show that under which condition D_1 will be pd.

Lemma 3.1 (See Jibo Wu 2014) Suppose that M is a positive definite matrix and N is a nonnegative definite matrix. Then,

$$M - N \geq 0 \leftrightarrow \ \lambda_{max}(NM^{-1}) \ \leq 1.$$

According to Lemma 3.1; $D_1 \ \geq 0$ if and only if $\lambda_{max}\left(A(F_{kd}S + w^2H'W^{-1}H)A \ \ F_{kd}{}^{-1}S\right) \ \leq 1$. Now, we can state the following theorem:

Theorem 3.2: The WMUTP estimator is superior to the UTP estimator in the MMSE sense if and only if $\lambda_{max}\left(A(F_{kd}S + w^2H'W^{-1}H)A \ \ F_{kd}{}^{-1}S\right) \ \leq 1$.

3.3. The comparison between the OLS and WMUTP estimators

The MMSE difference between OLS and WMUTP estimators is given as follows:

$$\begin{aligned}\Delta_3 &= MMSE(\hat{\beta}) - MMSE\left(\hat{\beta}_w(F_{kd}, J)\right)\\ &= Cov(\hat{\beta}) - Cov\left(\hat{\beta}_w(F_{kd}, J)\right)\\ &= \sigma^2S^{-1} - \sigma^2A\left(F_{kd}S + w^2H'W^{-1}H\right)A\\ &= \ \ \sigma^2S^{-1} + \sigma^2A - \sigma^2A + MMSE\left(\hat{\beta}_w\right) - MMSE\left(\hat{\beta}_w\right)\\ &\quad -\sigma^2A\left(F_{kd}S + w^2H'W^{-1}H\right)A;\\ &= \sigma^2S^{-1} + \sigma^2A - \sigma^2A - \sigma^2A(S + w^2H'W^{-1}H)A + \Delta_1;\end{aligned}$$

where $\Delta_1 = MMSE\left(\hat{\beta}_w\right) - MMSE\left(\hat{\beta}_w(F_{kd}, J)\right)$.

Since,

$\sigma^2S^{-1} - \sigma^2A = \ \sigma^2S^{-1} - \sigma^2(S + wH'W^{-1}H)^{-1}$ and using the fact that,

$$\begin{gathered}(\mathrm{A} + \mathrm{BCD}) - 1 = A^{-1} - A^{-1}\mathrm{B}(\ C^{-1} + \mathrm{DA}^{-1}\mathrm{B}) - 1\mathrm{DA}^{-1}\\ (S + wH'W^{-1}H)^{-1} = S^{-1} - wS^{-1}H' \ \ (W + wHS^{-1}H')^{-1}HS^{-1}\end{gathered}$$

Then, $\sigma^2S^{-1} - \sigma^2A = \ \sigma^2wS^{-1}H' \ \ (W + wHS^{-1}H')^{-1}H \ S^{-1}$, which is positive definite.

Also,

$$\sigma^2 A - \sigma^2 A(S + w^2 H' W^{-1} H)A = \sigma^2 A A^{-1} A - \sigma^2 A(S + w^2 H' W^{-1} H)A$$
$$= \sigma^2 A(A^{-1} - S - w^2 H' W^{-1} H)A$$
$$= \sigma^2 w(1-w)\ A\ H'\ W^{-1} H,$$

which is also pd. and Δ_1 is proved as pd. Therefore, we may state the following theorem:

Theorem 3.3: The WMUTP estimator is always better than the OLS estimator in the sense MMSE criterion.

3.4. The comparison between the WMRR and WMUTP estimators

The MMSE of WMRR estimator is given as follows:

$$MMSE\left(\hat{\beta}_w(k)\right) = \sigma^2 A\left(T_k\ S\ T_k + w^2 H' W^{-1} H\right)A + B_1 B_1';$$

where $B_1 = A(T_k\ - I)S\beta$.

$$\Delta_4 = MMSE\left(\hat{\beta}_w(F_{kd}, J)\right) - MMSE\left(\hat{\beta}_w(k)\right)$$
$$= \sigma^2 A\left(F_{kd} S + w^2 H' W^{-1} H\right)A - \sigma^2 A\left(T_k\ S\ T_k + w^2 H' W^{-1} H\right)A - B_1 B_1'$$
$$= \sigma^2 AMA\ - B_1 B_1',$$

where $M = (F_{kd} S - T_k\ S\ T_k)$

The following Lemma can be helpful

Lemma 3.2 (See F.W. Farebrother 1976) *Let M be a pd matrix and let α be a vector, then $M - \alpha\alpha' \geq 0$ if and only if $\alpha' M^{-1} \alpha \leq 1$.*

Using the canonical form $F_{kd} S - T_k\ S\ T_k$, can be written as

$$F_{kd} S - T_k\ S\ T_k = P\Gamma^* P' = Pdiag(\tau_1, \ldots, \tau_p)P',$$

where $\Gamma^* = (\Lambda + kI)^{-1}(\Lambda + kdI)\Lambda - \left(1 + k\Lambda^{-1}\right)^{-1}\Lambda\left(1 + k\Lambda^{-1}\right)^{-1}$ and

$$\tau_i = (\lambda_i + k)^{-1}(\lambda_i + kd)\lambda_i - \frac{\lambda_i}{\left(1 + k\lambda_i^{-1}\right)^2}.$$

After some mathematical simplifications, we obtain the following

$$\tau_i = \frac{\lambda_i^2\left(k\left(1 + d\left(1 + k\lambda_i^{-1}\right)\right)\right)}{(\lambda_i + k)^2} > 0$$

Therefore $F_{kd} S - T_k\ S\ T_k$ is pd and as well as A and that means that M is pd and by applying Lemma 3.2 we can state the following theorem:

Theorem 3.4: The WMUTP estimator is superior to the WMRR estimator in the MMSE sense; Namely $\Delta_4 < 0$ if and only if $B_1' M^{-1} B_1 > 1$.

Table 1. The estimated mean squares error for OLS, UTP, WME, MWRR and WMUTP estimators.

	w = 0.1			w = 0.1		
	k = 0.2			k = 0.5		
Estimated MMSE values	d = 0.1	d = 0.5	d = 0.9	d = 0.1	d = 0.5	d = 0.9
$\hat{\beta}$	4912.1	4912.1	4912.1	4912.1	4912.1	4912.1
$\hat{\beta}(F_{kd}, J)$	518.03	2470.9	4423.9	502.01	2462	4422.1
$\hat{\beta}_w$	7.2799	7.2799	7.2799	7.2799	7.2799	7.2799
$\hat{\beta}_w(k)$	7.0917	7.0917	7.0917	7.087	7.087	7.087
$\hat{\beta}_w(F_{kd}, J)$	6.5323	6.8646	7.1968	6.5293	6.8629	7.1965

Table 2. The estimated mean squares error for OLS, UTP, WME, MWRR and WMUTP estimators.

	w = 0.1			w = 0.1		
	k = 0.7			k = 0.9		
Estimated MMSE values	d = 0.1	d = 0.5	d = 0.9	d = 0.1	d = 0.5	d = 0.9
$\hat{\beta}$	4912.1	4912.1	4912.1	4912.1	4912.1	4912.1
$\hat{\beta}(F_{kd}, J)$	498.95	2460.3	4421.7	497.24	2459.4	4421.6
$\hat{\beta}_w$	7.2799	7.2799	7.2799	7.2799	7.2799	7.2799
$\hat{\beta}_w(k)$	7.0817	7.0817	7.0817	7.0761	7.0761	7.0761
$\hat{\beta}_w(F_{kd}, J)$	6.5286	6.8625	7.1964	6.5281	6.8622	7.1963

4. A numerical example

In order to illustrate the theoretical results of this paper, we provide a numerical example to examine the performance of the proposed estimator comparing to other estimators in this section. In this regard, the dataset on Portland cement originally due to Woods et al. (1932), has been used and widely analyzed by many researchers, for examples, Kibria (2005) and Li and Yang (2011) among others. Since our theoretical results of this study depend on the unknown parameters β and σ^2, we cannot assume that our results will be held. Thus, we replace them by their corresponding unbiased estimators.

The ordinary least square estimator of β is obtained as:

$$\hat{\beta} = (62.4054,\ 1.5511,\ 0.5102,\ 0.1019, -0.1441).$$

Also the mse of $\hat{\beta}$ is:

$$mse\left(\hat{\beta}\right) = 4912.09, \text{ where } \hat{\sigma}^2 = 5.983$$

According to Li and Yang (2011), we use the following stochastic linear restrictions:

$$\text{h} = \text{HB} + \text{e},\ \ \text{H} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & 1 & 1 \end{bmatrix},\ h = \begin{bmatrix} 63.9498 \\ 2.5648 \end{bmatrix} \text{ and } e \sim \left(0, \hat{\sigma}^2\right).$$

Crouse et al. (1995) suggested $\bar{J} = \left[\sum_{i=1}^{5} \hat{\beta}_i \Big/ 5 \right] 1_{5\times 1}$ as a realistic empirical prior information where 1 is the vector of ones. Using the above information, we computed MMSE for all estimators (their estimated MSE values are obtained by replacing the corresponding theoretical MSE expressions and all unknown model parameters with their respective least squares estimators) and for different values of w, d and k and provided them in Tables 1–6 when $\sigma^2 = \hat{\sigma}^2 = 5.983$.

Table 3. The estimated mean squares error for OLS, UTP, WME, MWRR and WMUTP estimators.

	w = 0.4			w = 0.4		
	k = 0.2			k = 0.5		
Estimated MMSE values	d = 0.1	d = 0.5	d = 0.9	d = 0.1	d = 0.5	d = 0.9
$\hat{\beta}$	4912.1	4912.1	4912.1	4912.1	4912.1	4912.1
$\hat{\beta}(F_{kd}, J)$	518.03	2470.9	4423.9	502.01	2462	4422.1
$\hat{\beta}_w$	6.6242	6.6242	6.6242	6.6242	6.6242	6.6242
$\hat{\beta}_w(k)$	6.6106	6.6106	6.6106	6.6081	6.6081	6.6081
$\hat{\beta}_w(F_{kd}, J)$	6.5761	6.5975	6.6189	6.5756	6.5972	6.6188

Table 4. The estimated mean squares error for OLS, UTP, WME, MWRR and WMUTP estimators.

	w = 0.4			w = 0.4		
	k = 0.7			k = 0.9		
Estimated MMSE values	d = 0.1	d = 0.5	d = 0.9	d = 0.1	d = 0.5	d = 0.9
$\hat{\beta}$	4912.1	4912.1	4912.1	4912.1	4912.1	4912.1
$\hat{\beta}(F_{kd}, J)$	498.95	2460.3	4421.7	497.24	2459.4	4421.6
$\hat{\beta}_w$	6.6242	6.6242	6.6242	6.6242	6.6242	6.6242
$\hat{\beta}_w(k)$	6.6063	6.6063	6.6063	6.6046	6.6046	6.6046
$\hat{\beta}_w(F_{kd}, J)$	6.5754	6.5971	6.6188	6.5752	6.597	6.6188

Table 5. The estimated mean squares error for OLS, UTP, WME, MWRR and WMUTP estimators.

	w = 0.8			w = 0.8		
	k = 0.2			k = 0.5		
Estimated MMSE values	d = 0.1	d = 0.5	d = 0.9	d = 0.1	d = 0.5	d = 0.9
$\hat{\beta}$	4912.1	4912.1	4912.1	4912.1	4912.1	4912.1
$\hat{\beta}(F_{kd}, J)$	518.03	2470.9	4423.9	502.01	2462	4422.1
$\hat{\beta}_w$	6.6054	6.6054	6.6054	6.6054	6.6054	6.6054
$\hat{\beta}_w(k)$	6.6013	6.6013	6.6013	6.5997	6.5997	6.5997
$\hat{\beta}_w(F_{kd}, J)$	6.5931	6.5986	6.6041	6.5928	6.5984	6.604

Table 6. The estimated mean squares error for OLS, UTP, WME, MWRR and WMUTP estimators.

	w = 0.8			w = 0.8		
	k = 0.7			k = 0.9		
Estimated MMSE values	d = 0.1	d = 0.5	d = 0.9	d = 0.1	d = 0.5	d = 0.9
$\hat{\beta}$	4912.1	4912.1	4912.1	4912.1	4912.1	4912.1
$\hat{\beta}(F_{kd}, J)$	498.95	2460.3	4421.7	497.24	2459.4	4421.6
$\hat{\beta}_w$	6.6054	6.6054	6.6054	6.6054	6.6054	6.6054
$\hat{\beta}_w(k)$	6.5986	6.5986	6.5986	6.5976	6.5976	6.5976
$\hat{\beta}_w(F_{kd}, J)$	6.5926	6.5983	6.604	6.5923	6.5982	6.604

Table 7. The estimated mean squares error for OLS, UTP, WME, MWRR and WMUTP, $\sigma^2 = 0.5$.

	w = 0.1			w = 0.1		
	k = 0.2			k = 0.5		
Estimated MMSE values	d = 0.1	d = 0.5	d = 0.9	d = 0.1	d = 0.5	d = 0.9
$\hat{\beta}$	410.507	410.507	410.507	410.507	410.507	410.507
$\hat{\beta}(F_{kd}, J)$	369.5124	287.523	369.5124	369.6229	287.8545	369.6229
$\hat{\beta}_w$	0.552	0.552	0.552	0.552	0.552	0.552
$\hat{\beta}_w(k)$	0.5585	0.5585	0.5585	0.5604	0.5604	0.5604
$\hat{\beta}_w(F_{kd}, J)$	0.5519	0.5517	0.5519	0.5519	0.5517	0.5519

Table 8. The estimated mean squares error for OLS, UTP, WME, MWRR and WMUTP, $\sigma^2 = 0.5$.

	w = 0.1			w = 0.1		
	k = 0.7			k = 0.9		
Estimated MMSE values	d = 0.1	d = 0.5	d = 0.9	d = 0.1	d = 0.5	d = 0.9
$\hat{\beta}$	410.507	410.507	410.507	410.507	410.507	410.507
$\hat{\beta}(F_{kd}, J)$	369.5282	287.5705	369.5282	369.5282	287.5705	369.5282
$\hat{\beta}_w$	0.552	0.552	0.552	0.552	0.552	0.552
$\hat{\beta}_w(k)$	0.5591	0.5591	0.5591	0.5581	0.5581	0.5581
$\hat{\beta}_w(F_{kd}, J)$	0.5519	0.5517	0.5519	0.5518	0.5516	0.5518

Table 9. The estimated mean squares error for OLS, UTP, WME, MWRR and WMUTP, $\sigma^2 = 0.5$.

	w = 0.4			w = 0.4		
	k = 0.2			k = 0.5		
Estimated MMSE values	d = 0.1	d = 0.5	d = 0.9	d = 0.1	d = 0.5	d = 0.9
$\hat{\beta}$	410.507	410.507	410.507	410.507	410.507	410.507
$\hat{\beta}(F_{kd}, J)$	369.5124	287.523	369.5124	369.6229	287.8545	369.6229
$\hat{\beta}_w$	0.5527	0.5527	0.5527	0.5527	0.5527	0.5527
$\hat{\beta}_w(k)$	0.5719	0.5719	0.5719	0.575	0.575	0.575
$\hat{\beta}_w(F_{kd}, J)$	0.5524	0.5518	0.5524	0.5524	0.5518	0.5524

Table 10. The estimated mean squares error for OLS, UTP, WME, MWRR and WMUTP, $\sigma^2 = 0.5$.

	w = 0.4			w = 0.4		
	k = 0.7			k = 0.9		
Estimated MMSE values	d = 0.1	d = 0.5	d = 0.9	d = 0.1	d = 0.5	d = 0.9
$\hat{\beta}$	410.507	410.507	410.507	410.507	410.507	410.507
$\hat{\beta}(F_{kd}, J)$	369.5282	287.5705	369.5282	369.5282	287.5705	369.5282
$\hat{\beta}_w$	0.5527	0.5527	0.5527	0.5527	0.5527	0.5527
$\hat{\beta}_w(k)$	0.5729	0.5729	0.5729	0.5728	0.5728	0.5728
$\hat{\beta}_w(F_{kd}, J)$	0.5524	0.5518	0.5524	0.5523	0.5517	0.5523

Table 11. The estimated mean squares error for OLS, UTP, WME, MWRR and WMUTP, $\sigma^2 = 0.5$.

	w = 0.8			w = 0.8		
	k = 0.2			k = 0.5		
Estimated MMSE values	d = 0.1	d = 0.5	d = 0.9	d = 0.1	d = 0.5	d = 0.9
$\hat{\beta}$	410.507	410.507	410.507	410.507	410.507	410.507
$\hat{\beta}(F_{kd}, J)$	369.5124	287.523	369.5124	369.6229	287.8545	369.6229
$\hat{\beta}_w$	0.5544	0.5544	0.5544	0.5544	0.5544	0.5544
$\hat{\beta}_w(k)$	0.5963	0.5963	0.5963	0.6009	0.6009	0.6009
$\hat{\beta}_w(F_{kd}, J)$	0.5539	0.5527	0.5539	0.5539	0.5527	0.5539

Table 12. The estimated mean squares error for OLS, UTP, WME, MWRR and WMUTP, $\sigma^2 = 0.5$.

	w = 0.8			w = 0.8		
	k = 0.7			k = 0.9		
Estimated MMSE values	d = 0.1	d = 0.5	d = 0.9	d = 0.1	d = 0.5	d = 0.9
$\hat{\beta}$	410.507	410.507	410.507	410.507	410.507	410.507
$\hat{\beta}(F_{kd}, J)$	369.5282	287.5705	369.5282	369.5282	287.5705	369.5282
$\hat{\beta}_w$	0.5544	0.5544	0.5544	0.5544	0.5544	0.5544
$\hat{\beta}_w(k)$	0.5978	0.5978	0.5978	0.5979	0.5979	0.5979
$\hat{\beta}_w(F_{kd}, J)$	0.5539	0.5527	0.5539	0.5535	0.5524	0.5535

From Tables 1 to 6, we can see that for different values of w, k and d, the estimated MSE values of the WMUTP are smaller than those of the OLS, UTP WME and MWRR estimators. However, we also observed that the estimated MSE value of our new estimator is bigger than those of the other estimators for large value of d, specially, when $d = 0.9$.

In order to see the effect of σ^2 on the estimators, we have computed MSE of the proposed estimators for $\sigma^2 = 0.5$ and 1.0. However, to save the space of the journal, we have provided only for $\sigma^2 = 0.5$ in Tables 7–12. From these Tables, we observed that as value of σ^2 increases, the estimated MSE of the estimators also increases. However, our proposed estimator most of the time produces smaller MSE compare to the rest. Therefore, for different values of σ^2, w, k, and d, we can say that the $\hat{\beta}_w(F_{kd}, J)$ is useful to used as an estimator for the regression parameters for the model under consideration.

5. Concluding remarks

In this paper, we introduced a weighted mixed two-parameter unbiased estimator with prior information. We explored the conditions for the superiority of the new estimator over the OLS estimator, the UTP estimator, the WME estimator, and the MWRR estimator in the MMSE sense. Furthermore, A numerical example has been given to illustrate the performance of the proposed estimator comparing to other estimators in this paper.

Acknowledgments

The authors are thankful to the Editor and three referees for their valuable and constructive comments/suggestions, which certainly improved the presentation and quality of the paper.

Dedication

The second author, B. M. Golam Kibria dedicates this paper to all who, sacrificed themselves during the liberation war that started on March 26, 1971 and ended on December 16, 1971 to bring out the freedom of our beautiful Bangladesh.

Disclosure statement

The author declares that there is no conflict of interests regarding the publication of this paper.

Data availability

The [Portland cement] data supporting this study are from previously reported studies and datasets, which have been cited. The processed data are available [Woods, H., Steinour, H. H., Starke, H. R./Indust. Eng. Chem. 24:1207–1241].

References

Crouse, R. H., C. Jin, and R. C. Hanumara. 1995. Unbiased ridge estimation with prior information and ridge trace. *Communication in Statistics- Theory and Methods* 24:2341–54.

Farebrother, R. W. 1976. Further results on the mean square error of ridge regression. *Journal of the Royal Statistical Association C* 38 (3):248–50. doi:10.1111/j.2517-6161.1976.tb01588.x.

Hoerl, A. E., and R. W. Kennard. 1970. Ridge regression: biased estimation for non-orthogonal problems. *Technometrics* 12 (1):55–67.

Kibria, B. M. G. 2005. Applications of some improved estimators in linear regression. *Journal of Modern Applied Statistical Methods* 5 (2):367–80. doi:10.22237/jmasm/1162354200.

Li, Y., and H. Yang. 2011. A new ridge-type estimator in stochastic restricted linear regression. *Statistics* 45 (2):123–30. doi:10.1080/02331880903573153.

Özkale, M. R., and S. Kaçiranlar. 2007. Comparisons of the unbiased ridge estimation to the other estimations. *Communications in Statistics—Theory and Methods* 36 (4):707–23.

Schaffrin, B., and H. Toutenburg. 1990. Weighted mixed regression. *Zeitschrift Fur Angewandte Mathematik Und Mechanik* 70:735–8.

Theil, H. 1963. On the use of incomplete prior information in regression analysis. *Journal of the American Statistical Association* 58 (302):401–14. doi:10.2307/2283275.

Theil, H., and A. S. Goldberger. 1961. On pure and mixed statistical estimation in economics. *International Economic Review* 2 (1):65–78.

Woods, H., H. H. Steinour, and H. R. Starke. 1932. Effect of composition of Portland cement on heat evolved during hardening. Industrial & Engineering Chemistry 24 (11):1207–41. doi:10.1021/ie50275a002.

Wu, J. 2014. An unbiased two-parameter estimation with prior information in linear regression model. *The Scientific World Journal* 2014:1–8. Volume Article ID 206943, doi:10.1155/2014/206943.

Yang, H., and X. Chang. 2010. A new two-parameter estimator in linear regression. *Communication in Statistics- Theory and Methods* 39 (6):923–34. doi:10.1080/03610920902807911.