A new class of efficient and debiased two-step shrinkage estimators: method and application

Muhammad Qasim a, Kristofer Månsson a, Pär Sjölander a and B. M. Golam Kibria b

aDepartment of Economics, Finance and Statistics, Jönköping University, Jönköping, Sweden; bDepartment of Mathematics and Statistics, Florida International University, Miami, FL, USA

ABSTRACT
This paper introduces a new class of efficient and debiased two-step shrinkage estimators for a linear regression model in the presence of multicollinearity. We derive the proposed estimators’ mean square error and define the necessary and sufficient conditions for superiority over the existing estimators. In addition, we develop an algorithm for selecting the shrinkage parameters for the proposed estimators. The comparison of the new estimators versus the traditional ordinary least squares, ridge regression, Liu, and the two-parameter estimators is done by a matrix mean square error criterion. The Monte Carlo simulation results show the superiority of the proposed estimators under certain conditions. In the presence of high but imperfect multicollinearity, the two-step shrinkage estimators’ performance is relatively better. Finally, two real-world chemical data are analyzed to demonstrate the advantages and the empirical relevance of our newly proposed estimators. It is shown that the standard errors and the estimated mean square error decrease substantially for the proposed estimator. Hence, the precision of the estimated parameters is increased, which of course is one of the main objectives of the practitioners.

1. Introduction
To describe the problem, we consider the following classical linear regression model (LRM):

\[ Y = X\beta + \varepsilon, \]  

where \( Y \) is a \((n \times 1)\) vector of observations on the response variable, \( X \) is a \((n \times p)\) full ranked design matrix consisting of the explanatory variables, \( \beta \) is a \((p \times 1)\) column vector of unknown regression coefficients and \( \varepsilon \) is a \((n \times 1)\) vector of random errors assumed to be normally distributed with \( E(\varepsilon) = 0 \) and \( E(\varepsilon\varepsilon^T) = \sigma^2I_n \) where \( I_n \) is a \((n \times n)\) identity matrix. The ordinary least square estimator (OLSE) of the unknown parameter vector \( \beta \) is:

\[ \hat{\beta}_{OLS} = (X^TX)^{-1}X^TY. \]
Multicollinearity causes inflated variance in the model, making the OLSE is unstable, and OLSE becomes sensitive to minor changes in the model. Therefore, Hoerl and Kennard [9] suggested the ridge regression estimator (RRE) to mitigate the multicollinearity problem by reducing the estimator’s variance with the cost of accepting (a small) bias. The RRE is obtained by augmenting Equation (1) with $0 = k^{1/2} \beta + \varepsilon$, and then use the OLSE and derived following form of the estimator:

$$\hat{\beta}_{RR} = (XTX + kIp)^{-1}XTY, (k > 0).$$

Özkale and Kaçiranlar [29] stated that as $k$ becomes larger for the RRE, the distance between $k^{1/2} \beta$ and 0 increases and the RRE have an excessive amount of bias. Therefore, they proposed a two-parameter estimator (TPE) by augmenting Equation (1) with $(kd)\hat{\beta}_{OLS} = k\beta + \varepsilon$, and then using the OLSE. Kibria and Lukman [17] proposed one-parameter estimator by minimizing the objective function $(Y - X\beta)^T(Y - X\beta) + k((\beta + \hat{\beta})^T(\beta + \hat{\beta}) - c)$ with respect to $\beta$.

Estimating the optimal ridge parameter $k$ is a crucial issue for practitioners. See, for instance, the following approaches to choose $k$; McDonald and Galarneau [26], Gibbon [8], Kibria [15], Kibria and Banik [16], Månsson et al. [24], Lukman et al. [20], Lukman and Ayinde [19], Saleh et al. [35], Amin et al. [2], and Naveed et al. (2020) among others. Many other directions have been taken in the literature to improve the classical ridge regression method suggested by Hoerl and Kennard [9]. One such trend is to apply different algorithms such as the kidney-inspired algorithm for finding the optimal shrinkage parameter by Algamal [1], the particle swarm algorithm by Uslu et al. [38] and the boosting algorithm by Tutz and Binder [37]. Yet, another way of improving this method was proposed by Kejian [14]), which is known as the Liu estimator. Its main advantage over the ridge regression estimator is that it is a linear function of the shrinkage parameter.$^2$

Lately, shrinkage estimators have been used for different types of datasets such as Mandal et al. [22] where the Gamma regression model was considered to analyze the prostate cancer data, Maronna [25] suggesting methods for high dimensional data, and Peterson and Kuhn [30] developed methods to deal with noise variables. Different estimators have been proposed in the literature to improve the original ridge estimator by Hoerl and Kennard [9]. Also, the use of ridge regression in different situations has been considered (see, e.g. [4,5,3,6,10,11,34]).

The primary aim of this article is to introduce a new two-step shrinkage estimator (TSSE) that provides an alternative method to mitigate the problem of multicollinearity in the multiple linear regression model. This new method encompasses OLSE and RRE as exceptional cases. In addition, we proposed an almost unbiased version of the TSSE, and this estimator will be called debiased TSSE (DTSSE). We compare the matrix mean square error (MMSE) properties analytically and prove the superiority of our new methods under certain conditions. Then, we show the superiority of the proposed estimator in finite samples using a Monte Carlo simulation study. Finally, we apply the methods on two different chemometric datasets. Regression models are widely used in chemistry to build efficient and robust prediction models. In the first example, we use the classical Portland cement data analyzed by Lukman et al. [21]. This example models the heat evolved after 180 days of curing cement, measured in calories per gram of cement by four highly correlated variables. In the second illustration, we use a dataset from Qasim et al. [32] where the
dependent variable corresponds to the boll weight during the cropping season and experimental material consisted of thirty-two upland cotton accessions. Five highly correlated explanatory variables explain the biochemical traits. In both examples, the benefit of the new estimator is compared to the OLSE and other different shrinkage estimators.

2. Proposed estimators

In the presence of multicollinearity, the estimated regression coefficients using the OLSE are too large in absolute values. Therefore, Hoerl and Kennard [9] and Liu (1993) proposed RRE and Liu regression estimator (LRE), respectively as a remedy when the Euclidean length of the OLSE (||βOLS||) is too large. Both the ordinary RRE and LRE improves the estimation in the presence of multicollinearity since they have a smaller Euclidean length than the OLSE, i.e. ||βRR|| < ||βOLS|| and ||βLR|| < ||βOLS||.

In many real-world chemometrical problems, we expect a situation where the multicollinearity is high but imperfect, and the value of the ridge parameter k becomes too small and the performance of TPE is not satisfactory. On the other hand, for large values of k it decreases the bias problem but the distance between k^{1/2}β and 0 still increase (sometimes substantially). Therefore, we propose another class of TSSE by augmenting Equation (1) with \[ \left( \frac{-kd}{k^{1/2}} \right) \hat{β}_{OLS} = k^{1/2}β + ε, \] for 0 ≤ d < 1, and then apply the OLSE. The advantage of the new estimator over the existing estimators is that augmenting equation still gives a better fit by choosing appropriate values of k and d. In addition, the proposed estimator will give minimum standard errors. We proposed two new estimators, namely \( \hat{β}_{TSS} \) and \( \hat{β}_{BATSS} \) which are given in the following subsections.

2.1. New class of two-step shrinkage estimators, \( \hat{β}_{TSS} \)

Following Hoerl and Kennard [9], Liu [18], and Kaciranlar et al. [12], we proposed the new estimator as follows. Let \( \hat{β}_{TSS} \) be the new estimator of the vector β then we derive it from the following function:

Minimize the squared distance \( \{ \hat{β}_{TSS} - (-dβ_{OLS})\}^T \{ \hat{β}_{TSS} - (-dβ_{OLS})\} \)

subject to \( (Y - X\hat{β}_{TSS})^T(Y - X\hat{β}_{TSS}) = K \),

As a Lagrangian problem this is

\[ \left( \hat{β}_{TSS} + dβ_{OLS} \right)^T \left( \hat{β}_{TSS} + dβ_{OLS} \right) + \left( 1/k \right) \{ (Y - X\hat{β}_{TSS})^T(Y - X\hat{β}_{TSS}) - K \}, \]  (4)

where 1/k is a Lagrangian multiplier and K is a constant term. In addition, the equivalent statement to Equation (4) is that the propose estimator has minimum residual sum of squares in the equivalence class of estimators of the parameter vector β, which are equal distance from dβ_{OLS}. Consequently, the optimization problem in Equation (4) can be rewritten as:

Minimize \( (Y - X\hat{β}_{TSS})^T(Y - X\hat{β}_{TSS}) \)

subject to \( (\hat{β}_{TSS} + dβ_{OLS})^T(\hat{β}_{TSS} + dβ_{OLS}) = K \).

\[ (Y - X\hat{β}_{TSS})^T(Y - X\hat{β}_{TSS}) + k\{ (\hat{β}_{TSS} + dβ_{OLS})^T(\hat{β}_{TSS} + dβ_{OLS}) - K \}. \]  (5)
By differentiating the above function with respect to $\hat{\beta}_{kq}$ and equating to zero, we receive:

$$(X^TX + kI_p)\hat{\beta}_{TSS} = X^TY - dk\beta_{OLS}. \tag{6}$$

The final form of the proposed estimator $\hat{\beta}_{TSS}$ (TSS) is defined as:

$$\hat{\beta}_{TSS} = (I_p - k(1 + d)(X^TX + kI_p)^{-1})\hat{\beta}_{OLS}, \quad k > 0, \quad 0 \leq d < 1, \tag{7}$$

where $0 < k < \infty$ and $d$ are the shrinkage parameters. The proposed TSSE can also be found as a solution to the linear stochastic restriction problem. By considering the prior information for $\beta$ in the form of linear stochastic restriction as follows:

$$\left(-\frac{kd}{k^{1/2}}\right)\hat{\beta}_{OLS} = k^{1/2}\beta + \varepsilon,$$

where $k$ and $d$ are the shrinkage parameters, $\varepsilon$ is a $(p \times 1)$ vector of random errors with mean zero and variance matrix $\sigma^2I_p$. The $\hat{\beta}_{TSS}$ is the TSSE which includes the following estimators as special cases:

$$\lim_{k \to 0} \hat{\beta}_{TSS} = \hat{\beta}_{OLS} = (X^TX)^{-1}X^TY, \text{ the OLS estimator.}$$

$$\lim_{d \to 0} \hat{\beta}_{TSS} = \hat{\beta}_{RR} = (X^TX + kI_p)^{-1}X^TY, \text{ the RR estimator.}$$

### 2.2. A new class of debiased TSSE, $\hat{\beta}_{DTSS}$

Since unbiasedness is a desirable property in real-world applications, we propose a debiased TSSE (DTSSE) with adjustment for the bias of the $\hat{\beta}_{TSS}$. Following Kadiyala [13], we introduce a new version of $\hat{\beta}_{TSS}$ based on the bias of the estimator.

**Definition 2.1:** Consider $\hat{\beta}_{TSS}$ which is a biased estimator of $\beta$ where the bias is given by $\text{Bias}(\hat{\beta}_{TSS}) = E(\hat{\beta}_{TSS}) - \beta = M\beta$, which implies that $E(\hat{\beta}_{TSS} - M\beta) = \beta$. The new estimator $\hat{\beta}_{DTSS} = \hat{\beta}_{TSS} - M\hat{\beta}_{TSS} = (I - M)\hat{\beta}_{TSS}$ is called the debiased two-step shrinkage estimator based on the biased estimator, $\hat{\beta}_{TSS}$.

The new class of DTSSE using Definition 2.1 may be written as:

$$\text{Bias}(\hat{\beta}_{TSS}) = E(\hat{\beta}_{TSS}) - \beta.$$

$$\text{Bias}(\hat{\beta}_{TSS}) = -k(1 + d)(X^TX + kI_p)^{-1}\beta.$$

Using the above expression, it is possible to define the $\hat{\beta}_{DTSS}$ based on the Bias($\hat{\beta}_{TSS}$):

$$\hat{\beta}_{DTSS} = \hat{\beta}_{TSS} - \text{Bias}(\hat{\beta}_{TSS}) = \hat{\beta}_{TSS} + k(1 + d)(X^TX + kI_p)^{-1}\beta.$$

The $\hat{\beta}_{DTSS}$ may now be defined by following the methods in Ohtani [28] where the parameter vector $\beta$ is replaced with $\hat{\beta}_{TSS}$ as follows:

$$\hat{\beta}_{DTSS} = \hat{\beta}_{TSS} + k(1 + d)(X^TX + kI_p)^{-1}\hat{\beta}_{TSS}.$$
which may be simplified to the following equations:

\[ \hat{\beta}_{DTSS} = \{I_p + k(1 + d)(X^T X + kI_p)^{-1}\} \hat{\beta}_{TSS} \]

or

\[ \hat{\beta}_{DTSS} = (2I_p - H_{kd})H_{kd} \hat{\beta}_{OLS}, \quad (8) \]

where \( H_{kd} = \{I_p - k(1 + d)(X^T X + kI_p)^{-1}\} \) and \( H_{kd} \hat{\beta}_{OLS} = \hat{\beta}_{TSS} \).

### 3. Properties of the proposed estimators

In this section, we show the superiority of the proposed estimators over the existing estimators in the sense of matrix mean square error (MMSE). We also illustrate the bias comparison between the TSSE and the DTSSE. The comparisons of different estimators are listed in Table 1. First, rewrite the model in Equation (1) into canonical form as

\[ Y = Z\gamma + \varepsilon, \quad (9) \]

where \( Z = X^T, \gamma = T^T \beta \) and \( T \) is the orthogonal matrix whose columns comprise the eigenvectors of the matrix \( X^T X \). Since \( X^T X \) is symmetric, and there exists a \( p \times p \) orthogonal matrix such that \( X^T X = T^T (X^T X) T = \Lambda = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_p) \), where \( \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p > 0 \) are the eigenvalues of \( X^T X \). The OLSE, RRE, LRE, TPE, TSSE and DTSSE are defined for model (9) as

\[ \hat{\gamma}_{OLS} = (\Lambda)^{-1} Z^T Y. \quad (10) \]
\[ \hat{\gamma}_{RR} = (\Lambda + kI_p)^{-1} Z^T Y. \quad (11) \]
\[ \hat{\gamma}_{LR} = (\Lambda + I_p)^{-1}(\Lambda + dI_p) \hat{\gamma}_{OLS}. \quad (12) \]
\[ \hat{\gamma}_{TP} = (\Lambda + kI_p)^{-1}(\Lambda + kdI_p) \hat{\gamma}_{OLS}. \quad (13) \]
\[ \hat{\gamma}_{TSS} = h_{kd} \hat{\gamma}_{OLS}. \quad (14) \]
\[ \hat{\gamma}_{DTSS} = (2I_p - h_{kd})h_{kd} \hat{\gamma}_{OLS}, \quad (15) \]

where \( h_{kd} = \{I_p - k(1 + d)(\Lambda + kI_p)^{-1}\} \) and \( h_{kd} \hat{\gamma}_{OLS} = \hat{\gamma}_{TSS} \). It is also possible to find the OLSE, RRE, LRE, TPE, TSSE and DTSSE for Equation (1) by multiplying \( T \) (orthogonal matrix) in conjunction with the above equations so that, \( \hat{\beta}_{OLS} = T \hat{\gamma}_{OLS}, \hat{\beta}_{RR} = T \hat{\gamma}_{RR}, \hat{\beta}_{LR} = T \hat{\gamma}_{LR}, \hat{\beta}_{TP} = T \hat{\gamma}_{TP}, \hat{\beta}_{TSS} = T \hat{\gamma}_{TSS} \) and \( \hat{\beta}_{DTSS} = T \hat{\gamma}_{DTSS} \).

### Table 1. Theoretical comparisons names of different estimators.

| Theorem’s name | Estimators | Criteria |
|----------------|------------|----------|
| Theorem 3.1    | DTSS and TSSE | Bias     |
| Theorem 3.2    | OLSE and TSSE | MMSE    |
| Theorem 3.3    | RRE and TSSE | MMSE    |
| Theorem 3.4    | LRE and TSSE | MMSE    |
| Theorem 3.5    | TPE and TSSE | MMSE    |
| Theorem 3.6    | DTSS and TSSE | MMSE    |
3.1. Bias comparison of DTSSE and TSSE

In this section, we compare the bias of the $\hat{\beta}_{TSS}$ and $\hat{\beta}_{DTSS}$. Predominantly, the debiased estimator always provides a smaller bias than the biased estimator, however, it does not give a minimum variance of the regression coefficient. Let $\hat{y}$ be any type of estimator of $\gamma$, the squared bias (SB) of $\hat{y}$ is specified as $SB(\hat{y}) = (Bias(\hat{y}))^2$. Therefore, the bias and the SB of $\hat{y}_{TSS}$ can be defined as:

$$\Bias(\hat{y}_{TSS}) = E(\hat{y}_{TSS}) - \gamma = (h_{kd} - I_p)\gamma,$$

$$SB(\hat{y}_{TSS}) = (h_{kd} - I_p)\gamma^T(h_{kd} - I_p)^T = k^2(1 + d)^2 \sum_{j=1}^{p} \frac{\gamma_j^2}{(\lambda_j + k)^2}. \quad (16)$$

Using the above formulas, the bias and SB of the DTSSE are defined as:

$$\Bias(\hat{y}_{DTSS}) = E(\hat{y}_{DTSS}) - \gamma$$
$$= \{(2I_p - h_{kd})h_{kd} - I\}\gamma$$
$$= -k^2(1 + d)^2(\Lambda + kI_p)^{-2}\gamma$$

$$\Bias(\hat{y}_{DTSS}) = (I_p - h_{kd})^2\gamma$$

$$SB(\hat{y}_{BATSS}) = (I_p - h_{kd})^T\gamma^T((I_p - h_{kd})^2)^T = k^4(1 + d)^4 \sum_{j=1}^{p} \frac{\gamma_j^2}{(\lambda_j + k)^4}. \quad (17)$$

One can compare the SB of the estimators by considering the SB differences between the estimators as $\Theta_1 = SB(\hat{y}_{TSS}) - SB(\hat{y}_{DTSS}) > 0$:

$$\Theta_1 = k^2(1 + d)^2 \sum_{j=1}^{p} \frac{\gamma_j^2}{(\lambda_j + k)^2} - k^4(1 + d)^4 \sum_{j=1}^{p} \frac{\gamma_j^2}{(\lambda_j + k)^4}$$

$$\Theta_1 = \sum_{j=1}^{p} \frac{k^2(1 + d)^2\gamma_j^2((\lambda_j + k)^2 - k^2(1 + d)^2)}{(\lambda_j + k)^4}.$$

where $\lambda_j$ is the eigenvalue of $\Lambda$, $k$ and $d$ are the biasing parameters. Reduction of bias in DTSSE is observed once we consider $|Bias(\hat{y}_{TSS})| - |Bias(\hat{y}_{DTSS})| = \frac{k^2(1 + d)(\lambda_j - kd)}{(\lambda_j + k)^2}|\gamma_j|$. It can be easily seen that $\Theta_1$ is positive if the expression $\min_j (\lambda_j + k)^2 > k^2(1 + d)^2$ when $k > 0$ and $0 \leq d < 1$ and this condition is milder. Therefore, we can define that $SB(\hat{y}_{TSS}) - SB(\hat{y}_{DTSS}) > 0$. Hence, the bias of TSSE is higher than the bias of DTSSE. Therefore, based on the theoretical comparison, we can define the following theorem:

**Theorem 3.1:** Under the multiple linear regression model defined in Equation (9), we have $||Bias(\hat{y}_{DTSS})|| < ||Bias(\hat{y}_{TSS})||$ for $k > 0$ and $0 \leq d < 1$ if $\min_j (\lambda_j + k)^2 > k^2(1 + d)^2$. 
3.2. MMSE comparisons of the estimators

In this section, we compare the performance of the proposed TSSE with the existing estimators in the sense of matrix mean square error (MMSE) criteria. The MMSE of an estimator \( \hat{\alpha} \) of \( \alpha \) can be defined as

\[
\text{MMSE}(\hat{\alpha}) = E(\hat{\alpha} - \alpha)^T(\hat{\alpha} - \alpha) = \text{Cov}(\hat{\alpha}) + \text{Bias}(\hat{\alpha})^T \text{Bias}(\hat{\alpha}),
\]

where \( \text{Cov}(\hat{\alpha}) \) is the covariance matrix of \( \hat{\alpha} \) and \( \text{Bias}(\hat{\alpha}) = E(\hat{\alpha}) - \alpha \) is the bias vector. The scalar MSE of \( \hat{\alpha} \) can be found by employing the trace (\( \text{tr}(\cdot) \)) operator as

\[
\text{MSE}(\hat{\alpha}) = \text{tr}[\text{MSE}(\hat{\alpha})] = \text{tr}[\text{Cov}(\hat{\alpha})] + [\text{Bias}(\hat{\alpha})]^T \text{Bias}(\hat{\alpha})).
\]

Let \( \hat{\alpha}_1 \) and \( \hat{\alpha}_2 \) be the two estimators of \( \alpha \), the estimator \( \hat{\alpha}_2 \) is said to be superior to the estimator \( \hat{\alpha}_1 \) if and only if

\[
\Theta = \text{MSE}(\hat{\alpha}_1) - \text{MSE}(\hat{\alpha}_2) \geq 0.
\]

We define a variety of lemmas to illustrate the MMSE properties of the proposed estimators:

**Lemma 3.1:** Let \( n \times n \) matrices \( \mathbf{M} > 0, \mathbf{N} \geq 0 \), then \( \mathbf{M} > \mathbf{N} \Leftrightarrow \lambda_1(\mathbf{NM}^{-1}) < 1 \), where \( \lambda_1(\mathbf{NM}^{-1}) \) is the largest eigenvalue of the matrix \( \mathbf{NM}^{-1} \).

**Proof:** See Wang et al. [39] for more details.

**Lemma 3.2:** Let \( \hat{\alpha}_j = \mathbf{A}_j \gamma, j = 1, 2 \) be two competing estimators of \( \alpha \). Suppose \( \Theta = \text{Cov}(\hat{\alpha}_1) - \text{Cov}(\hat{\alpha}_2) > 0 \), where \( \text{Cov}(\hat{\alpha}_j), j = 1, 2 \) denotes the covariance matrix of \( \hat{\alpha}_j \). Then \( \Theta(\hat{\alpha}_1, \hat{\alpha}_2) = \text{MMSE}(\hat{\alpha}_1) - \text{MMSE}(\hat{\alpha}_2) \geq 0 \Leftrightarrow b_2^T(\sigma^2\Theta + b_1b_1^T)^{-1}b_2 \leq 1 \), where \( b_j \) denote the bias vector of \( \hat{\alpha}_j, j = 1, 2 \).

**Proof:** See Trenkler and Toutenburg [36] for more details.

**Lemma 3.3:** Let \( \mathbf{M} (\mathbf{M} > 0) \) be a positive definite (p.d.) matrix, \( \alpha \) be a vector of nonzero constants, then \( \mathbf{M} - \alpha\alpha^T \) is a non-negative definite (n.n.d.) matrix if and only if \( \alpha^T \mathbf{M}^{-1} \alpha \leq 1 \).

**Proof:** See Farebrother [7] for more details.

### 3.2.1. MMSE comparison of \( \hat{\gamma}_{TSS} \) and \( \hat{\gamma}_{OLS} \)

From Equations (10) and (14), we can compute the MMSE of \( \hat{\gamma}_{OLS} \) and \( \hat{\gamma}_{TSS} \) by using Equation (18). The MMSE of \( \hat{\gamma}_{OLS} \) can be defined as:

\[
\text{MMSE}(\hat{\gamma}_{OLS}) = \sigma^2(\Lambda)^{-1}.
\]

The MMSE of \( \hat{\gamma}_{TSS} \) can be computed as:

\[
\text{MMSE}(\hat{\gamma}_{TSS}) = \text{Cov}(\hat{\gamma}_{TSS}) + \text{Bias}(\hat{\gamma}_{TSS})^T \text{Bias}(\hat{\gamma}_{TSS}),
\]

\[
\text{MMSE}(\hat{\gamma}_{TSS}) = \sigma^2 h_{kd}(\Lambda)^{-1}(h_{kd})^T + (h_{kd} - I_p)\gamma\gamma^T(h_{kd} - I_p)^T.
\]
From Equations (21) and (22), we find the difference between MMSEs as

$$\Theta_2 = \Theta(\hat{y}_{OLS}, \hat{y}_{TSS}) = \text{MMSE}(\hat{y}_{OLS}) - \text{MMSE}(\hat{y}_{TSS})$$

$$= \sigma^2(\Lambda^{-1} - h_{kd}(\Lambda)^{-1}(h_{kd})^T) - (h_{kd} - I_p)\gamma T (h_{kd} - I_p)^T. \quad (23)$$

In the following theorem, we define the essential and adequate conditions for $\hat{y}_{TSS}$ to be superior to the $\hat{y}_{OLS}$.

**Theorem 3.2:** Let $k > 0$ and $0 \leq d < 1$ under the multiple linear regression model with correlated regressors, the $\hat{y}_{TSS}$ is superior to the $\hat{y}_{OLS}$ in the MMSE sense, namely, $\Theta(\hat{y}_{TSS}, \hat{y}_{OLS})$, if and only if $\gamma^T (h_{kd} - I_p)^T [M_1]^{-1} (h_{kd} - I_p)\gamma \leq \sigma^2$, where $M_1 = \{(\Lambda)^{-1} - h_{kd}(\Lambda)^{-1}(h_{kd})^T\}$.

**Proof:** We rewrite Equation (23) as:

$$\Theta_2 = \sigma^2 \text{diag}\left\{\frac{1}{\lambda_j} - \frac{(\lambda_j - kd)^2}{(\lambda_j + k)^2 \lambda_j}\right\}^p - (h_{kd} - I_p)\gamma T (h_{kd} - I_p)^T.$$

The matrix $M_1 = \{(\Lambda)^{-1} - h_{kd}(\Lambda)^{-1}(h_{kd})^T\}$ is p.d. if and only if

$$(\lambda_j + k)^2 - (\lambda_j - kd)^2 > 0 \text{ or } (\lambda_j + k) - (\lambda_j - kd) > 0$$

For $k > 0$ and $0 \leq d < 1$, it is noted that $(\lambda_j + k)^2 - (\lambda_j - kd)^2 = k(1 + d)[k(1 + d) + 2\lambda_j k] > 0$ or $(\lambda_j + k) - (\lambda_j - kd) = k(1 + d) > 0$. Therefore, $(\Lambda)^{-1} - h_{kd}(\Lambda)^{-1}(h_{kd})^T$ is p.d. The proof is completed by using Lemma 3.3.

**3.2.2. MMSE comparison of $\hat{y}_{TSS}$ and $\hat{y}_{RR}$**

This subsection describes the MMSE comparison between $\hat{y}_{TSS}$ and $\hat{y}_{RR}$, and show the superiority of the $\hat{y}_{TSS}$ to the $\hat{y}_{RR}$. First, we define the MMSE of $\hat{y}_{RR}$ as:

$$\text{Cov}(\hat{y}_{RR}) = \sigma^2(\Lambda + kI_p)^{-1} \Lambda (\Lambda + kI_p)^{-1},$$

$$\text{MMSE}(\hat{y}_{RR}) = \sigma^2 h_{RR}(\Lambda)^{-1}(h_{RR})^T + (h_{RR} - I_p)\gamma T (h_{RR} - I_p)^T. \quad (24)$$

where $h_{RR} = \{I_p + k(\Lambda)^{-1}\}^{-1}$ and $(h_{RR} - I_p) = -k(\Lambda + kI_p)^{-1}$.

From Equations (22) and (24), we get the MMSE difference of $\hat{y}_{TSS}$ and $\hat{y}_{RR}$:

$$\Theta_3 = \Theta(\hat{y}_{RR}, \hat{y}_{TSS}) = \text{MMSE}(\hat{y}_{RR}) - \text{MMSE}(\hat{y}_{TSS})$$

$$= \sigma^2 \{h_{RR}(\Lambda)^{-1}(h_{RR})^T - h_{kd}(\Lambda)^{-1}(h_{kd})^T\}$$

$$+ (h_{RR} - I_p)\gamma T (h_{RR} - I_p)^T - (h_{kd} - I_p)\gamma T (h_{kd} - I_p)^T. \quad (25)$$
In the following theorem, we define the essential and adequate conditions for $\hat{\gamma}_{TSS}$ to be superior to the $\hat{\gamma}_{RR}$.

**Theorem 3.3:** Let $k > 0$ and $0 \leq d < 1$ under the multiple linear regression model with correlated regressors, then $\text{MMSE}(\hat{\gamma}_{RR}) - \text{MMSE}(\hat{\gamma}_{TSS}) > 0$ if $\gamma^T(h_{kd} - I_p)^T[M_2]^{-1}(h_{kd} - I_p)\gamma \leq 1$, where $M_2 = \sigma^2\{h_{RR}(\Lambda)^{-1}(h_{RR})^T - h_{kd}(\Lambda)^{-1}(h_{kd})^T\}$.

**Proof:** We rewrite Equation (25) as:

$$\Theta_3 = \sigma^2 \text{diag} \left\{ \frac{\lambda_j}{(\lambda_j + k)^2} - \frac{(\lambda_j - kd)^2}{(\lambda_j + k)^2\lambda_j} \right\}_{j=1}^p,$$

$$+ (h_{RR} - I_p)\gamma\gamma^T(h_{RR} - I_p)^T - (h_{kd} - I_p)\gamma\gamma^T(h_{kd} - I_p)^T.$$  

Since $(h_{RR} - I_p)\gamma\gamma^T(h_{RR} - I_p)^T$ is n.n.d., therefore, it is enough evidence to prove that $\sigma^2\{h_{RR}(\Lambda)^{-1}(h_{RR})^T - h_{kd}(\Lambda)^{-1}(h_{kd})^T\} = (h_{kd} - I_p)\gamma\gamma^T(h_{kd} - I_p)^T$ is p.d.

It is clear when $k > 0$ and $0 \leq d < 1$, the matrix $M_2$ is p.d. if $(\lambda_j)^2 - (\lambda_j - kd)^2 > 0$. Correspondingly, the last inequality, one finds that $2\lambda_jkd - (kd)^2 > 0 \forall j = 1, 2, \ldots, p$. ■

### 3.2.3. MMSE comparison of $\hat{\gamma}_{TSS}$ and $\hat{\gamma}_{LR}$

First, we define the MMSE of $\hat{\gamma}_{LR}$ as follows:

$$\hat{\gamma}_{LR} = (\Lambda + I_p)^{-1}(\Lambda + dI_p)\hat{\gamma}_{OLS}$$

$$\text{Cov}(\hat{\gamma}_{LR}) = \sigma^2(\Lambda + I_p)^{-1}(\Lambda + dI_p)(\Lambda^{-1})((\Lambda + I_p)^{-1}(\Lambda + dI_p))^T$$

$$\text{MMSE}(\hat{\gamma}_{LR}) = \sigma^2[h_{LR}(\Lambda)^{-1}(h_{LR})^T + (d - 1)^2(\Lambda + I_p)^{-1}\gamma\gamma^T((\Lambda + I_p)^{-1})^T, \quad (26)$$

where $h_{LR} = (\Lambda + I_p)^{-1}(\Lambda + dI_p)$.  

From Equations (22) and (26), we find the difference between $\text{MMSE}(\hat{\gamma}_{TSS})$ and $\text{MMSE}(\hat{\gamma}_{LR})$:

$$\Theta_4 = \Theta(\hat{\gamma}_{LR}, \hat{\gamma}_{TSS}) = \text{MMSE}(\hat{\gamma}_{LR}) - \text{MMSE}(\hat{\gamma}_{TSS})$$

$$\Theta_4 = \sigma^2\{h_{LR}(\Lambda)^{-1}(h_{LR})^T - h_{kd}(\Lambda)^{-1}(h_{kd})^T\}$$

$$+ (d - 1)^2(\Lambda + I_p)^{-1}\gamma\gamma^T((\Lambda + I_p)^{-1})^T - (h_{kd} - I_p)\gamma\gamma^T(h_{kd} - I_p)^T.$$  

(27)

In the following theorem, we explain the essential and adequate conditions for $\hat{\gamma}_{TSS}$ to be superior to the $\hat{\gamma}_{LR}$.

**Theorem 3.4:** Let us consider two biased competing estimators, namely $\hat{\gamma}_{LR}$ and $\hat{\gamma}_{TSS}$ of $\gamma$. If $k > 0$ and $0 \leq d < 1$ under the multiple linear regression model with correlated regressors, the estimator $\hat{\gamma}_{TSS}$ is superior to the estimator $\hat{\gamma}_{LR}$ in the MMSE form, namely $\text{MMSE}(\hat{\gamma}_{LR}) - \text{MMSE}(\hat{\gamma}_{TSS}) > 0$ if $\gamma^T(h_{kd} - I_p)^T[M_3]^{-1}(h_{kd} - I_p)\gamma < 1$, where $M_3 = \sigma^2\{h_{RR}(\Lambda)^{-1}(h_{RR})^T - h_{kd}(\Lambda)^{-1}(h_{kd})^T\}$.  


Proof: We can also define the MMSE difference between $\hat{\gamma}_{\text{LR}}$ and $\hat{\gamma}_{\text{TSS}}$ as below:

$$\Theta_4 = \sigma^2 \text{diag} \left\{ \frac{(\lambda_j + d)^2}{\lambda_j(\lambda_j + 1)^2} - \frac{(\lambda_j - kd)^2}{(\lambda_j + k)^2 \lambda_j} \right\}^p_{j=1}$$

$$+ (d - 1)^2 (\Lambda + I_p)^{-1} \gamma \gamma^T \{((\Lambda + I_p)^{-1})^T - (h_{kd} - I_p) \gamma \gamma^T (h_{kd} - I_p)^T \}.$$  

It is evident that the $\text{Bias}(\hat{\gamma}_{\text{TSS}})\{\text{Bias}(\hat{\gamma}_{\text{TSS}})\}^T = (d - 1)^2 (\Lambda + I_p)^{-1} \gamma \gamma^T \{(\Lambda + I_p)^{-1}\}^T > 0$ for $0 \leq d < 1$, therefore, it is enough evidence to prove that $\sigma^2 \{h_{\text{LR}}(\Lambda)^{-1}(h_{\text{LR}})^T - h_{kd}(\Lambda)^{-1}(h_{kd})^T\} - (h_{kd} - I_p) \gamma \gamma^T (h_{kd} - I_p)^T$ is p.d., if and only if $(\lambda_j + k)(\lambda_j + d)^2 > (\lambda_j + 1)^2(\lambda_j - kd)^2 + (\lambda_j + 1)(\lambda_j - kd) > 0$. Simplify the last inequality, one can find as $\lambda_j(d + k + kd - 1) + kd > 0 \Leftrightarrow k > 0$ and $0 \leq d \leq 1$. Consequently, $\sigma^2 \{h_{\text{LR}}(\Lambda)^{-1}(h_{\text{LR}})^T - h_{kd}(\Lambda)^{-1}(h_{kd})^T\} - (h_{kd} - I_p) \gamma \gamma^T (h_{kd} - I_p)^T$ is p.d., and the proof is completed by using Lemma 3.3.

3.2.4. MMSE comparison of $\hat{\gamma}_{\text{TSS}}$ and $\hat{\gamma}_{\text{TP}}$

From the TPE, we compute the bias and covariance matrix of $\hat{\gamma}_{\text{TP}}$ as below:

$$\text{Bias}(\hat{\gamma}_{\text{TP}}) = E(\hat{\gamma}_{\text{TP}}) - \gamma = k(d - 1)(\Lambda + kI_p)^{-1}\gamma$$

$$\text{Cov}(\hat{\gamma}_{\text{TP}}) = \sigma^2(\Lambda + kI_p)^{-1}(\Lambda + kdI_p)(\Lambda^{-1})\{(\Lambda + kI_p)^{-1}(\Lambda + kdI_p)\}^T$$

$$\text{MMSE}(\hat{\gamma}_{\text{TP}}) = \sigma^2 h_{TP}(\Lambda^{-1})(h_{TP})^T + k^2(d - 1)^2(\Lambda + kI_p)^{-1} \gamma \gamma^T \{(\Lambda + kI_p)^{-1}\}^T,$$

(28)

where $h_{TP} = (\Lambda + kI_p)^{-1}(\Lambda + kdI_p)$.

The difference between the MMSE functions of $\hat{\gamma}_{\text{TP}}$ and $\hat{\gamma}_{\text{TSS}}$ is obtained as

$$\Theta_5 = \Theta(\hat{\gamma}_{\text{TP}}, \hat{\gamma}_{\text{TSS}}) = \text{MMSE}(\hat{\gamma}_{\text{TP}}) - \text{MMSE}(\hat{\gamma}_{\text{TSS}})$$

$$\Theta_5 = \sigma^2 \{h_{TP}(\Lambda^{-1})(h_{TP})^T - h_{kd}(\Lambda)^{-1}(h_{kd})^T\}$$

$$+ k^2(d - 1)^2(\Lambda + kI_p)^{-1} \gamma \gamma^T \{(\Lambda + kI_p)^{-1}\}^T - b_{kd}b_{kd}^T.$$  

(29)

In Theorem 3.5, we explain the essential and adequate conditions for $\hat{\gamma}_{\text{TSS}}$ to be superior to the $\hat{\gamma}_{\text{TP}}$.

Theorem 3.5: Under the multiple linear regression model with correlated regressors, if $(b_{kd})^T [M_4]^{-1} b_{kd} < 1$ for $k > 0$ and $0 < d < 1$, then $\text{MMSE}(\hat{\gamma}_{\text{TP}}) - \text{MMSE}(\hat{\gamma}_{\text{TSS}}) > 0$, where $b_{kd} = (h_{kd} - I_p) \gamma$ and $M_4 = \sigma^2 \{h_{TP}(\Lambda^{-1})(h_{TP})^T - h_{kd}(\Lambda)^{-1}(h_{kd})^T\}$.

Proof: We can write the expression (29) as

$$= \sigma^2 \text{diag} \left\{ \frac{(\lambda_j + kd)^2}{\lambda_j(\lambda_j + k)^2} - \frac{(\lambda_j - kd)^2}{(\lambda_j + k)^2 \lambda_j} \right\}^p_{j=1}$$

$$+ k^2(d - 1)^2(\Lambda + kI_p)^{-1} \gamma \gamma^T \{(\Lambda + kI_p)^{-1}\}^T - b_{kd}b_{kd}^T.$$
\[
\begin{align*}
\sigma^2 \text{diag} \left\{ \frac{4\lambda_j kd}{\lambda_j (\lambda_j + k)^2} \right\}^P_{j=1} \\
+ k^2 (d-1)^2 (\Lambda + kI_p)^{-1} \gamma \gamma^T \{ (\Lambda + kI_p)^{-1} \}^T - b_{kd} b_{kd}^T.
\end{align*}
\]

Since \( b_{kd} b_{kd}^T \) is n.n.d., then it is noticeable that \((h_{TP}(\Lambda^{-1})(h_{TP})^T - h_{kd}(\Lambda^{-1})(h_{kd})^T) + k^2 (d-1)^2 (\Lambda + kI_p)^{-1} \gamma \gamma^T \{ (\Lambda + kI_p)^{-1} \}^T \) will be p.d. It can be easily shown that \( \text{Cov}(\hat{y}_{TP}) \sim \text{Cov}(\hat{y}_{TSS}) \) is a p.d. matrix for \( k > 0 \) and \( 0 \leq d < 1 \). Hence, we can state that \( \hat{y}_{TSS} \) has a smaller sampling variance and covariance matrix than the \( \hat{y}_{TP} \). Thus, the proof is completed through Lemmas 3.1 and 3.3.

**3.2.5. MMSE comparison of \( \hat{y}_{DTSS} \) and \( \hat{y}_{TSS} \)**

First, we compute the MMSE and scaler MSE of the estimators:

\[
\begin{align*}
\text{MMSE}(\hat{y}_{TSS}) &= \sigma^2 h_{kd}(\Lambda)^{-1}(h_{kd})^T + (h_{kd} - I_p) \gamma \gamma^T (h_{kd} - I_p)^T \\
\text{MSE}(\hat{y}_{kd}) &= \sum_{j=1}^{P} \frac{\sigma^2 (1 - \Phi_j)^2 + \lambda_j \Phi_j^2 \gamma_j^2}{\lambda_j}.
\end{align*}
\]

(30)

where \( \Phi_j = kd^*/(\lambda_j + k) \) and \( d^* = d + 1. \)

\[
\begin{align*}
\text{MMSE}(\hat{y}_{DTSS}) &= \sigma^2 (2I_p - h_{kd}) h_{kd}(\Lambda^{-1}) \{ (2I_p - h_{kd}) h_{kd} \}^T \\
&\quad + (I_p - h_{kd})^2 \gamma \gamma^T (I_p - h_{kd})^T. \quad (31)
\end{align*}
\]

\[
\begin{align*}
\text{MSE}(\hat{y}_{DTSS}) &= \sum_{j=1}^{P} \frac{\sigma^2 (1 - \Phi_j^2)^2 + \lambda_j \Phi_j^4 \gamma_j^2}{\lambda_j}.
\end{align*}
\]

(32)

**Theorem 3.6:** The estimator \( \hat{y}_{DTSS} \) is superior to the \( \hat{y}_{TSS} \) in the linear regression model if \( \sigma^2 < [k \gamma_j^2 ((\lambda_j + 2k)/(2\lambda_j + 3k))] \forall j = 1, 2, \ldots, p \), when \( 1 - \Phi_j > 0 \) and, then for a fixed \( k > 0 \), \( \text{MSE}(\hat{y}_{TSS}) - \text{MSE}(\hat{y}_{DTSS}) > 0 \) for \( 0 < \min(d_{ij}^*) < \max(d_{ij}^*) < 2 \), where \( d_{ij}^* = \left\lfloor \frac{\left( \lambda_j + k \right) \left( \sqrt{(\lambda_j \gamma_j^2)^2 + 10\lambda_j (\sigma^2 \gamma_j^2)^2 + (3\sigma^2)^2} - \lambda_j \gamma_j^2 - \sigma^2 \right) }{2k(\gamma_j^2 + \sigma^2)} \right\rfloor \) and \( d^* = d + 1. \)

**Proof:** The difference between Equations (30) and (31) is:

\[
\begin{align*}
\text{MSE}(\hat{y}_{TSS}) - \text{MSE}(\hat{y}_{DTSS}) &= \sum_{j=1}^{P} \frac{\sigma^2 (1 - \Phi_j)^2 + \lambda_j \Phi_j^2 \gamma_j^2}{\lambda_j} - \sum_{j=1}^{P} \frac{\sigma^2 (1 - \Phi_j^2)^2 + \lambda_j \Phi_j^4 \gamma_j^2}{\lambda_j} \\
&\quad = \sum_{j=1}^{P} \frac{\sigma^2 (1 - \Phi_j)^2 + \lambda_j \Phi_j^2 \gamma_j^2 - \sigma^2 (1 - \Phi_j^2)^2 - \lambda_j \Phi_j^4 \gamma_j^2}{\lambda_j} \\
&\quad = \sum_{j=1}^{P} \frac{\sigma^2 (1 - \Phi_j)^2 [1 - (1 + \Phi_j)^2] + \lambda_j \Phi_j^2 \gamma_j^2 - \lambda_j \Phi_j^4 \gamma_j^2}{\lambda_j}
\end{align*}
\]
\[
= \sum_{j=1}^{p} \left[ \frac{(\Phi_j - \Phi_j^2)\gamma_j^2}{\lambda_j} \left\{ (\Phi_j + \Phi_j^2)\lambda_j - \frac{\sigma^2}{\gamma_j^2}(2 - \Phi_j - \Phi_j^2) \right\} \right]
\]

The difference \( \text{MSE}(\hat{\gamma}_{TSS}) - \text{MSE}(\hat{\gamma}_{DTSS}) \) will be n.n.d., when

(i) \((\Phi_j + \Phi_j^2)\lambda_j - \frac{\sigma^2}{\gamma_j^2}(2 - \Phi_j - \Phi_j^2) \geq 0.\)

(ii) \(1 - \Phi_j > 0 \forall j = 1, 2, \ldots, p.\)

The inequality, \((\Phi_j + \Phi_j^2)\lambda_j - \frac{\sigma^2}{\gamma_j^2}(2 - \Phi_j - \Phi_j^2) \geq 0\) is fulfilled for \(0 < d^* < 2\) and it can be noticed that:

\[
d^*_j = \left( \frac{\gamma_j^2}{\sqrt{\lambda_j \gamma_j^2} + (2\lambda_j + 3k)} \right) \left( \frac{\lambda_j + k}{\lambda_j + 2k} \right), \quad (33)
\]

Since the acceptable range of \(d^*\) and using Equation (33), we have \(\sigma^2 < \frac{k\gamma_j^2(\lambda_j + 2k)}{(2\lambda_j + 3k)} \forall j = 1, 2, \ldots, p.\) Then \(\text{MSE}(\hat{\gamma}_{TSS}) - \text{MSE}(\hat{\gamma}_{DTSS}) > 0\) when \(1 - \Phi_j > 0 \forall j = 1, 2, \ldots, p\) and when \(0 < \min(d^*_j) < \max(d^*_j) < 2.\)

4. Selection of shrinkage parameters \(k\) and \(d\)

The performance of the proposed estimator depends on the suitable value of the shrinkage parameters \(k\) and \(d.\) Therefore, we derive optimal \(k\) and \(d,\) and suggest an iterative method for the determination of \(k\) and \(d.\) The optimal value of \(d\) is obtained by taking the derivatives of \(\text{MSE}(\hat{\gamma}_{TSS})\) with respect to \(d\) for fixed \(k\) as follows:

\[
\frac{\partial \text{MSE}(\hat{\gamma}_{TSS})}{\partial d} = \sum_{j=1}^{p} \frac{2\gamma_j^2k^2\lambda_j(1 + d) - 2k\sigma^2(\lambda_j - kd)}{\lambda_j(\lambda_j + k)^2}.
\]

For \(\frac{\partial \text{MSE}(\hat{\gamma}_{TSS})}{\partial d} = 0,\) simplifying the numerator of the above expression and solving for \(d\) as:

\[
d = \frac{\sum_{j=1}^{p} (\sigma^2 - \gamma_j^2k)}{\sum_{j=1}^{p} \left( \frac{k\sigma^2}{\lambda_j} + \gamma_j^2k \right)},
\]

where \(\sigma^2\) and \(\gamma_j^2\) are the unknown parameters and we replace these unknown parameters with their unbiased estimators and propose the following estimator:

\[
\hat{d} = \min \left( 1, \frac{(\hat{\sigma}^2 - \hat{\gamma}_{min}^2k)}{\left( \frac{k\sigma^2}{\lambda_{min}} + \hat{\gamma}_{min}^2k \right)} \right), \quad (34)
\]
The condition \( \lambda_j \hat{\sigma}^2 - \hat{\gamma}_j^2 k \lambda_j > 0 \) should hold for the value of \( \hat{d} \) to be positive and therefore, we propose the following restriction for \( \hat{d} \) as

\[
\hat{k} = \min \left( \frac{\hat{\sigma}^2}{\hat{\gamma}_j^2} \right)_{j=1}^p
\]  

(35)

The optimal value of \( k \) is determined by differentiating \( \text{MSE}(\hat{\gamma}_{TSS}) \) for \( k \) and equating it to be zero;

\[
\frac{\partial \text{MSE}(\hat{\gamma}_{TSS})}{\partial k} = \sum_{j=1}^p \frac{2(1 + d)[(d \sigma^2 + (\gamma_j^2 d + \gamma_j^2) \lambda_j)k - \lambda_j \sigma^2]}{(\lambda_j + k)^3},
\]

\[
k_j = \frac{\lambda_j \sigma^2}{d \sigma^2 + \gamma_j^2 (1 + d) \lambda_j}.
\]  

(36)

When \( d = 0 \), the expression, \( k_j = \lambda_j \sigma^2 / (d \sigma^2 + \gamma_j^2 (1 + d) \lambda_j) \) reduces to \( k_j = \sigma^2 / \gamma_j^2 \), which is suggested by Hoerl and Kennard [9] to estimate the ridge parameter \( k \). It can be noted that the value of \( k_j \) is always positive. The expression in Equation (36) depends on the unknown parameters \( \sigma^2 \) and \( \gamma_j^2 \), and we replaced them with their corresponding unbiased estimators and proposed the following ridge estimator as:

\[
\hat{k}_{opt} = \min \left[ \frac{\lambda_j \hat{\sigma}^2}{d \hat{\sigma}^2 + \hat{\gamma}_j^2 (1 + d) \lambda_j} \right]
\]  

(37)

Following, Månsson et al. [23] and Qasim et al. [31], we propose the following estimators to estimate the value of \( \hat{d} \).

\[
\hat{d}_1 = \min \left[ 1, \left( \sum_{j=1}^p \frac{1}{q_j} \right) / p \right] \], \hat{d}_2 = \min \left[ 1, \left( \min_{j=1}^p \frac{1}{q_j} \right) \right],
\]

\[
\hat{d}_3 = \min \left[ 1, \left( \text{median} \left( \frac{1}{q_j} \right) \right) \right], \hat{d}_4 = \sum_{j=1}^p \left( \frac{\sigma^2 - \gamma_j^2 k^*}{\lambda_{max}^2 + \gamma_j^2 k^*} \right) / p,
\]

where \( q_j = \frac{\hat{\gamma}_j^2 k_{opt}}{\hat{k}_{opt}} \).

Finally, we define the following procedure to determine the value of the biasing parameters, \( k \) and \( d \):

**Step 1:** Compute the value of \( k^* \) using Equation (35).

**Step 2:** Estimate \( \hat{d} \) from Equation (34) by using \( k^* \) in step 1.

**Step 3:** Calculate \( \hat{d}_1 - \hat{d}_4 \) by substituting in the value of \( \hat{k}_{opt} \) and \( k^* \).
Table 2. Affective parameters in the simulation studies.

| Factors                        | Symbol | Values                   |
|--------------------------------|--------|--------------------------|
| Number of explanatory variables| $p$    | 4, 8                     |
| Sample size                    | $n$    | 15, 25, 50, 100, 200     |
| Degree of correlation          | $\rho^2$ | 0.75, 0.90, 0.95, 0.99  |
| Values of the error variance   | $\sigma^2$ | 0.50, 1.00, 2.00         |
| Number of replicates           | $R$    | 5000                     |

5. Monte Carlo simulations

This Monte Carlo simulation is carried out to compare the finite sample properties of the proposed estimators with the traditional estimators in different empirically relevant situations. The average means square error (AMSE) of the estimator is determined based on 5000 replications and the entire process executed 5000 times to compute the simulated AMSE as follows:

$$\text{AMSE}(\hat{\gamma}) = \frac{\sum_{r=1}^{5000} (\hat{\gamma}_r - \gamma)^T (\hat{\gamma}_r - \gamma)}{5000}$$

where $\hat{\gamma}$ is any estimator ($\hat{\gamma}_{\text{OLS}}, \hat{\gamma}_{\text{RRE}}, \hat{\gamma}_{\text{LRE}}, \hat{\gamma}_{\text{TPE}}, \hat{\gamma}_{\text{TSSE}}, \hat{\gamma}_{\text{DTSS}}$) for each replicate and $r$ denotes the number of replication.

5.1. The design of the experiment

The correlated explanatory variables are generated by following Gibbon [8] and Kibria [15] as follows:

$$x_{ij} = (1 - \rho^2)^{1/2} Z_{ij} + \rho Z_{i(j+1)},$$

where $Z_{ij}$ are the independent standard normal pseudo-random numbers and $\rho$ is the degree of correlation between two regressors which is given by $\rho^2$. The performance of the proposed estimators depends on different factors which are demonstrated in Table 2. The response variable is generated as follows:

$$y_i = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_p x_{ip} + \varepsilon_i, i = 1, 2, \ldots, n, j = 1, 2, \ldots, p.$$ 

where $y_i$ represent the $n$th observations of the dependent variable $\beta_j$ are the regression coefficients and $\varepsilon_i$ is the independent identically normally distributed error term with mean zero and variance $\sigma^2$. The values of $\beta$ are chosen such that $\beta^T \beta = 1$ (see, [27]). The AMSE is minimized when $\beta$ is the normalized eigenvector corresponding to the largest eigenvalue of the matrix $X^T X$. Therefore, we selected the regression coefficients $\beta_j(\beta_1, \ldots, \beta_p)$ as the normalized eigenvector corresponding to the largest eigenvalue of the matrix $X^T X$. Besides, we assume zero intercept for the model in Equation (10) without loss of any generality. Then the variables are standardized so that $X^T y$ represents the vector of correlations between the explanatory variables and the dependent variables.

5.2. Simulation results

To demonstrate the finite sample properties of the estimators, the simulation results are summarized in Tables 3–6. We computed the AMSE of the OLSE, RRE, LRE, TPE, TSSE...
and DTSS under different situations that are common in a real-world application by changing the sample size \((n)\), population variance \((\sigma^2)\), degree of correlation \((\rho^2)\) and the number of explanatory variables \((p)\). In almost all cases, the proposed class of TSSE performed well. Especially the performance of the \(\hat{\gamma}_{TSS}(\hat{d}_4)\) estimator is extremely satisfactory in most of the cases, and it is always as good as (and usually superior) to the \(\hat{\gamma}_{OLS}\), \(\hat{\gamma}_{LR}\) and the \(\hat{\gamma}_{TP}\). Though, in a few cases, the performance of the \(\hat{\gamma}_{LR}\) is reasonably good when \(n\) is small, and there is a limited number of explanatory variables. The \(\hat{\gamma}_{LR}\) does not perform well when the \(n\), \(\rho^2\) and \(p\) increase. For example, when \(n = 15\), \(p = 8\), \(\rho^2 = 0.99\), \(\sigma^2 = 0.50\); AMSE \((\hat{\gamma}_{LR}) = 8.190\) and AMSE \((\hat{\gamma}_{TSS}(\hat{d}_4)) = 6.492\). As the parameters \(n\), \(p\), \(\rho^2\) and \(\sigma^2\) increase in size, the relative performance of TSSE is substantially improved.

From Tables 3–6, it can be seen that the simulated AMSE values decrease when \(n\) becomes larger and the AMSE values increase when \(p\), \(\rho^2\) and \(\sigma^2\) increase. The performance of the \(\hat{\gamma}_{TSS}(\hat{d}_4)\) is quite well than other estimators in the sense of AMSE. When looking at the performance of the debiased two-step shrinkage estimator (namely, \(\hat{\gamma}_{DTSS}\)) is not always superior to the \(\hat{\gamma}_{OLS}\) in the sense of AMSE.

In this section, we show the benefit of proposed estimators through two real-life applications.

### 6.1. Portland cement dataset

The Portland dataset, which was initially adopted by Woods et al. (1932) and also used in Lukman et al. [21] is used as a first illustration in this paper to demonstrate the performance of the new estimator. The dependent variable is defined as the heat evolved after 180 days of curing measured in calories per gram of cement. This variable is modeled using four correlated explanatory variables corresponding to \(x_1\) that represents tricalcium aluminate, \(x_2\) that represents tricalcium silicate, \(x_3\) that represents tetracalcium aluminoferrite, and \(x_4\) that represents \(\beta\)-dicalcium silicate. In Table 7, the correlation matrix is shown that indicates a multicollinearity problem since the pairwise correlations reach up to an absolute value of 0.9730. Also, the condition index defined as \(\sqrt{\lambda_{max}/\lambda_{min}}\) equals 6056, indicating...
Table 3. Simulated AMSE values of the estimators when $\rho^2 = 0.75$.

| $n$  | $\sigma^2$ | $\gamma_{OLS}$ | $\gamma_{RR}$ | $\gamma_{LR}$ | $\gamma_{TSS}(d_1)$ | $\gamma_{TSS}(d_2)$ | $\gamma_{TSS}(d_3)$ | $\gamma_{TSS}(d_4)$ | $\gamma_{TSS}(d_4)$ |
|------|------------|----------------|---------------|---------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| 15   | 0.50       | 3.254          | 0.176         | 0.156         | 0.188               | 0.194               | 0.196               | 0.193               | 0.174               |
|      | 1.00       | 3.862          | 0.646         | 0.648         | 0.735               | 0.771               | 0.774               | 0.776               | 0.740               |
|      | 2.00       | 6.201          | 2.229         | 2.496         | 2.701               | 6.808               | 5.054               | 5.682               | 2.607               |
| 25   | 0.50       | 3.182          | 0.148         | 0.133         | 0.155               | 0.163               | 0.162               | 0.164               | 0.141               |
|      | 1.00       | 3.692          | 0.558         | 0.566         | 0.619               | 0.643               | 0.656               | 0.645               | 0.508               |
|      | 2.00       | 5.624          | 1.960         | 2.221         | 2.305               | 7.427               | 3.009               | 6.007               | 2.320               |
| 50   | 0.50       | 2.986          | 0.053         | 0.050         | 0.054               | 0.056               | 0.056               | 0.056               | 0.051               |
|      | 1.00       | 3.148          | 0.194         | 0.202         | 0.206               | 0.215               | 0.218               | 0.216               | 0.182               |
|      | 2.00       | 3.813          | 0.700         | 0.820         | 0.796               | 1.082               | 0.876               | 1.079               | 0.814               |
| 100  | 0.50       | 1.061          | 0.029         | 0.029         | 0.030               | 0.030               | 0.030               | 0.030               | 0.029               |
|      | 1.00       | 1.552          | 0.110         | 0.114         | 0.115               | 0.118               | 0.119               | 0.118               | 0.106               |
|      | 2.00       | 1.507          | 0.398         | 0.458         | 0.437               | 0.485               | 0.466               | 0.482               | 0.456               |
| 200  | 0.50       | 2.989          | 0.014         | 0.014         | 0.014               | 0.014               | 0.014               | 0.014               | 0.014               |
|      | 1.00       | 3.028          | 0.052         | 0.053         | 0.053               | 0.055               | 0.055               | 0.055               | 0.051               |
|      | 2.00       | 3.195          | 0.197         | 0.209         | 0.209               | 0.218               | 0.218               | 0.218               | 0.186               |
| 500  | 0.50       | 1.387          | 0.067         | 0.067         | 0.069               | 0.071               | 0.071               | 0.071               | 0.066               |
|      | 1.00       | 1.596          | 0.251         | 0.274         | 0.273               | 0.281               | 0.281               | 0.282               | 0.279               |
|      | 2.00       | 2.455          | 0.898         | 1.078         | 1.044               | 1.050               | 1.079               | 1.062               | 1.019               |
| 1000 | 0.50       | 1.543          | 0.013         | 0.013         | 0.013               | 0.013               | 0.013               | 0.013               | 0.013               |
|      | 1.00       | 2.745          | 0.123         | 0.127         | 0.127               | 0.128               | 0.128               | 0.128               | 0.122               |
|      | 2.00       | 3.151          | 0.451         | 0.512         | 0.500               | 0.508               | 0.513               | 0.511               | 0.498               |
Table 4. Simulated AMSE values of the estimators when $\rho^2 = 0.90$.

| $n$  | $\sigma^2$ | $\hat{\gamma}_{OLS}$ | $\hat{\gamma}_{RR}$ | $\hat{\gamma}_{LR}$ | $\hat{\gamma}_{TSS}(\hat{d}_1)$ | $\hat{\gamma}_{TSS}(\hat{d}_2)$ | $\hat{\gamma}_{TSS}(\hat{d}_3)$ | $\hat{\gamma}_{TSS}(\hat{d}_4)$ | $\hat{\gamma}_{TSS}(\hat{d}_5)$ | $\hat{\gamma}_{TSS}(\hat{d}_6)$ |
|------|------------|-----------------------|----------------------|----------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|
| 15   | 0.50       | 1.524                 | 0.424                | 0.336                | 0.481                         | 0.506                         | 0.497                         | 0.489                         | 0.383                         | 0.413                         | 0.395                         | 0.373                         |
| 1.00 | 3.078      | 1.461                 | 1.317                | 1.779                | 3.295                         | 1.911                         | 3.435                         | 1.704                         | 1.313                         | 1.420                         | 1.366                         | 1.200                         |
| 2.00 | 9.412      | 5.551                 | 5.281                | 7.003                | 9.791                         | 7.628                         | 8.290                         | 6.498                         | 7.077                         | 5.494                         | 6.906                         | 4.422                         |
| 25   | 0.50       | 3.439                 | 0.358                | 0.284                | 0.391                         | 0.419                         | 0.416                         | 0.419                         | 0.326                         | 0.350                         | 0.331                         | 0.326                         |
| 1.00 | 4.665      | 1.257                 | 1.155                | 1.473                | 2.924                         | 1.680                         | 3.147                         | 1.496                         | 1.092                         | 1.211                         | 1.125                         | 1.065                         |
| 2.00 | 9.778      | 4.698                 | 4.692                | 5.753                | 7.938                         | 13.536                        | 10.644                        | 5.591                         | 4.736                         | 4.837                         | 4.846                         | 3.815                         |
| 50   | 0.50       | 3.115                 | 0.117                | 0.104                | 0.124                         | 0.129                         | 0.130                         | 0.129                         | 0.111                         | 0.116                         | 0.112                         | 0.111                         |
| 1.00 | 3.495      | 0.424                 | 0.436                | 0.477                | 0.799                         | 0.856                         | 5.988                         | 0.493                         | 0.390                         | 0.414                         | 0.399                         | 0.376                         |
| 2.00 | 5.106      | 1.554                 | 1.790                | 1.861                | 3.763                         | 2.026                         | 3.840                         | 1.825                         | 1.393                         | 1.511                         | 1.445                         | 1.291                         |
| 100  | 0.50       | 3.069                 | 0.073                | 0.068                | 0.076                         | 0.078                         | 0.078                         | 0.078                         | 0.071                         | 0.073                         | 0.071                         | 0.071                         |
| 1.00 | 3.305      | 0.267                 | 0.279                | 0.290                | 0.332                         | 0.320                         | 0.305                         | 0.245                         | 0.262                         | 0.248                         | 0.248                         | 0.246                         |
| 2.00 | 4.222      | 0.933                 | 1.110                | 1.088                | 1.719                         | 1.614                         | 1.205                         | 0.817                         | 0.906                         | 0.838                         | 0.795                         | 0.795                         |
| 200  | 0.50       | 1.045                 | 0.034                | 0.034                | 0.035                         | 0.035                         | 0.035                         | 0.035                         | 0.033                         | 0.034                         | 0.033                         | 0.033                         |
| 1.00 | 1.150      | 0.127                 | 0.131                | 0.133                | 0.138                         | 0.139                         | 0.138                         | 0.138                         | 0.121                         | 0.125                         | 0.121                         | 0.121                         |
| 2.00 | 1.576      | 0.454                 | 0.528                | 0.505                | 0.507                         | 0.537                         | 0.528                         | 0.526                         | 0.410                         | 0.443                         | 0.418                         | 0.407                         |
| 50   | 0.50       | 3.138                 | 0.354                | 0.344                | 0.395                         | 0.402                         | 0.408                         | 0.405                         | 0.395                         | 0.337                         | 0.352                         | 0.344                         | 0.317                         |
| 1.00 | 4.397      | 1.298                 | 1.410                | 1.549                | 1.536                         | 1.589                         | 1.560                         | 1.474                         | 1.197                         | 1.286                         | 1.240                         | 1.083                         |
| 2.00 | 9.434      | 4.947                 | 5.762                | 6.159                | 6.505                         | 6.219                         | 6.106                         | 5.603                         | 4.436                         | 4.873                         | 4.636                         | 3.967                         |
| 100  | 0.50       | 2.881                 | 0.161                | 0.159                | 0.174                         | 0.179                         | 0.180                         | 0.179                         | 0.177                         | 0.154                         | 0.161                         | 0.156                         | 0.149                         |
| 1.00 | 3.460      | 0.606                 | 0.673                | 0.700                | 0.726                         | 0.714                         | 0.686                         | 0.565                         | 0.601                         | 0.581                         | 0.524                         | 0.524                         |
| 2.00 | 5.694      | 2.251                 | 2.725                | 2.751                | 2.695                         | 2.790                         | 2.740                         | 2.545                         | 2.042                         | 2.217                         | 2.136                         | 1.840                         |
| 200  | 0.50       | 2.767                 | 0.080                | 0.082                | 0.083                         | 0.083                         | 0.083                         | 0.079                         | 0.080                         | 0.080                         | 0.080                         | 0.078                         |
| 1.00 | 3.008      | 0.293                 | 0.317                | 0.322                | 0.327                         | 0.329                         | 0.329                         | 0.322                         | 0.284                         | 0.292                         | 0.290                         | 0.266                         |
| 2.00 | 4.037      | 1.054                 | 1.271                | 1.246                | 1.243                         | 1.278                         | 1.262                         | 1.190                         | 0.981                         | 1.044                         | 1.017                         | 0.889                         |
Table 5. Simulated AMSE values of the estimators when $\rho^2 = 0.95$.

| n  | $\sigma^2$ | $\hat{\gamma}_{OLS}$ | $\hat{\gamma}_{RR}$ | $\hat{\gamma}_{LR}$ | $\hat{\gamma}_{TP}$ | $\hat{\gamma}_{TSS}(\hat{d}_1)$ | $\hat{\gamma}_{TSS}(\hat{d}_2)$ | $\hat{\gamma}_{TSS}(\hat{d}_3)$ | $\hat{\gamma}_{TSS}(\hat{d}_4)$ | $\hat{\gamma}_{TSS}(\hat{d}_5)$ | $\hat{\gamma}_{TSS}(\hat{d}_6)$ |
|----|-----------|-----------------------|----------------------|----------------------|----------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|
| 15 | 0.50      | 4.063                 | 0.790                | 0.575                | 0.940                | 1.138                         | 0.993                         | 1.101                         | 0.922                         | 0.689                         | 0.764                         | 0.716                         | 0.662                         |
|    |           | 1.00                  | 7.394                | 2.942                | 3.250                | 3.690                         | 3.970                         | 3.647                         | 3.138                         | 3.439                         | 2.941                         | 3.890                         | 3.025                         | 2.348                         |
|    |           | 2.00                  | 20.871               | 11.542               | 9.517                | 14.812                        | 27.641                        | 19.906                        | 26.003                        | 13.505                        | 15.091                        | 11.292                        | 14.274                        | 9.064                         |
| 25 | 0.50      | 3.857                 | 0.677                | 0.487                | 0.770                | 0.824                         | 0.822                         | 0.825                         | 0.796                         | 0.599                         | 0.659                         | 0.611                         | 0.591                         |
|    |           | 1.00                  | 6.502                | 2.456                | 1.989                | 2.964                         | 8.937                         | 3.289                         | 8.479                         | 2.923                         | 2.161                         | 2.342                         | 2.230                         | 2.021                         |
|    |           | 2.00                  | 17.121               | 9.475                | 8.103                | 11.767                        | 22.355                        | 13.246                        | 20.356                        | 11.265                        | 10.056                        | 9.334                         | 10.322                        | 7.603                         |
| 50 | 0.50      | 3.257                 | 0.220                | 0.182                | 0.239                | 0.250                         | 0.254                         | 0.251                         | 0.250                         | 0.202                         | 0.215                         | 0.206                         | 0.202                         |
|    |           | 1.00                  | 4.066                | 0.790                | 0.769                | 0.925                         | 1.021                         | 0.976                         | 1.022                         | 0.930                         | 0.686                         | 0.764                         | 0.709                         | 0.671                         |
|    |           | 2.00                  | 7.150                | 2.867                | 3.077                | 3.546                         | 7.022                         | 6.081                         | 8.666                         | 3.368                         | 2.789                         | 2.827                         | 2.922                         | 2.313                         |
| 100| 0.50      | 3.159                 | 0.147                | 0.128                | 0.155                | 0.162                         | 0.163                         | 0.162                         | 0.162                         | 0.137                         | 0.144                         | 0.139                         | 0.139                         |
|    |           | 1.00                  | 3.639                | 0.500                | 0.507                | 0.565                         | 0.654                         | 0.607                         | 0.621                         | 0.587                         | 0.441                         | 0.486                         | 0.490                         | 0.439                         |
|    |           | 2.00                  | 5.537                | 1.809                | 2.072                | 2.183                         | 3.119                         | 2.333                         | 3.042                         | 2.143                         | 1.560                         | 1.743                         | 1.622                         | 1.489                         |
| 200| 0.50      | 3.076                 | 0.070                | 0.066                | 0.072                | 0.074                         | 0.074                         | 0.074                         | 0.074                         | 0.068                         | 0.069                         | 0.068                         | 0.068                         |
|    |           | 1.00                  | 3.279                | 0.248                | 0.258                | 0.267                         | 0.278                         | 0.283                         | 0.279                         | 0.279                         | 0.230                         | 0.243                         | 0.233                         | 0.229                         |
|    |           | 2.00                  | 4.141                | 0.871                | 1.029                | 1.007                         | 1.122                         | 1.069                         | 1.104                         | 1.026                         | 0.763                         | 0.844                         | 0.785                         | 0.749                         |

When $p = 4$

When $p = 8$
Table 6. Simulated AMSE values of the estimators when $\rho^2 = 0.99$.

| $n$ | $\sigma^2$ | $\hat{y}_{OLS}$ | $\hat{y}_{RR}$ | $\hat{y}_{LR}$ | $\hat{y}_{TSS}(d_1)$ | $\hat{y}_{TSS}(d_2)$ | $\hat{y}_{TSS}(d_3)$ | $\hat{y}_{TSS}(d_4)$ | $\hat{y}_{TSS}(d_5)$ |
|-----|-----------|----------------|------------|------------|----------------|----------------|----------------|----------------|----------------|
| 15  | 0.50      | 6.901         | 3.908      | 2.425      | 4.952         | 67.547         | 7.603          | 88.463         | 4.559          |
|     | 1.00      | 23.959        | 14.711     | 9.543      | 19.054        | 27.705         | 25.961         | 27.255         | 17.192         |
|     | 2.00      | 92.146        | 57.799     | 37.692     | 75.354        | 95.526         | 90.772         | 98.003         | 67.644         |
| 25  | 0.50      | 7.360         | 3.071      | 1.769      | 3.743         | 35.965         | 4.349          | 38.964         | 3.654          |
|     | 1.00      | 20.646        | 11.675     | 6.952      | 14.607        | 18.790         | 23.292         | 25.557         | 13.892         |
|     | 2.00      | 73.922        | 46.486     | 28.046     | 58.649        | 94.364         | 71.007         | 91.652         | 55.319         |
| 50  | 0.50      | 2.320         | 0.938      | 0.665      | 0.819         | 1.239          | 1.968          | 1.106          | 0.822          |
|     | 1.00      | 6.345         | 3.602      | 2.773      | 4.495         | 15.097         | 18.467         | 4.229          | 3.527          |
|     | 2.00      | 21.157        | 13.022     | 10.591     | 16.734        | 29.738         | 41.822         | 28.886         | 15.321         |
| 100 | 0.50      | 3.820         | 0.630      | 0.453      | 0.724         | 0.837          | 0.780          | 0.808          | 0.744          |
|     | 1.00      | 6.442         | 2.356      | 1.944      | 2.873         | 21.470         | 4.424          | 25.922         | 2.784          |
|     | 2.00      | 16.504        | 9.116      | 7.975      | 11.419        | 22.079         | 14.224         | 19.654         | 9.729          |
| 200 | 0.50      | 3.382         | 0.313      | 0.254      | 0.357         | 0.363          | 0.358          | 0.307          | 0.293          |
|     | 1.00      | 4.509         | 1.104      | 1.035      | 1.293         | 1.558          | 1.457          | 1.513          | 1.306          |
|     | 2.00      | 8.889         | 4.078      | 4.185      | 5.014         | 18.344         | 6.225          | 16.778         | 4.030          |
| 200 | 0.50      | 3.382         | 0.313      | 0.254      | 0.357         | 0.363          | 0.358          | 0.307          | 0.293          |
|     | 1.00      | 8.889         | 4.078      | 4.185      | 5.014         | 18.344         | 6.225          | 16.778         | 4.030          |

When $p = 4$

When $p = 8$
Table 7. Correlation matrix.

| Variables | $x_1$ | $x_2$ | $x_3$ | $x_4$ |
|-----------|-------|-------|-------|-------|
| $x_1$     | 1.0000|       |       |       |
| $x_2$     | 0.2286| 1.0000|       |       |
| $x_3$     | −0.8241| −0.1392| 1.0000|       |
| $x_4$     | −0.2454| −0.9730| 0.0295| 1.0000|

Note: $P$-values are given in parenthesis.

Table 8. Estimated coefficients and MSE of Portland cement dataset.

| Estimators | $\hat{\beta}_0$ | $\hat{\beta}_1$ | $\hat{\beta}_2$ | $\hat{\beta}_3$ | $\hat{\beta}_4$ | MSE |
|------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----|
| $\hat{\beta}_{OLS}$ | 62.4054 | 1.5511 | 0.5102 | 0.1019 | −0.1441 | 4912.09 |
| $\hat{\beta}_{RR}$ | 42.9860 | 1.7509 | 0.7103 | 0.3062 | 0.0521 | 2706.36 |
| $\hat{\beta}_{LR}$ | 49.9266 | 1.6767 | 0.6394 | 0.2312 | −0.0176 | 3298.65 |
| $\hat{\beta}_{TP}$ | 27.4575 | 1.9106 | 0.8703 | 0.4696 | 0.2090 | 4333.39 |
| $\hat{\beta}_{DTSS(d_4)}$ | 38.0793 | 1.8013 | 0.7609 | 0.3579 | 0.1586 | 2694.80 |
| $\hat{\beta}_{TSS(d_1)}$ | 38.0793 | 1.8013 | 0.7609 | 0.3579 | 0.1017 | 2419.22 |
| $\hat{\beta}_{TSS(d_2)}$ | 42.9859 | 1.7509 | 0.7103 | 0.3062 | 0.0521 | 2706.35 |
| $\hat{\beta}_{TSS(d_3)}$ | 42.9812 | 1.7529 | 0.7124 | 0.3067 | 0.0530 | 2706.02 |
| $\hat{\beta}_{TSS(d_4)}$ | 27.4575 | 1.9106 | 0.8703 | 0.4696 | 0.2090 | 2171.01 |

Note: The OLSE of $\sigma^2$; $\hat{\sigma}^2 = 5.98$.

a severe multicollinearity problem. In Table 8 the result is shown. We can see that among the unbiased and the almost unbiased estimators the $\hat{\beta}_{OLS}$ performs the worst. There is a substantial decrease in the MSE using the $\hat{\beta}_{DTSS(d_4)}$ as compared to the $\hat{\beta}_{OLS}$. Among the biased estimators, the $\hat{\beta}_{LR}$ shows the worst while the $\hat{\beta}_{TSS(d_4)}$ has the lowest MSE. This result is in line with the simulated result since the Liu estimator did not perform well when the error variance is large (in this application it is 5.98). Furthermore, the simulated results demonstrated that the $\hat{\beta}_{TSS(d_4)}$ is the best performing estimator in terms of MSE. Hence, the $\hat{\beta}_{TSS(d_4)}$ and $\hat{\beta}_{DTSS(d_4)}$ are the best options among the biased and the biased-corrected estimators, respectively.

6.2. Biochemical structure dataset

Our second example is taken from Qasim et al. [32,33], where the dependent variable corresponds to the boll weight during the cropping season and experimental material consisted of $n = 32$ upland cotton accessions. The dependent variable is modeled using five highly correlated explanatory variables corresponding to $x_1$ that represents chlorophyll a, $x_2$ signifies chlorophyll b, $x_3$ is defined the total chlorophyll, $x_4$ measures the total soluble protein and $x_5$ implies the total soluble sugar. The linear regression is used to measure the effects of biochemical traits ($x_1$-$x_5$) on boll, and the model estimation can be negatively affected by the multicollinearity problem due to the linear relationship among traits. The condition index, being a measure of the degree of multicollinearity, is considered corresponding to $\sqrt{\lambda_{max}/\lambda_{min}} = 331.44$, which indicates severe multicollinearity [15]. Therefore, the design matrix $X^TX$ is ill-conditioned and the OLSE is no longer a good
approach when the multicollinearity is present. As a remedy, the biased and debiased estimation methods should be used.

In Table 9, the results from the regression model are shown. Here we can see that the MSE is most extensive for the \( \hat{\beta}_{OLS} \) indicating the poor performance in case of severe multicollinearity and many explanatory variables as shown in the simulation study. The second worst estimator is the \( \hat{\beta}_{TP} \). The estimator that minimizes the MSE is the \( \hat{\beta}_{TSS(\hat{d}_4)} \) which is also in line with the simulated results. Among the debiased estimator, the \( \hat{\beta}_{DTSS(\hat{d}_4)} \) works well. Hence, if (almost) unbiasedness is desired, this estimator can also play a relevant role for the analysis of this type of problems. Figure 1 depicts the comparison of DTSSE and TSSE in the form of SB. It is noted that the SB increases as the value of \( d \) increase. However, as expected, the DTSSE gives lower bias than the TSSE. As we mentioned earlier, the DTSSE is a class of debiased estimators which always exhibit the minimum bias. Figure 2 demonstrates the advantages of TSSE over the TPE and OLSE in the sense of scalar MSE. The performance of TSSE is satisfactory as the value of \( d \) increases from 0 to 1. Whereas the TPE demonstrate higher scalar MSE values when \( d \to 1 \). Thus, the proposed class of TSSE

**Table 9. Estimated coefficients and MSE of biochemical structure dataset.**

| Estimators          | \( \hat{\beta}_0 \) | \( \hat{\beta}_1 \) | \( \hat{\beta}_2 \) | \( \hat{\beta}_3 \) | \( \hat{\beta}_4 \) | \( \hat{\beta}_5 \) | MSE     |
|---------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|---------|
| \( \hat{\beta}_{OLS} \) | 2.2243               | 2.1920               | 2.0345               | -1.9738              | -0.0026              | 0.0113               | 3.5693  |
| \( \hat{\beta}_{RR} \)  | 2.2263               | 1.9681               | 1.8211               | -1.7526              | -0.0025              | 0.0103               | 3.0155  |
| \( \hat{\beta}_{LR} \)  | 1.9820               | 1.7506               | 1.5481               | -1.4696              | 0.0112               | 0.0205               | 2.8784  |
| \( \hat{\beta}_{TP} \)  | 2.2247               | 2.1387               | 1.9836               | -1.9211              | -0.0026              | 0.0111               | 3.4039  |
| \( \hat{\beta}_{TSS(\hat{d}_4)} \) | 2.2267               | 1.9228               | 1.7786               | -1.7117              | -0.0025              | 0.0120               | 3.3432  |
| \( \hat{\beta}_{TSS(\hat{d}_1)} \) | 2.2267               | 1.9226               | 1.7777               | -1.7076              | -0.0025              | 0.0101               | 2.9480  |
| \( \hat{\beta}_{TSS(\hat{d}_2)} \) | 2.2263               | 1.9678               | 1.8208               | -1.7523              | -0.0025              | 0.0103               | 3.0150  |
| \( \hat{\beta}_{TSS(\hat{d}_3)} \) | 2.2263               | 1.9626               | 1.8159               | -1.7471              | -0.0025              | 0.0103               | 3.0065  |
| \( \hat{\beta}_{TSS(\hat{d}_4)} \) | 2.2278               | 1.7975               | 1.6585               | -1.5840              | -0.0025              | 0.0095               | 2.8411  |

Note: The OLSE of \( \sigma^2 : \hat{\sigma}^2 = 0.086 \).
Figure 2. Left: Scalar MSE values of OLSE, TPE and TSSE for $\hat{k}_{opt} = 0.000506508$ and $0 \leq d \leq 1$ for Portland cement dataset. Right: Scalar MSE values of OLSE, TPE and TSSE for $\hat{k}_{opt} = 0.003015998$ and $0 \leq d < 1$ for biochemical structure dataset.

depends on the optimal value of $d$. TSSE ($d_4$) yield the lowest MSE, obviously considerably lower than the traditional OLSE.

7. Conclusions

This article introduces a new class of efficient and debiased two-step shrinkage estimators. Proposed estimator includes the OLSE and RRE as special cases to achieve the minimum bias. The proposed estimators are compared theoretically with the OLSE, RRE, LRE and TPE in the sense of MMSE. The proposed TSSE has an advantage over the existing estimators since it exhibits the minimum MMSE under certain conditions. The TSSE exhibit a minimum variance and scalar MSE compared to the TPE suggested by Özkale and Kaciranlar [29]. The TPE does not perform well under certain conditions, and this claim is substantiated by Theorem 3.5 and Figure 2. However, the proposed estimator’s performance is based on the appropriate selection of biasing parameters $k$ and $d$. Therefore, we also proposed an algorithm for selecting the shrinkage parameters. The Monte Carlo simulations are also carried out to evaluate the performance of the proposed estimator where the AMSE is considered as a performance criterion. The performance of the proposed TSSE is satisfactory compared to the other estimator, at least in terms of a smaller bias and MSE. Even though the DTSSE has higher MSE in some situation as compared to other biased estimators, the DTSSE with $\hat{d}_4$ performs better in the sense of smaller SB and MSE. Consequently, based on the theoretical comparisons, simulation results, and empirically relevant real-world applications, we conclude that the class of TSSE with combinations of ($\hat{k}_{opt}, \hat{d}_4$) is performing considerably better than the OLSE, RRE, LRE and the TPE. Therefore, these can be recommended for practitioners.

Notes

1. As explained in Algamal [1] the etymology of 'kidney-inspired algorithm' is related to the organ in the human body. The solutions of the algorithm are filtered in a rate that is calculated based
on the mean of objective functions of all solutions in the current population of each iteration. Thus, the filtered solutions represent the better solutions and is transferred to the filtered blood, while the rest, worse solution, is transferred to waste.

2. A standard Liu estimator generally exhibit a lower mean squared error than the standard ridge regression estimator. Moreover, the Liu estimator can fully address the ill-conditioning problem ([18]).

Acknowledgements

Authors are dedicating this article to those who lost their lives during COVID-19. The Norwegian Research Council / Finance Market Fund is gratefully acknowledged (grant award number 274569) by Pär Sjölander. This article was partially completed while Pär Sjölander was visiting Professor B. M. Golam Kibria at the Department of Mathematics and Statistics, generously hosted by the Florida International University between December 2019 to January 2020. The authors would like to thank the Editor, associate Editor and anonymous referees for their valuable comments and suggestions that improved the quality of this paper greatly.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Muhammad Qasim http://orcid.org/0000-0002-6471-8500
Kristofer Månsson http://orcid.org/0000-0002-4535-3630
B. M. Golam Kibria http://orcid.org/0000-0002-6073-1978

References

[1] Z.Y. Algamal, *A new method for choosing the biasing parameter in ridge estimator for generalized linear model*. Chemom. Intell. Lab. Syst. 183 (2018), pp. 96–101.

[2] M. Amin, M. Qasim, S. Afzal, and K. Naveed, *New ridge estimators in the inverse Gaussian regression: Monte Carlo simulation and application to chemical data*. Commun. Stat. – Simul. Comput. (2020), pp. 1–18. https://doi.org/10.1080/03610918.2020.1797794

[3] M. Arashi, Y. Asar, and B. Yüzbaşı, *SLASSO: A scaled LASSO for multicollinear situations*. J. Stat. Comput. Simul. (2021a), pp. 1–14. https://doi.org/10.1080/00949655.2021.1924174

[4] M. Arashi, M. Norouzirad, S.E. Ahmed, and B. Yüzbaşı, *Rank-based Liu regression*. Comput. Stat. 33 (2018), pp. 1525–1561.

[5] M. Arashi, and M. Roozbeh, *Some improved estimation strategies in high-dimensional semi-parametric regression models with application to riboflavin production data*. Stat. Pap. 60 (2019), pp. 667–686.

[6] M. Arashi, M. Roozbeh, N.A. Hamzah, M. Gasparini, and A.D. Hutson, *Ridge regression and its applications in genetic studies*. Plos one 16 (2021b), pp. e0245376.

[7] R.W. Farebrother, *Further results on the mean square error of ridge regression*. J. R. Stat. Soc. Series B Methodol. 38 (1976), pp. 248–250.

[8] D.G. Gibbons, *A simulation study of some ridge estimators*. J. Am. Stat. Assoc. 76 (1981), pp. 131–139.

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

[10] R.A. Ibikunle, A.F. Lukman, I.F. Titiladunayo, E.A. Akeju, and S.O. Dahunsi, *Modeling and robust prediction of high heating values of municipal solid waste based on ultimate analysis*. Energy Sources Part A (2020a), pp. 1–18. https://doi.org/10.1080/15567036.2020.1841343
[11] R.A. Ibikunle, I.F. Titiladunayo, A.F. Lukman, S.O. Dahunsi, and E.A. Akeju, Municipal solid waste sampling, quantification and seasonal characterization for power evaluation: Energy potential and statistical modelling. Fuel 277 (2020b), pp. 118122.

[12] S. Kaçiranlar, S. Sakallioğlu, F. Akdeniz, G.P. Styan, and H.J. Werner, A new biased estimator in linear regression and a detailed analysis of the widely-analysed dataset on Portland Cement. Sankhyā: The Indian J. Stat. Series B. 61 (1999), pp. 443–459.

[13] K. Kadiyala, A class of almost unbiased and efficient estimators of regression coefficients. Econ. Lett. 16 (1984), pp. 293–296.

[14] L. Kejian, A new class of biased estimate in linear regression. Commun. Stat. – Theory Methods 22 (1993), pp. 393–402.

[15] B.M.G. Kibria, Performance of some new ridge regression estimators. Commun. Stat. – Simul. Comput. 32 (2003), pp. 419–435.

[16] B.M.G. Kibria, and S. Banik, Some ridge regression estimators and their performances. J. Mod. Appl. Stat. Methods. 15 (2016), pp. 206–238.

[17] B.M.G. Kibria, and A.F. Lukman, A new ridge-type estimator for the linear regression model: Simulations and applications. Scientifica 2020 (2020), pp. 1–16.

[18] K. Liu, Using Liu-type estimator to combat collinearity. Commun. Stat. – Theory Methods 32 (2003), pp. 1009–1020.

[19] A.F. Lukman, and K. Ayinde, Review and classi cations of the ridge parameter estimation techniques. Hacettepe J. Math. Stat. 46 (2017), pp. 953–967.

[20] A.F. Lukman, K. Ayinde, and A.S. Ajiboye, Monte Carlo study of some classi cation-based ridge parameter estimators. J. Mod. Appl. Stat. Methods. 16 (2017), pp. 428–451.

[21] A.F. Lukman, K. Ayinde, S. Binuomote, and O.A. Clement, Modified ridge-type estimator to combat multicollinearity: Application to chemical data. J. Chemom. 33 (2019), p. e3125.

[22] S. Mandal, R. Arabi Belagli, A. Mahmoudi, and M. Aminnejad, Stein-type shrinkage estimators in gamma regression model with application to prostate cancer data. Stat. Med. 38 (2019), pp. 4310–4322.

[23] K. Månsson, B.M.G. Kibria, and G. Shukur, On Liu estimators for the logit regression model. Econ. Model. 29 (2012), pp. 1483–1488.

[24] K. Månsson, G. Shukur, and B.M.G. Kibria, A simulation study of some ridge regression estimators under different distributional assumptions. Commun. Stat. – Simul. Comput. 39 (2010), pp. 1639–1670.

[25] R.A. Maronna, Robust ridge regression for high-dimensional data. Technometrics. 53 (2011), pp. 44–53.

[26] G.C. McDonald, and D.I. Galarneau, A Monte Carlo evaluation of some ridge-type estimators. J. Am. Stat. Assoc. 70 (1975), pp. 407–416.

[27] J.P. Newhouse, and S.D. Oman, An evaluation of ridge estimators. A report prepared for United States Air Force project RAND, 1971.

[28] K. Ohtani, On small sample properties of the almost unbiased generalized ridge estimator. Commun. Stat. – Theory Methods 15 (1986), pp. 1571–1578.

[29] M.R. Özkale, and S. Kaçiranlar, The restricted and unrestricted two-parameter estimators. Commun. Stat. – Theory Methods 36 (2007), pp. 2707–2725.

[30] J.J. Peterson, and A.M. Kuhn, Ridge analysis with noise variables. Technometrics. 47 (2005), pp. 274–283.

[31] M. Qasim, M. Amin, and M. Amanullah, On the performance of some new Liu parameters for the gamma regression model. J. Stat. Comput. Simul. 88 (2018), pp. 3065–3080.

[32] M. Qasim, M. Amin, and T. Omer, Performance of some new Liu parameters for the linear regression model. Commun. Stat. – Theory Methods 49 (2020a), pp. 4178–4196.

[33] M. Qasim, M. Amin, and M.K.S. Sarwar, Effect of different biochemical traits on seed cotton yield: An application of Liu linear regression. J. Anim. Plant Sci. 30 (2020b), pp. 1533–1539.

[34] M. Qasim, K. Månsson, and B.M.G. Kibria, On some beta ridge regression estimators: Method, simulation and application. J. Stat. Comput. Simul. 91 (2021), pp. 1699–1712.

[35] A.K.M.E. Saleh, M. Arashi, and B.M.G. Kibria, Theory of Ridge Regression Estimation with Applications, Wiley, New York, 2019.
[36] G. Trenkler, and H. Toutenburg, *Mean squared error matrix comparisons between biased estimators—An overview of recent results*. Stat. Pap. 31(1) (1990), pp. 165–179.

[37] G. Tutz, and H. Binder, *Boosting ridge regression*. Comput. Stat. Data. Anal. 51 (2007), pp. 6044–6059.

[38] V.R. Uslu, E. Egrioglu, and E. Bas, *Finding optimal value for the shrinkage parameter in ridge regression via particle swarm optimization*. Am. J. Intell. Syst. 4 (2014), pp. 142–147.

[39] S.G. Wang, M.X. Wu, and Z.Z. Jia, *Matrix Inequalities*. 2nd ed., Chinese Science Press, Beijing, 2006.

[40] H. Woods, H.H. Steinour, and H.R. Starke, *Effect of composition of Portland cement on heat evolved during hardening*. Ind. Eng. Chem. 24 (1932), pp. 1207–1214.

[41] J. Wu, and H. Yang, *Efficiency of an almost unbiased two-parameter estimator in linear regression model*. Statistics 47 (2013), pp. 535–545.