Research Article

On the Weighted Mixed Almost Unbiased Ridge Estimator in Stochastic Restricted Linear Regression

Chaolin Liu, Hu Yang, and Jibo Wu

College of Mathematics and Statistics, Chongqing University, Chongqing 401331, China

Correspondence should be addressed to Chaolin Liu; lcl@cqu.edu.cn

We introduce the weighted mixed almost unbiased ridge estimator (WMAURE) based on the weighted mixed estimator (WME) (Trenkler and Toutenburg 1990) and the almost unbiased ridge estimator (AURE) (Akdeniz and Erol 2003) in linear regression model. We discuss superiorities of the new estimator under the quadratic bias (QB) and the mean square error matrix (MSEM) criteria. Additionally, we give a method about how to obtain the optimal values of parameters k and w. Finally, theoretical results are illustrated by a real data example and a Monte Carlo study.

1. Introduction

Consider the linear regression model

$$y = X\beta + \varepsilon, \tag{1}$$

where $y = (y_1, y_2, \ldots, y_n)^T$ is an n-dimensional response vector, $X = (X_1, X_2, \ldots, X_n)^T$ with $X_i = (x_{i1}, x_{i2}, \ldots, x_{ip})^T$ is a known $n \times p$ matrix of full column rank, β is a $p \times 1$ vector of unknown parameters, and ε is an $n \times 1$ vector of errors with expectation $E(\varepsilon) = 0$ and covariance matrix $\text{Cov}(\varepsilon) = \sigma^2 I_n$, I_n is an identity matrix of order $n \times n$.

It is well known that the ordinary least squares estimator (LS) for β is given by

$$\widehat{\beta}_{\text{LS}} = \left(X^T X\right)^{-1} X^T y, \tag{2}$$

which has been treated as the best estimator for a long time. However, many results have proved that LS is no longer a good estimator when the multicollinearity is present in model (1). To tackle this problem, some suitable biased estimators have been developed, such as principal component regression estimator (PCR) [1], ordinary ridge estimator (RE) [2], $r - k$ class estimator [3], Liu estimator (LE) [4], and $r - d$ class estimator [5]. Kadiyala [6] introduced a class of almost unbiased shrinkage estimator which can be not only almost unbiased but also more efficient than the LS. Singh et al. [7]

introduced the almost unbiased generalized ridge estimator by the jackknife procedure, and Akdeniz and Kaçiranlar [8] studied the almost unbiased generalized Liu estimator. By studying bias corrected estimators of the RE and the LE, Akdeniz and Erol [9] discussed the almost unbiased ridge estimator (AURE) and the almost unbiased Liu estimator (AULE).

An alternative technique to tackle the multicollinearity is to consider the parameter estimator in addition to the sample information, such as some exact or stochastic restrictions on unknown parameters. When additional stochastic linear restrictions on unknown parameters are assumed to be held, Durbin [10], Theil and Goldberger [11], and Theil [12] proposed the ordinary mixed estimator (OME). Hubert and Wijekoon [13] proposed the stochastic restricted Liu estimator, and Yang and Xu [14] obtained a new stochastic restricted Liu estimator. By grafting the RE into the mixed estimation procedure, Li and Yang [15] introduced the stochastic restricted ridge estimator. When the prior information and the sample information are not equally important, Schaffrin and Toutenburg [16] studied the weighted mixed regression and developed the weighted mixed estimator (WME). Li and Yang [17] grafted the RE into the weighted mixed estimation procedure and proposed the weighted mixed ridge estimator (WMRE).

In this paper, by combining the WME and the AURE, we propose a weighted mixed almost unbiased ridge estimator

(WMAURE) for unknown parameters in a linear regression model when additional stochastic linear restriction is supposed to be held. Furthermore, we discuss the performance of the new estimator over the LS, WME, AURE, and WMRE with respect to the quadratic bias (QB) and the mean square error matrix (MSEM) criteria.

The rest of the paper is organized as follows. In Section 2, we describe the statistical model and propose the weighted mixed almost unbiased ridge estimator. We compare the new estimator with the weighted mixed ridge estimator and the almost unbiased ridge estimator under the quadratic bias criterion in Section 3. In Section 4, superiorities of the proposed estimator over relative estimators are considered under the mean square error matrix criterion. In Section 5, the selection of parameters k and w is discussed. Finally, to justify the superiority of the new estimator, we perform a real data example and a Monte Carlo simulation study in Section 6. We give some conclusions in Section 7.

2. The Proposed Estimator

The ordinary ridge estimator proposed by Hoerl and Kennard [2] is defined as

$$\widehat{\beta}_{\mathrm{RE}}(k) = (S + kI)^{-1} X^T y = S_k^{-1} X^T y, \tag{3}$$

where $S_k = S + kI, k > 0$. Let $T_k = (I + kS^{-1})^{-1} = S_k^{-1} S = S S_k^{-1}$; we may rewrite $\widehat{\beta}_{\mathrm{RE}}(k)$ as

$$\widehat{\beta}_{\mathrm{RE}}(k) = T_k \widehat{\beta}_{\mathrm{LS}} = T_k S^{-1} X^T y = S^{-1} T_k X^T y. \tag{4}$$

The almost unbiased ridge estimator obtained by Akdeniz and Erol [9] is denoted as

$$\widehat{\beta}_{\mathrm{AURE}}(k) = \left(I - k^2 S_k^{-2}\right) \widehat{\beta}_{\mathrm{LS}} = \left(I + k S_k^{-1}\right) T_k \widehat{\beta}_{\mathrm{LS}}. \tag{5}$$

In addition to model (1), let us give some prior information about β in the form of a set of J which is independent stochastic linear restriction

$$r = R\beta + e, \quad e \sim \left(0, \sigma^2 W\right), \tag{6}$$

where R is a $J \times p$ known matrix of rank J, e is a $J \times 1$ vector of disturbances with expectation 0 and covariance matrix $\sigma^2 W$, W is supposed to be known and positive definite, and the $J \times 1$ vector r can be interpreted as a random variable with expectation $E(r) = R\beta$. Then, we can derive that (6) does not hold exactly but in the mean. We assume r to be a realized value of the random vector, so that all expectations are conditional on r [18]. We will not separately mention this in the following discussions. Furthermore, it is also supposed that ε is stochastically independent of e.

For the restricted model specified by (1) and (6), the OME introduced by Durbin [10], Theil and Goldberger [11], and Theil [12] is defined as

$$\widehat{\beta}_{\mathrm{OME}} = \left(S + R^T W^{-1} R\right)^{-1} \left(X^T y + R^T W^{-1} r\right). \tag{7}$$

When the prior information and the sample information are not equally important, Schaffrin and Toutenburg [16] considered the WME which is denoted as

$$\widehat{\beta}_{\mathrm{WME}}(w) = \left(S + w R^T W^{-1} R\right)^{-1} \left(X^T y + w R^T W^{-1} r\right), \tag{8}$$

where w ($0 \le w \le 1$) is a nonstochastic and nonnegative scalar weight.

Note that

$$\left(S + w R^T W^{-1} R\right)^{-1} = S^{-1} - w S^{-1} R^T \left(W + w R S^{-1} R^T\right)^{-1} R S^{-1}, \tag{9}$$

$$\left(S^{-1} - w S^{-1} R^T \left(W + w R S^{-1} R^T\right)^{-1} R S^{-1}\right) w R^T W^{-1} r = w S^{-1} R^T \left(W + w R S^{-1} R^T\right)^{-1} r. \tag{10}$$

Then, the WME (8) can be rewritten as

$$\widehat{\beta}_{\mathrm{WME}}(w) = \widehat{\beta}_{\mathrm{LS}} + w S^{-1} R^T \left(W + w R S^{-1} R^T\right)^{-1} \left(r - R\widehat{\beta}_{\mathrm{LS}}\right). \tag{11}$$

Additionally, by combining the WME and RE, Li and Yang [17] obtained the WMRE which is defined as

$$\widehat{\beta}_{\mathrm{WMRE}}(w, k) = \widehat{\beta}_{\mathrm{RE}}(k) + w S^{-1} R^T \left(W + w R S^{-1} R^T\right)^{-1} \left(r - R\widehat{\beta}_{\mathrm{RE}}(k)\right). \tag{12}$$

The WMRE also can be rewritten as

$$\widehat{\beta}_{\mathrm{WMRE}}(w, k) = \left(S + w R^T W^{-1} R\right)^{-1} \left(T_k X^T y + w R^T W^{-1} r\right). \tag{13}$$

Now, based on the WME [16] and the AURE [9], we can define the following weighted mixed almost unbiased ridge estimator:

$$\begin{aligned}
\widehat{\beta}_{\mathrm{WMAURE}}(w, k) &= \widehat{\beta}_{\mathrm{AURE}}(k) + w S^{-1} R^T \left(W + w R S^{-1} R^T\right)^{-1} \\
&\quad \times \left(r - R\widehat{\beta}_{\mathrm{AURE}}(k)\right) \\
&= \left(I - k^2 S_k^{-2}\right) \widehat{\beta}_{\mathrm{LS}} + w S^{-1} R^T \left(W + w R S^{-1} R^T\right)^{-1} \\
&\quad \times \left(r - R\left(I - k^2 S_k^{-2}\right) \widehat{\beta}_{\mathrm{LS}}\right) \\
&= \left(S^{-1} - w S^{-1} R^T \left(W + w R S^{-1} R^T\right)^{-1} R S^{-1}\right) \\
&\quad \times \left(I - k^2 S_k^{-2}\right) X^T y + w S^{-1} R^T \left(W + w R S^{-1} R^T\right)^{-1} r,
\end{aligned} \tag{14}$$

which is according to the way in [17].

Using (10), (14) can be rewritten as

$$\widehat{\beta}_{\text{WMAURE}}(w,k)$$

$$= \left(S^{-1} - wS^{-1}R^T\left(W + wRS^{-1}R^T\right)^{-1}RS^{-1}\right)$$

$$\times \left(\left(I - k^2S_k^{-2}\right)X^T y + wR^T W^{-1}r\right)$$

$$= \left(S + wR^T W^{-1}R\right)^{-1}\left(\left(I - k^2S_k^{-2}\right)X^T y + wR^T W^{-1}r\right). \tag{15}$$

From the definition of $\widehat{\beta}_{\text{WMAURE}}(w,k)$, it can be seen that $\widehat{\beta}_{\text{WMAURE}}(w,k)$ is a general estimator, and as special cases of it, the WME, LS, and AURE can be described as

$$\widehat{\beta}_{\text{WMAURE}}(w,0) = \widehat{\beta}_{\text{WME}}(w),$$

$$\widehat{\beta}_{\text{WMAURE}}(0,0) = \widehat{\beta}_{\text{LS}}, \tag{16}$$

and if $R = 0$,

$$\widehat{\beta}_{\text{WMAURE}}(w,k) = \widehat{\beta}_{\text{AURE}}(k). \tag{17}$$

It is easy to compute expectation values and covariance matrices of the LS, WME, WMRE, AURE, and WMAURE as

$$E\left(\widehat{\beta}_{\text{LS}}\right) = \beta,$$

$$\text{COV}\left(\widehat{\beta}_{\text{LS}}\right) = \sigma^2 S^{-1},$$

$$E\left(\widehat{\beta}_{\text{WME}}(w)\right) = \beta,$$

$$\text{COV}\left(\widehat{\beta}_{\text{WME}}(w)\right) = \sigma^2 A\left(S + w^2 R^T W^{-1}R\right)A,$$

$$E\left(\widehat{\beta}_{\text{WMRE}}(w,k)\right) = \beta + A\left(T_k - I\right)S\beta,$$

$$\text{COV}\left(\widehat{\beta}_{\text{WMRE}}(w,k)\right) = \sigma^2 A\left(T_k S T_k + w^2 R^T W^{-1}R\right)A,$$

$$E\left(\widehat{\beta}_{\text{AURE}}(k)\right) = \left(I - k^2 S_k^{-2}\right)\beta = \left(I + kS_k^{-1}\right)T_k\beta,$$

$$\text{COV}\left(\widehat{\beta}_{\text{AURE}}(k)\right) = \sigma^2\left(I - k^2 S_k^{-2}\right)S^{-1}\left(I - k^2 S_k^{-2}\right)$$

$$= \sigma^2\left(I + kS_k^{-1}\right)T_k S^{-1}T_k\left(I + kS_k^{-1}\right),$$

$$E\left(\widehat{\beta}_{\text{WMAURE}}(w,k)\right) = \beta - k^2 AS_k^{-2}S\beta,$$

$$\text{COV}\left(\widehat{\beta}_{\text{WMAURE}}(w,k)\right)$$

$$= \sigma^2 A\left(\left(I - k^2 S_k^{-2}\right)S\left(I - k^2 S_k^{-2}\right) + w^2 R^T W^{-1}R\right)A, \tag{18}$$

where $A = \left(S + wR^T W^{-1}R\right)^{-1}$.

In the rest of the sections, we intend to study the performance of the new estimator over relative estimators under the quadratic bias and the mean square error matrix criteria.

3. Quadratic Bias Comparison of Estimators

In this section, quadratic bias comparisons are performed among the AURE, WMRE, and WMAURE. Let $\widehat{\beta}$ be the estimator of β, then the quadratic bias of $\widehat{\beta}$ is defined as $\text{QB}(\widehat{\beta}) = \text{Bias}(\widehat{\beta})^T\text{Bias}(\widehat{\beta})$, where $\text{Bias}(\widehat{\beta}) = E(\widehat{\beta}) - \beta$. Based on the definition of quadratic bias, we can easily get quadratic biases of AURE, WMRE, and WMAURE:

$$\text{QB}\left(\widehat{\beta}_{\text{AURE}}(k)\right) = \text{Bias}\left(\widehat{\beta}_{\text{AURE}}(k)\right)^T\text{Bias}\left(\widehat{\beta}_{\text{AURE}}(k)\right)$$

$$= k^4\beta^T S_k^{-4}\beta,$$

$$\text{QB}\left(\widehat{\beta}_{\text{WMRE}}(w,k)\right)$$

$$= \text{Bias}\left(\widehat{\beta}_{\text{WMRE}}(w,k)\right)^T\text{Bias}\left(\widehat{\beta}_{\text{WMRE}}(w,k)\right)$$

$$= k^2\beta^T SS_k^{-1}A^2 S_k^{-1}S\beta,$$

$$\text{QB}\left(\widehat{\beta}_{\text{WMAURE}}(w,k)\right)$$

$$= \text{Bias}\left(\widehat{\beta}_{\text{WMAURE}}(w,k)\right)^T\text{Bias}\left(\widehat{\beta}_{\text{WMAURE}}(w,k)\right)$$

$$= k^4\beta^T SS_k^{-2}A^2 S_k^{-2}S\beta. \tag{19}$$

3.1. Quadratic Bias Comparison between the AURE and WMAURE.
Here, we focus on the quadratic bias comparison between the AURE and WMAURE. The difference of quadratic biases can be derived as

$$\Delta_1 = \text{QB}\left(\widehat{\beta}_{\text{AURE}}(k)\right) - \text{QB}\left(\widehat{\beta}_{\text{WMAURE}}(w,k)\right)$$

$$= k^4\beta^T S_k^{-4}\beta - k^4\beta^T SS_k^{-2}A^2 S_k^{-2}S\beta$$

$$= k^4\beta^T S_k^{-4}\beta - k^4\beta^T S_k^{-2}SA^2 SS_k^{-2}\beta$$

$$= k^4\beta^T S_k^{-2}\left(I - SA^2 S\right)S_k^{-2}\beta. \tag{20}$$

Firstly, we just consider $I - SA^2 S$. Note that $\left(I - SA^2 S\right)^T = I - SA^2 S$ and

$$I - SA^2 S = SS^{-2}S - SA^2 S$$

$$= S\left(S^{-2} - A^2\right)S. \tag{21}$$

It can be seen that $A^{-2} - S^2 = \left(S + wR^T W^{-1}R\right)^2 - S^2 > 0$, namely, $A^{-2} > S^2$. Then we can get

$$S^{-2} - A^2 > 0. \tag{22}$$

Thus, we have $I - SA^2 S > 0$ by (21) and (22). Therefore, we can derive that $\Delta_1 = \text{QB}(\widehat{\beta}_{\text{AURE}}(k)) - \text{QB}(\widehat{\beta}_{\text{WMAURE}}(w,k)) > 0$ and the $\widehat{\beta}_{\text{WMAURE}}(w,k)$ outperforms the $\widehat{\beta}_{\text{AURE}}(k)$ according to the quadratic bias criterion.

Based on the above analysis, we can derive the following theorem.

Theorem 1. *According to the quadratic bias criterion, the WMAURE performs better than the AURE.*

3.2. Quadratic Bias Comparison between the WMRE and WMAURE. Similarly, the quadratic bias comparison between the WMRE and WMAURE will be discussed. The difference of quadratic biases of both estimators can be obtained by

$$
\begin{aligned}
\Delta_2 &= \mathrm{QB}\left(\widehat{\beta}_{\mathrm{WMRE}}\left(w,k\right)\right) - \mathrm{QB}\left(\widehat{\beta}_{\mathrm{WMAURE}}\left(w,k\right)\right) \\
&= k^2\beta^T SS_k^{-1} A^2 S_k^{-1} S\beta - k^4\beta^T SS_k^{-2} A^2 S_k^{-2} S\beta \qquad (23) \\
&= k^2\beta^T SS_k^{-1}\left(A^2 - k^2 S_k^{-1} A^2 S_k^{-1}\right) S_k^{-1} S\beta.
\end{aligned}
$$

We consider $A^2 - k^2 S_k^{-1} A^2 S_k^{-1}$. Note that $(A^2 - k^2 S_k^{-1} A^2 S_k^{-1})^T = A^2 - k^2 S_k^{-1} A^2 S_k^{-1}$ and

$$
\begin{aligned}
&A^2 - k^2 S_k^{-1} A^2 S_k^{-1} \\
&\quad = A^2 - kS_k^{-1} A^2 + kS_k^{-1} A^2 - k^2 S_k^{-1} A^2 S_k^{-1} \qquad (24) \\
&\quad = \left(I - kS_k^{-1}\right) A^2 + kS_k^{-1} A^2 \left(I - kS_k^{-1}\right).
\end{aligned}
$$

For $S = X^T X > 0$, there exists an orthogonal matrix P such that $S = P\Lambda P^T$, where $\Lambda = \mathrm{diag}(\lambda_1,\ldots,\lambda_p)$ and $\lambda_1 \geq \cdots \geq \lambda_p$ denote the ordered eigenvalues of S. Therefore, we can easily compute that $I - kS_k^{-1} = P\,\mathrm{diag}(\tau_1,\ldots,\tau_p)P^T$, where $\tau_i = 1 - (k/(\lambda_i + k)) = \lambda_i/(\lambda_i + k), i = 1,\ldots,p$. For $k > 0, \lambda_i > 0$, we obtain $\tau_i > 0$ which means that $I - kS_k^{-1} > 0$. Thus, we have $(I - kS_k^{-1})^{1/2} A^2 (I - kS_k^{-1})^{1/2} > 0$ and $(I - kS_k^{-1})^{1/2} kS_k^{-1} A^2 (I - kS_k^{-1})^{1/2} > 0$. Note that $(I - kS_k^{-1})A^2$, $kS_k^{-1} A^2 (I - kS_k^{-1})$ and $(I - kS_k^{-1})^{1/2} A^2 (I - kS_k^{-1})^{1/2}$, $(I - kS_k^{-1})^{1/2} kS_k^{-1} A^2 (I - kS_k^{-1})^{1/2}$ have same nonzero eigenvalues, respectively, which means that $A^2 - k^2 S_k^{-1} A^2 S_k^{-1} > 0$.

Therefore, we can derive that $\Delta_2 = \mathrm{QB}(\widehat{\beta}_{\mathrm{WMRE}}(w,k)) - \mathrm{QB}(\widehat{\beta}_{\mathrm{WMAURE}}(w,k)) > 0$ and the $\widehat{\beta}_{\mathrm{WMAURE}}(w,k)$ performs better than the $\widehat{\beta}_{\mathrm{WMRE}}(w,k)$ under the quadratic bias criterion.

We can get the following theorem.

Theorem 2. *According to the quadratic bias criterion, the WMAURE outperforms the WMRE.*

4. Mean Square Error Matrix Comparisons of Estimators

In this section, We will compare the proposed estimator with relative estimators under the mean square error matrix (MSEM) criterion.

For the sake of convenience, we list some lemmas needed in the following discussions.

Lemma 3. *Let M be a positive definite matrix, namely, $M > 0$; let α be a vector, then $M - \alpha\alpha^T \geq 0$ if and only if $\alpha^T M^{-1}\alpha \leq 1$.*

Proof. See [19]. □

Lemma 4. *Let $\widehat{\beta}_i = A_i y$, $i = 1,2$ be two competing homogeneous linear estimators of β. Suppose that $D = \mathrm{COV}(\widehat{\beta}_1) - \mathrm{COV}(\widehat{\beta}_2) > 0$, then $\mathrm{MSEM}(\widehat{\beta}_1) - \mathrm{MSEM}(\widehat{\beta}_2) \geq 0$ if and only if $d_2^T(D + d_1 d_1^T)^{-1} d_2 \leq 1$, where $\mathrm{MSEM}(\widehat{\beta}_i)$, d_i denote the MSEM and bias vector of $\widehat{\beta}_i$, respectively.*

Proof. See [20]. □

Lemma 5. *Let two $n \times n$ matrices $M > 0$, $N \geq 0$, then $M > N \Leftrightarrow \lambda_1(NM^{-1}) < 1$.*

Proof. See [18]. □

Firstly, the MSEM of an estimator $\widehat{\beta}$ is defined as

$$
\begin{aligned}
\mathrm{MSEM}\left(\widehat{\beta}\right) &= E\left(\left(\widehat{\beta} - \beta\right)\left(\widehat{\beta} - \beta\right)^T\right) \\
&= \mathrm{COV}\left(\widehat{\beta}\right) + \mathrm{Bias}\left(\widehat{\beta}\right)\mathrm{Bias}\left(\widehat{\beta}\right)^T.
\end{aligned} \qquad (25)
$$

For two given estimators $\widehat{\beta}_1$ and $\widehat{\beta}_2$, the estimator $\widehat{\beta}_2$ is said to be superior to $\widehat{\beta}_1$ under the MSEM criterion if and only if

$$
M\left(\widehat{\beta}_1, \widehat{\beta}_2\right) = \mathrm{MSEM}\left(\widehat{\beta}_1\right) - \mathrm{MSEM}\left(\widehat{\beta}_2\right) \geq 0. \qquad (26)
$$

The scalar mean square error (MSE) is defined as $\mathrm{MSE}(\widehat{\beta}) = \mathrm{tr}(\mathrm{MSEM}(\widehat{\beta}))$. It is well known that the MSEM criterion is superior over the MSE criterion; we just compare the MSEM of the WMAURE with other relative estimators.

Now, from (18), we can easily obtain the MSEM of the WME, WMRE, AURE, and WMAURE as follows:

$$
\mathrm{MSEM}\left(\widehat{\beta}_{\mathrm{LS}}\right) = \sigma^2 S^{-1}, \qquad (27)
$$

$$
\mathrm{MSEM}\left(\widehat{\beta}_{\mathrm{WME}}\left(w\right)\right) = \sigma^2 A\left(S + w^2 R^T W^{-1} R\right) A, \qquad (28)
$$

$$
\begin{aligned}
&\mathrm{MSEM}\left(\widehat{\beta}_{\mathrm{WMRE}}\left(w,k\right)\right) \\
&\quad = \sigma^2 A\left(T_k S T_k + w^2 R^T W^{-1} R\right) A + b_1 b_1^T,
\end{aligned} \qquad (29)
$$

$$
\begin{aligned}
&\mathrm{MSEM}\left(\widehat{\beta}_{\mathrm{AURE}}\left(k\right)\right) \\
&\quad = \sigma^2 \left(I + kS_k^{-1}\right) T_k S^{-1} T_k \left(I + kS_k^{-1}\right) + b_2 b_2^T,
\end{aligned} \qquad (30)
$$

$$
\begin{aligned}
&\mathrm{MSEM}\left(\widehat{\beta}_{\mathrm{WMAURE}}\left(w,k\right)\right) \\
&\quad = \sigma^2 A\left(\left(I + kS_k^{-1}\right) T_k S T_k \left(I + kS_k^{-1}\right) + w^2 R^T W^{-1} R\right) \\
&\qquad \times A + b_3 b_3^T,
\end{aligned} \qquad (31)
$$

where $b_1 = A(T_k - I)S\beta$, $b_2 = -k^2 S_k^{-2}\beta$ and $b_3 = -k^2 AS_k^{-2} S\beta$.

4.1. MSEM Comparison of the WME and WMAURE. To compare MSEM values between the WMAURE and WME.

Firstly, from (28) and (31), the difference of MSEM values between the WME and WMAURE can be gained by

$$
\begin{aligned}
&M\left(\widehat{\beta}_{\mathrm{WME}}(w), \widehat{\beta}_{\mathrm{WMAURE}}(w, k)\right)\\
&\quad = \mathrm{MSEM}\left(\widehat{\beta}_{\mathrm{WME}}(w)\right) - \mathrm{MSEM}\left(\widehat{\beta}_{\mathrm{WMAURE}}(w, k)\right)\\
&\quad = \sigma^2 A\left(S + w^2 R^T W^{-1} R\right) A - \sigma^2 A\\
&\qquad \times \left(\left(I + kS_k^{-1}\right) T_k ST_k \left(I + kS_k^{-1}\right) + w^2 R^T W^{-1} R\right) A - b_3 b_3^T\\
&\quad = \sigma^2 A\left(S - \left(I + kS_k^{-1}\right) T_k ST_k \left(I + kS_k^{-1}\right)\right) A - b_3 b_3^T\\
&\quad = \sigma^2 D_1 - b_3 b_3^T,
\end{aligned}
\tag{32}
$$

where $D_1 = A(S - (I + kS_k^{-1}) T_k ST_k (I + kS_k^{-1})) A$.

Theorem 6. *The WMAURE $\widehat{\beta}_{WMAURE}(w, k)$ is superior to the WME $\widehat{\beta}_{WME}(w)$ under MSEM criterion, namely, $M(\widehat{\beta}_{WME}(w), \widehat{\beta}_{WMAURE}(w, k)) \geq 0$ if and only if $b_3^T D_1^{-1} b_3 \leq \sigma^2$.*

Proof. Note that $S = P\Lambda P^T$. We can easily compute that $S - (I + kS_k^{-1}) T_k ST_k (I + kS_k^{-1}) = P\Gamma^{(1)} P^T = P\operatorname{diag}(\tau_1^{(1)}, \ldots, \tau_p^{(1)}) P^T$, where $\Gamma^{(1)} = \Lambda - (I + k(\Lambda + kI)^{-1})(I + k\Lambda^{-1})^{-1} \Lambda (I + k\Lambda^{-1})^{-1}(I + k(\Lambda + kI)^{-1})$ and

$$
\begin{aligned}
\tau_i^{(1)} &= \lambda_i - \left(1 + k(\lambda_i + k)^{-1}\right)\left(1 + k\lambda_i^{-1}\right)^{-1}\\
&\quad \times \lambda_i \left(1 + k\lambda_i^{-1}\right)^{-1}\left(1 + k(\lambda_i + k)^{-1}\right)\\
&= \lambda_i - \frac{\lambda_i + 2k}{\lambda_i + k} \frac{\lambda_i}{\lambda_i + k} \lambda_i \frac{\lambda_i}{\lambda_i + k} \frac{\lambda_i + 2k}{\lambda_i + k}\\
&= \frac{\left(k^4 + 4k^3 \lambda_i + 2k^2 \lambda_i^2\right) \lambda_i}{\left(\lambda_i + k\right)^4} > 0, \quad i = 1, \ldots, p,
\end{aligned}
\tag{33}
$$

which means $S - (I + kS_k^{-1}) T_k ST_k (I + kS_k^{-1}) > 0$. Observing that $A = (S + wR^T W^{-1} R)^{-1} > 0$, we have $D_1 = A(S - (I + kS_k^{-1}) T_k ST_k (I + kS_k^{-1})) A > 0$.

Applying Lemma 3, we can get that $M(\widehat{\beta}_{\mathrm{WME}}(w), \widehat{\beta}_{\mathrm{WMAURE}}(w, k)) \geq 0$ if and only if $b_3^T D_1^{-1} b_3 \leq \sigma^2$. This completes the proof. □

4.2. MSEM Comparison of the WMRE and WMAURE.

Similarly, we compare MSEM values between the WMAURE and WMRE. Firstly, from (29) and (31), the difference of MSEM values between the WMRE and the WMAURE can be computed by

$$
\begin{aligned}
&M\left(\widehat{\beta}_{\mathrm{WMAURE}}(w, k), \widehat{\beta}_{\mathrm{WMRE}}(w, k)\right)\\
&\quad = \mathrm{MSEM}\left(\widehat{\beta}_{\mathrm{WMAURE}}(w, k)\right) - \mathrm{MSEM}\left(\widehat{\beta}_{\mathrm{WMRE}}(w, k)\right)
\end{aligned}
$$

$$
\begin{aligned}
&= \sigma^2 A\left(\left(I + kS_k^{-1}\right) T_k ST_k \left(I + kS_k^{-1}\right) + w^2 R^T W^{-1} R\right) A\\
&\quad + b_3 b_3^T - \sigma^2 A\left(T_k ST_k + w^2 R^T W^{-1} R\right) A - b_1 b_1^T\\
&= \sigma^2 A\left(\left(I + kS_k^{-1}\right) T_k ST_k \left(I + kS_k^{-1}\right) - T_k ST_k\right) A - b_1 b_1^T\\
&= \sigma^2 D_2 + b_3 b_3^T - b_1 b_1^T,
\end{aligned}
\tag{34}
$$

where $D_2 = A((I + kS_k^{-1}) T_k ST_k (I + kS_k^{-1}) - T_k ST_k) A$.

Theorem 7. *The WMAURE $\widehat{\beta}_{WMAURE}(w, k)$ $(k > 0)$ is superior to the WMRE $\widehat{\beta}_{WMRE}(w, k)$ under MSEM criterion, namely, $M(\widehat{\beta}_{WMAURE}(w, k), \widehat{\beta}_{WMRE}(w, k)) < 0$ if and only if $b_1^T(\sigma^2 D_2 + b_3 b_3^T)^{-1} b_1 > 1$.*

Proof. Firstly, we prove $(I + kS_k^{-1}) T_k ST_k (I + kS_k^{-1}) - T_k ST_k > 0$.

Note that $S = P\Lambda P^T$, we can compute $(I + kS_k^{-1}) T_k ST_k (I + kS_k^{-1}) - T_k ST_k = P\Gamma^{(2)} P^T = P\operatorname{diag}(\tau_1^{(2)}, \ldots, \tau_p^{(2)}) P^T$, where

$$
\begin{aligned}
\Gamma^{(2)} &= \left(I + k(\Lambda + kI)^{-1}\right)\left(1 + k\Lambda^{-1}\right)^{-1} \Lambda \left(1 + k\Lambda^{-1}\right)^{-1}\\
&\quad \times \left(I + k(\Lambda + kI)^{-1}\right) - \left(1 + k\Lambda^{-1}\right)^{-1} \Lambda \left(1 + k\Lambda^{-1}\right)^{-1},\\
\tau_i^{(2)} &= \left(1 + k(\lambda_i + k)^{-1}\right)\left(1 + \frac{k}{\lambda_i}\right)^{-1} \lambda_i \left(1 + \frac{k}{\lambda_i}\right)^{-1}\\
&\quad \times \left(1 + k(\lambda_i + k)^{-1}\right) - \left(1 + \frac{k}{\lambda_i}\right)^{-1} \lambda_i \left(1 + \frac{k}{\lambda_i}\right)^{-1}\\
&= \left(1 + \frac{k}{\lambda_i}\right)^{-1} \lambda_i \left(1 + \frac{k}{\lambda_i}\right)^{-1}\left(\frac{\lambda_i + 2k}{\lambda_i + k}\right)^2 > 0,\\
&\hspace{6em} i = 1, \ldots, p.
\end{aligned}
\tag{35}
$$

Thus, we can get $(I + kS_k^{-1}) T_k ST_k (I + kS_k^{-1}) - T_k ST_k > 0$. Observing that $A > 0$, we have $D_2 = A((I + kS_k^{-1}) T_k ST_k (I + kS_k^{-1}) - T_k ST_k) A > 0$.

Applying Lemma 4, we can get $M(\widehat{\beta}_{\mathrm{WMAURE}}(w, k), \widehat{\beta}_{\mathrm{WMRE}}(w, k)) < 0$ if and only if $b_1^T(\sigma^2 D_2 + b_3 b_3^T)^{-1} b_1 > 1$. This completes the proof. □

4.3. MSEM Comparison of the AURE and WMAURE.

Now, we compare MSEM values between the WMAURE and AURE. Firstly, from (30) and (31), the difference of MSEM values between the AURE and WMAURE can be obtained by

$$
\begin{aligned}
&M\left(\widehat{\beta}_{\mathrm{AURE}}(k), \widehat{\beta}_{\mathrm{WMAURE}}(w, k)\right)\\
&\quad = \mathrm{MSEM}\left(\widehat{\beta}_{\mathrm{AURE}}(k)\right) - \mathrm{MSEM}\left(\widehat{\beta}_{\mathrm{WMAURE}}(w, k)\right)\\
&\quad = \sigma^2\left(I + kS_k^{-1}\right) T_k ST_k \left(I + kS_k^{-1}\right)\\
&\qquad + b_2 b_2^T - b_3 b_3^T\\
&\qquad - \sigma^2 A\left(\left(I + kS_k^{-1}\right) T_k ST_k \left(I + kS_k^{-1}\right) + w^2 R^T W^{-1} R\right) A\\
&\quad = \sigma^2 D_3 + b_2 b_2^T - b_3 b_3^T,
\end{aligned}
\tag{36}
$$

where $D_3 = (I + kS_k^{-1})T_kST_k(I + kS_k^{-1}) - A((I + kS_k^{-1})T_kST_k(I + kS_k^{-1}) + w^2R^TW^{-1}R)A$.

Theorem 8. *When* $\lambda_1(A((I + kS_k^{-1})T_kST_k(I + kS_k^{-1}) + w^2R^TW^{-1}R)A((I + kS_k^{-1})T_kST_k(I + kS_k^{-1}))^{-1}) < 1$, *the WMAURE is superior to the AURE in the MSEM sense, namely,* $M(\widehat{\beta}_{\mathrm{AURE}}(k), \widehat{\beta}_{\mathrm{WMAURE}}(w,k)) \geq 0$ *if and only if* $b_3^T(\sigma^2 D_3 + b_2 b_2^T)^{-1}b_3 \leq 1$.

Proof. Note that $(I + kS_k^{-1})T_kST_k(I + kS_k^{-1}) > 0$ and $A((I + kS_k^{-1})T_kST_k(I + kS_k^{-1}) + w^2R^TW^{-1}R)A > 0$. When $\lambda_1(A((I + kS_k^{-1})T_kST_k(I + kS_k^{-1}) + w^2R^TW^{-1}R)A((I + kS_k^{-1})T_kST_k(I + kS_k^{-1}))^{-1}) < 1$, we can get that $D_3 > 0$ by applying Lemma 5. Thus, from (36) and applying Lemma 4, we have $M(\widehat{\beta}_{\mathrm{AURE}}(k), \widehat{\beta}_{\mathrm{WMAURE}}(w,k)) \geq 0$ if and only if $b_3^T(\sigma^2 D_3 + b_2 b_2^T)^{-1}b_3 \leq 1$.

The proof is completed. $\qquad\square$

4.4. MSEM Comparison of the LS and WMAURE. Finally, we compare MSEM values between the LS and WMAURE. From (27) and (31), the difference of MSEM values between the LS and WMAURE can be computed by

$$
\begin{aligned}
&M\left(\widehat{\beta}_{\mathrm{LS}}, \widehat{\beta}_{\mathrm{WMAURE}}(w,k)\right) \\
&= \mathrm{MSEM}\left(\widehat{\beta}_{\mathrm{LS}}\right) - \mathrm{MSEM}\left(\widehat{\beta}_{\mathrm{WMAURE}}(w,k)\right) \\
&= \sigma^2 S^{-1} - \sigma^2 A\left(\left(I + kS_k^{-1}\right)T_kST_k\left(I + kS_k^{-1}\right)\right. \\
&\qquad\qquad \left. + w^2R^TW^{-1}R\right)A - b_3 b_3^T \\
&= \sigma^2 S^{-1} - \sigma^2 A + \sigma^2 A - \mathrm{MSEM}\left(\widehat{\beta}_{\mathrm{WME}}(w)\right) \\
&\quad + \mathrm{MSEM}\left(\widehat{\beta}_{\mathrm{WME}}(w)\right) - \mathrm{MSEM}\left(\widehat{\beta}_{\mathrm{WMAURE}}(w,k)\right) \\
&= \left(\sigma^2 S^{-1} - \sigma^2 A\right) + \left(\sigma^2 A - \sigma^2 A\left(S + w^2R^TW^{-1}R\right)A\right) \\
&\quad + \sigma^2 D_1 - b_3 b_3^T,
\end{aligned}
\tag{37}
$$

where $A = (S + wR^TW^{-1}R)^{-1}$ and $D_1 = A(S - (I + kS_k^{-1})T_kST_k(I + kS_k^{-1}))A > 0$ according to Section 4.1.

Firstly, using (9), we can compute that

$$
\sigma^2 S^{-1} - \sigma^2 A = \sigma^2 w S^{-1} R^T \left(W + wRS^{-1}R^T\right)^{-1} RS^{-1} > 0.
\tag{38}
$$

Moreover, it can be computed that

$$
\begin{aligned}
&\sigma^2 A - \sigma^2 A\left(S + w^2R^TW^{-1}R\right)A \\
&\quad = \sigma^2 AA^{-1}A - \sigma^2 A\left(S + w^2R^TW^{-1}R\right)A \\
&\quad = \sigma^2 A\left(S + wR^TW^{-1}R - S - w^2R^TW^{-1}R\right)A \\
&\quad = w(1-w)\sigma^2 A\left(R^TW^{-1}R\right)A > 0.
\end{aligned}
\tag{39}
$$

Therefore, $\sigma^2 S^{-1} - \sigma^2 A((I + kS_k^{-1})T_kST_k(I + kS_k^{-1}) + w^2R^TW^{-1}R)A > 0$. Applying Lemma 3, we can get that $M(\widehat{\beta}_{\mathrm{LS}}, \widehat{\beta}_{\mathrm{WMAURE}}(w,k)) \geq 0$ if and only if $b_3^T(\sigma^2 S^{-1} - \sigma^2 A((I + kS_k^{-1})T_kST_k(I + kS_k^{-1}) + w^2R^TW^{-1}R)A)b_3 \leq 0$.

Based on the above analysis, we can state the following theorem.

Theorem 9. *The WMAURE is superior to the LS according to the MSEM criterion, namely,* $M(\widehat{\beta}_{\mathrm{LS}}, \widehat{\beta}_{\mathrm{WMAURE}}(w,k)) \geq 0$ *if and only if* $b_3^T(\sigma^2 S^{-1} - \sigma^2 A((I + kS_k^{-1})T_kST_k(I + kS_k^{-1}) + w^2R^TW^{-1}R)A)b_3 \leq 0$.

5. Selection of Parameters k and w

In this section, we give a method about how to choose parameters k and w. Firstly, a linear regression model can be transformed to a canonical form by the orthogonal transformation. Let Q be an orthogonal matrix such that $Q^TX^TXQ = \Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_p)$, where λ_i is the eigenvalue of X^TX, and $\widetilde{X} = XQ$, $\widetilde{\beta} = Q^T\beta = (\widetilde{\beta}_1, \ldots, \widetilde{\beta}_p)^T$. Then, we get a canonical form of model (1) as

$$
y = \widetilde{X}\widetilde{\beta} + \varepsilon.
\tag{40}
$$

Note that $\widehat{\widetilde{\beta}}_{\mathrm{WMAURE}}(w,k) = Q^T\widehat{\beta}_{\mathrm{WMAURE}}(w,k)$ and $\mathrm{MSEM}(\widehat{\widetilde{\beta}}_{\mathrm{WMAURE}}(w,k)) = Q^T\mathrm{MSEM}(\widehat{\beta}_{\mathrm{WMAURE}}(w,k))Q$. It is supposed that S and $R^TW^{-1}R$ are commutative, then we have

$$
\begin{aligned}
&\mathrm{MSEM}\left(\widehat{\widetilde{\beta}}_{\mathrm{WMAURE}}(w,k)\right) \\
&= \sigma^2(\Lambda + w\Psi)^{-1}\left[\left(I + k(\Lambda + kI)^{-1}\right)(\Lambda + kI)^{-1}\right. \\
&\qquad\qquad \left. \times \Lambda^3(\Lambda + kI)^{-1}\left(I + k(\Lambda + kI)^{-1}\right) + w^2\Psi\right] \\
&\quad \times (\Lambda + w\Psi)^{-1} + k^4(\Lambda + w\Psi)^{-1}(\Lambda + kI)^{-2} \\
&\quad \times \Lambda\widetilde{\beta}\widetilde{\beta}^T\Lambda(\Lambda + kI)^{-2}(\Lambda + w\Psi)^{-1},
\end{aligned}
\tag{41}
$$

where $Q^TR^TW^{-1}RQ = \Psi = \mathrm{diag}(\xi_1, \ldots, \xi_p)$.

Optimal values for k and w can be derived by minimizing

$$
\begin{aligned}
g(w,k) &= \mathrm{tr}\left(\mathrm{MSEM}\left(\widehat{\widetilde{\beta}}_{\mathrm{WMAURE}}(w,k)\right)\right) \\
&= \sum_{i=1}^{p} \frac{\sigma^2\left[\lambda_i^3(\lambda_i + 2k)^2 + w\xi_i^2(\lambda_i + k)^4\right] + k^4\lambda_i^2\widetilde{\beta}_i^2}{(\lambda_i + w\xi_i)^2(\lambda_i + k)^4}.
\end{aligned}
\tag{42}
$$

For a fixed value of w, differentiating $g(w, k)$ with respect to k leads to

$$
\begin{aligned}
&\frac{\partial g(w, k)}{\partial k} \\
&= \sum_{i=1}^{p} \left(\left(\left[\sigma^2 \left[4 (\lambda_i + 2k) \lambda_i^3 + 4w \xi_i^2 (\lambda_i + k)^3 \right] + 4k^3 \lambda_i^2 \widetilde{\beta}_i^2 \right] \right. \right. \\
&\qquad \times (\lambda_i + k) - 4 \left[\sigma^2 \left[(\lambda_i + 2k)^2 \lambda_i^3 + w \xi_i^2 (\lambda_i + k)^4 \right] \right. \\
&\qquad \left. \left. + k^4 \lambda_i^2 \widetilde{\beta}_i^2 \right] \right) \left((\lambda_i + w \xi_i)^2 (\lambda_i + k)^5 \right)^{-1} \right) \\
&= \sum_{i=1}^{p} \frac{4 k \lambda_i^3 \left(k^2 \widetilde{\beta}_i^2 - 2 \sigma^2 k - \sigma^2 \lambda_i \right)}{(\lambda_i + w \xi_i)^2 (\lambda_i + k)^5},
\end{aligned}
\tag{43}
$$

and equating it to zero. Note that $k > 0$ and after unknown parameters σ^2 and $\widetilde{\beta}$ are replaced by their unbiased estimators, we obtain the optimal estimator of k for a fixed w value as

$$
\widehat{k} = \frac{\widehat{\sigma}^2 \sum_{i=1}^{p} \lambda_i^3 + \sqrt{\widehat{\sigma}^4 \left(\sum_{i=1}^{p} \lambda_i^3 \right)^2 + \widehat{\sigma}^2 \left(\sum_{i=1}^{p} \lambda_i^4 \right) \left(\sum_{i=1}^{p} \lambda_i^3 \widehat{\widetilde{\beta}}_i^2 \right)}}{\sum_{i=1}^{p} \lambda_i^3 \widehat{\widetilde{\beta}}_i^2}.
\tag{44}
$$

The w value which minimizes the function $g(w, k)$ can be found by differentiating $g(w, k)$ with respect to w when k is fixed

$$
\begin{aligned}
&\frac{\partial g(w, k)}{\partial w} \\
&= \sum_{i=1}^{p} \left(\left\{ \sigma^2 \xi_i^2 (\lambda_i + k)^4 (\lambda_i + w \xi_i) \right. \right. \\
&\qquad \left. - 2w \left[\sigma^2 \lambda_i^3 (\lambda_i + 2k)^2 + w \sigma^2 \xi_i^2 (\lambda_i + k)^4 + k^4 \lambda_i^2 \widetilde{\beta}_i^2 \right] \right\} \\
&\qquad \times \left((\lambda_i + w \xi_i)^3 (\lambda_i + k)^4 \right)^{-1} \right) \\
&= \sum_{i=1}^{p} \left(\left\{ -2w^2 \sigma^2 \xi_i^2 (\lambda_i + k)^4 + \sigma^2 \lambda_i \xi_i^2 (\lambda_i + k)^4 \right. \right. \\
&\qquad \left. + w \left[\sigma^2 \xi_i^3 (\lambda_i + k)^4 - 2 \sigma^2 \lambda_i^3 (\lambda_i + 2k)^2 - 2k^4 \lambda_i^2 \widetilde{\beta}_i^2 \right] \right\} \\
&\qquad \times \left((\lambda_i + w \xi_i)^3 (\lambda_i + k)^4 \right)^{-1} \right)
\end{aligned}
\tag{45}
$$

and equating it to zero. After unknown parameters σ^2 and $\widetilde{\beta}$ are replaced by their unbiased estimators, we get the optimal estimator of w for a fixed k value as

$$
\widehat{w} = \frac{\sqrt{\sum_{i=1}^{p} \widehat{h}_{2i}^2 + 8 \left(\sum_{i=1}^{p} \widehat{h}_{1i} \right)^2} - \sum_{i=1}^{p} \widehat{h}_{2i}}{\sum_{i=1}^{p} 4 \widehat{h}_{1i}},
\tag{46}
$$

where $\widehat{h}_{1i} = \widehat{\sigma}^2 \xi_i^2 (\lambda_i + k)^4$ and $\widehat{h}_{2i} = 2 \widehat{\sigma}^2 \lambda_i^3 (\lambda_i + 2k)^2 + 2k^4 \lambda_i^2 \widehat{\widetilde{\beta}}_i^2 - \widehat{\sigma}^2 \xi_i^3 (\lambda_i + k)^4$.

Table 1: Estimated QB values of the WMRE, AURE, and WMAURE.

| | $w = 0.1$ | | | $w = 0.4$ | | |
	$k = 0.01$	$k = 0.08$	$k = 0.1$	$k = 0.01$	$k = 0.08$	$k = 0.1$
WMRE	16.567	20.217	20.336	1.160	1.413	1.421
AURE	2456.43	3662.01	3706.28	2456.43	3662.01	3706.28
WMAURE	13.166	19.627	19.864	0.922	1.375	1.391
	$w = 0.8$			$w = 1$		
	$k = 0.01$	$k = 0.08$	$k = 0.1$	$k = 0.01$	$k = 0.08$	$k = 0.1$
WMRE	0.296	0.360	0.362	0.190	0.231	0.232
AURE	2456.43	3662.01	3706.28	2456.43	3662.01	3706.28
WMAURE	0.235	0.351	0.355	0.151	0.225	0.228

Table 2: Estimated MSE values of the WME, WMRE, AURE, and WMAURE.

| | $w = 0.1$ | | | $w = 0.4$ | | |
	$k = 0.01$	$k = 0.08$	$k = 0.1$	$k = 0.01$	$k = 0.08$	$k = 0.1$
LS	4912.13	4912.13	4912.13	4912.13	4912.13	4912.13
WME	59.771	59.771	59.771	39.278	39.278	39.278
WMRE	50.321	53.666	53.783	38.616	38.848	38.855
AURE	2663.66	3666.43	3709.16	2663.66	3666.43	3709.16
WMAURE	47.718	53.094	53.323	38.434	38.810	38.826
	$w = 0.8$			$w = 1$		
	$k = 0.01$	$k = 0.08$	$k = 0.1$	$k = 0.01$	$k = 0.08$	$k = 0.1$
LS	4912.13	4912.13	4912.13	4912.13	4912.13	4912.13
WME	38.639	38.639	38.639	38.620	38.620	38.620
WMRE	38.470	38.528	38.530	38.620	38.549	38.550
AURE	2663.66	3666.43	3709.16	2663.66	3666.43	3709.16
WMAURE	38.423	38.519	38.524	38.482	38.543	38.546

6. Numerical Example and Monte Carlo Simulation

In order to verify our theoretical results, we firstly conduct an experiment based on a real data set originally due to Woods et al. [21]. In this experiment, we replace the unknown parameters β and σ^2 by their unbiased estimators, which is according to the way in [17]. The result here and below is performed with R 2.14.1.

We can easily obtain that the condition number is about 3.66×10^7. This information indicates a serious multicollinearity among the regression vector. The ordinary least squares estimator of β is

$$
\widehat{\beta}_{\text{LS}} = S^{-1} X^T y = (62.4054, 1.5511, 0.5102, 0.1019, -0.1441)^T
\tag{47}
$$

with $\widehat{\sigma}_{\text{LS}}^2 = 5.983$.

Consider the following stochastic linear restrictions used in [17]:

$$
\begin{aligned}
r &= R\beta + e, \quad R = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & 1 & 1 \end{pmatrix}, \\
r &= \begin{pmatrix} 63.9498 \\ 2.5648 \end{pmatrix}, \quad e \sim \left(0, \widehat{\sigma}_{\text{LS}}^2 I_2 \right).
\end{aligned}
\tag{48}
$$

For the WMRE, AURE, and WMAURE, their quadratic bias values are given in Table 1 and their estimated MSE values are obtained in Table 2 by replacing all unknown parameters

Table 3: Estimated MSE values of the WME, WMRE, AURE, and WMAURE when $\rho = 0.8$.

| | $w = 0.1$ | | | $w = 0.4$ | | |
	$k = 0.01$	$k = 0.08$	$k = 0.1$	$k = 0.01$	$k = 0.08$	$k = 0.1$
LS	0.0898720	0.0898720	0.0898720	0.0898720	0.0898720	0.0898720
WME	0.0889010	0.0889010	0.0889010	0.0867887	0.0867887	0.0867887
WMRE	0.0889032	0.0902973	0.0911375	0.0867896	0.088059	0.0888259
AURE	0.0898720	0.0898718	0.0898717	0.0898720	0.0898718	0.0898717
WMAURE	0.0889010	0.0889008	0.0889007	0.0867887	0.0867885	0.0867884
	$w = 0.8$			$w = 1$		
	$k = 0.01$	$k = 0.08$	$k = 0.1$	$k = 0.01$	$k = 0.08$	$k = 0.1$
LS	0.0898720	0.0898720	0.0898720	0.0898720	0.0898720	0.0898720
WME	0.0854822	0.0854822	0.0854822	0.0853378	0.0853378	0.0853378
WMRE	0.0854817	0.0866088	0.0872931	0.0853367	0.0864015	0.0870495
AURE	0.0898720	0.0898718	0.0898717	0.0898720	0.0898718	0.0898717
WMAURE	0.0854822	0.0854821	0.0854820	0.0853378	0.0853376	0.0853375

Table 4: Estimated MSE values of the WME, WMRE, AURE, and WMAURE when $\rho = 0.9$.

| | $w = 0.1$ | | | $w = 0.4$ | | |
	$k = 0.01$	$k = 0.08$	$k = 0.1$	$k = 0.01$	$k = 0.08$	$k = 0.1$
LS	0.0987050	0.0987050	0.0987050	0.0987050	0.0987050	0.0987050
WME	0.0975363	0.0975363	0.0975363	0.0950143	0.0950143	0.0950143
WMRE	0.0975389	0.0992148	0.1002248	0.0950152	0.0965273	0.0974417
AURE	0.0987050	0.0987047	0.0987046	0.0987050	0.0987047	0.0987398
WMAURE	0.0975363	0.0975361	0.0975359	0.0950143	0.095014	0.0950139
	$w = 0.8$			$w = 1$		
	$k = 0.01$	$k = 0.08$	$k = 0.1$	$k = 0.01$	$k = 0.08$	$k = 0.1$
LS	0.0987050	0.0987050	0.0987050	0.0987050	0.0987050	0.0987050
WME	0.0934738	0.0934738	0.0934738	0.0933053	0.0933053	0.0933053
WMRE	0.0934729	0.0948017	0.0956092	0.0933036	0.0945530	0.0953142
AURE	0.0987050	0.0987047	0.0987398	0.0987050	0.0987047	0.0987398
WMAURE	0.0934738	0.0934736	0.0934735	0.0933053	0.0933050	0.0933049

in the corresponding theoretical MSE expressions by their least squares estimators.

It can be seen from Table 1 that the WMAURE has smaller quadratic bias values than the WMRE and AURE for every case, which agrees with our theoretical finding in Section 3. From Table 2, we can get that MSE values of our proposed estimator are the smallest among the LS, WME, WMRE, AURE, and WMAURE when w is fixed, which agrees with our theoretical finding in Theorems 6–9.

To further illustrate the behavior of our proposed estimator, we are to perform a Monte Carlo simulation study under different levels of multicollinearity. Following the way in [22, 23], we can get explanatory variables by the following equations:

$$x_{ij} = \left(1 - \rho^2\right)^{1/2} z_{ij} + \rho z_{i,p+1},$$

$$i = 1, 2, \ldots, n, \quad j = 1, 2, \ldots, p, \tag{49}$$

where z_{ij} is an independent standard normal pseudorandom number, and ρ is specified so that the theoretical correlation

between any two explanatory variables is given by ρ^2. A dependent variable is generated by

$$y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \varepsilon_i, \quad i = 1, 2, \ldots, n, \tag{50}$$

where ε_i is a normal pseudo-random number with mean zero and variance σ_i^2. In this study, we choose $(\beta_1, \beta_2, \beta_3, \beta_4)^T = (40, 1, 2, 3)^T$, $n = 60$, $p = 4$, $\sigma_i^2 = 1$, and the stochastic restriction $r = R\beta + e$, $R = \left(\begin{smallmatrix} 4 & 0 & -3 & 1 \\ 2 & 1 & 2 & 0 \end{smallmatrix}\right)$, $e \sim N(0, 0.1 I_2)$. Furthermore, we discuss three cases when $\rho = 0.8, 0.9, 0.99$.

For three different levels of multicollinearity, MSE values of LS, WME, AURE, WMRE, and WMAURE are obtained in Tables 3, 4, and 5, respectively. From Tables 3–5, we can derive the following results.

(1) With the increase of multicollinearity, MSE values of the LS, WME, WMRE, AURE, and WMAURE are increasing. And for all cases, the WMAURE has smaller estimated MSE values than the LS, AURE, and WME.

TABLE 5: Estimated MSE values of the WME, WMRE, AURE, and WMAURE when $\rho = 0.99$.

| | $w = 0.1$ | | | $w = 0.4$ | | |
	$k = 0.01$	$k = 0.08$	$k = 0.1$	$k = 0.01$	$k = 0.08$	$k = 0.1$
LS	0.1376903	0.1376903	0.1376903	0.1376903	0.1376903	0.1376903
WME	0.1354379	0.1354379	0.1354379	0.1307427	0.1307427	0.1307427
WMRE	0.1354425	0.138655	0.1405906	0.130743	0.1335329	0.1352221
AURE	0.1376903	0.1376896	0.1376892	0.1376903	0.1376896	0.1376892
WMAURE	0.1354379	0.1354372	0.1354368	0.1307426	0.130742	0.1307416
	$w = 0.8$			$w = 1$		
	$k = 0.01$	$k = 0.08$	$k = 0.1$	$k = 0.01$	$k = 0.08$	$k = 0.1$
LS	0.1376903	0.1376903	0.1376903	0.1376903	0.1376903	0.1376903
WME	0.1280238	0.1280238	0.1280238	0.127739	0.127739	0.127739
WMRE	0.1280196	0.1303649	0.1317949	0.127733	0.1298949	0.1312179
AURE	0.1376903	0.1376896	0.1376892	0.1376903	0.1376896	0.1376892
WMAURE	0.1280238	0.1280231	0.1280228	0.127739	0.1277384	0.127738

(2) The value of w is the level of the weight to the sample information and the prior information; we can see from three tables that estimated MSE values of the WME, WMRE, and WMAURE become more and more smaller when the value of w increases. It can be concluded that we get more exact estimator of the parameter with more depended prior information.

7. Conclusions

In this paper, we propose the WMAURE based on the WME [16] and the AURE [9] and discuss some properties of the new estimator over the relative estimators. In particular, we prove that the WMAURE has smaller quadratic bias than the AURE and WMRE and derive that the proposed estimator is superior to the LS, WME, WMRE, and AURE in the mean squared error matrix sense under certain conditions. The optimal values of parameters k and w are obtained. Furthermore, we perform a real data example and a Monte Carlo study to support the finding of our theoretical results.

Acknowledgments

References

[1] W. F. Massy, "Principal components regression in exploratory statistical research," *Journal of the American Statistical Association*, vol. 60, no. 309, pp. 234–266, 1965.

[2] A. E. Hoerl and R. W. Kennard, "Ridge regression: biased estimation for non-orthogonal problem," *Technometrics*, vol. 12, no. 1, pp. 55–67, 1970.

[3] M. R. Baye and D. F. Parker, "Combining ridge and principal component regression: a money demand illustration," *Communications in Statistics A*, vol. 13, no. 2, pp. 197–205, 1984.

[4] K.-J. Liu, "A new class of biased estimate in linear regression," *Communications in Statistics A*, vol. 22, no. 2, pp. 393–402, 1993.

[5] S. Kaçıranlar and S. Sakallıoğlu, "Combining the Liu estimator and the principal component regression estimator," *Communications in Statistics A*, vol. 30, no. 12, pp. 2699–2705, 2001.

[6] K. Kadiyala, "A class of almost unbiased and efficient estimators of regression coefficients," *Economics Letters*, vol. 16, no. 3-4, pp. 293–296, 1984.

[7] B. Singh, Y. P. Chaubey, and T. D. Dwivedi, "An almost unbiased ridge estimator," *Sankhyā B*, vol. 48, no. 3, pp. 342–346, 1986.

[8] F. Akdeniz and S. Kaçiranlar, "On the almost unbiased generalized Liu estimator and unbiased estimation of the bias and MSE," *Communications in Statistics A*, vol. 24, no. 7, pp. 1789–1797, 1995.

[9] F. Akdeniz and H. Erol, "Mean squared error matrix comparisons of some biased estimators in linear regression," *Communications in Statistics A*, vol. 32, no. 12, pp. 2389–2413, 2003.

[10] J. Durbin, "A note on regression when there is extraneous information about one of the coefficients," *Journal of the American Statistical Association*, vol. 48, no. 264, pp. 799–808, 1953.

[11] H. Theil and A. S. Goldberger, "On pure and mixed statistical estimation in economics," *International Economic Review*, vol. 2, no. 1, pp. 65–78, 1961.

[12] H. Theil, "On the use of incomplete prior information in regression analysis," *Journal of the American Statistical Association*, vol. 58, pp. 401–414, 1963.

[13] M. H. Hubert and P. Wijekoon, "Improvement of the Liu estimator in linear regression model," *Statistical Papers*, vol. 47, no. 3, pp. 471–479, 2006.

[14] H. Yang and J. Xu, "An alternative stochastic restricted Liu estimator in linear regression," *Statistical Papers*, vol. 50, no. 3, pp. 639–647, 2009.

[15] Y.-L. Li and H. Yang, "A new stochastic mixed ridge estimator in linear regression model," *Statistical Papers*, vol. 51, no. 2, pp. 315–323, 2010.

[16] B. Schaffrin and H. Toutenburg, "Weighted mixed regression," *Zeitschrift für Angewandte Mathematik und Mechanik*, vol. 70, pp. 735–738, 1990.

[17] Y.-L. Li and H. Yang, "A new ridge-type estimator in stochastic restricted linear regression," *Statistics*, vol. 45, no. 2, pp. 123–130, 2011.

[18] C. R. Rao, H. Toutenburg, and S. C. Heumann, *Linear Models and Generalizations: Least Squares and Alternatives*, Springer, New York, NY, USA, 2008.

[19] R. W. Farebrother, "Further results on the mean square error of ridge regression," *Journal of the Royal Statistical Society B*, vol. 38, no. 3, pp. 248–250, 1976.

[20] G. Trenkler and H. Toutenburg, "Mean squared error matrix comparisons between biased estimators-an overview of recent results," *Statistical Papers*, vol. 31, no. 3, pp. 165–179, 1990.

[21] H. Woods, H. H. Steinour, and H. R. Starke, "Effect of composition of Portland cement on heat evolved during hardening," *Industrial and Engineering Chemistry*, vol. 24, no. 11, pp. 1207–1214, 1932.

[22] K.-J. Liu, "Using Liu-type estimator to combat collinearity," *Communications in Statistics A*, vol. 32, no. 5, pp. 1009–1020, 2003.

[23] J.-W. Xu and H. Yang, "More on the bias and variance comparisons of the restricted almost unbiased estimators," *Communications in Statistics A*, vol. 40, no. 22, pp. 4053–4064, 2011.