

New estimators in a partial linear model depending on an unbiased ridge regression estimator

Yousif A. Al-Khazraji^{1*}, Mustafa I. Alheety²

¹Department of Mathematics, College of Education for Pure Sciences, University of Anbar, Anbar, Iraq.

²Department of Mathematics, College of Education for Pure Sciences, University of Anbar, Anbar, Iraq.

Abstract. This paper introduces two new estimators based on the philosophy of unbiased ridge regression estimation, where the parameters are part of a partial linear model suffering from multicollinearity. These proposed estimators are called the Difference-Based Unbiased Ridge Estimator $\hat{\beta}_{DB-URR}$ and the Difference-Based Modified Unbiased Ridge Estimator $\hat{\beta}_{DB-MUR}$ for the regression parameters β . The Mean Squared Error Matrix (MSEM) criterion is employed to compare the proposed estimators against the Difference-Based Ordinary Least Squares estimator $\hat{\beta}_{DB-OLS}$ and the Difference-Based Ordinary Ridge Estimator $\hat{\beta}_{DB-ORR}$. Finally, the performance of the new estimators is evaluated through a comprehensive simulation study and a numerical example.

Keywords: Partial Linear Model; Multicollinearity; Ridge Regression Estimator; Unbiased Ridge Regression Estimator; Difference-Based Estimator.

1. Introduction

In modern statistical and econometric analysis, the partial linear model has gained significant importance due to its ability to balance parametric precision with non-parametric flexibility. This model takes the general form:

$$y_i = \hat{x}_i\beta + f(t_i) + \epsilon_i, \quad i = 1, \dots, n \quad (1)$$

where y_i represents the i -th observation of the response variable, $\mathbf{x}_i = (x_{i1}, \dots, x_{ip})'$ is a known p -dimensional vector of explanatory variables with $p \leq n$, $\beta = (\beta_1, \dots, \beta_p)'$ is a vector of unknown parameters, and $f(\cdot)$ is an unknown smooth function. The non-parametric variables t_i are arranged and distributed so that $t_1 \leq t_2 \leq \dots, t_n$. The terms ϵ_i are independent random errors that follow a normal distribution with mean zero and variance σ^2 , i.e., $\epsilon_i \sim N(0, \sigma^2)$.

* Corresponding author: you24u2005@uoanbar.edu.iq

We rewrite the model in equation (1) in matrix form as follows:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{f}(\mathbf{t}) + \boldsymbol{\epsilon} \quad (2)$$

where $\mathbf{y} = (y_1, \dots, y_n)'$ is an $n \times 1$ vector of responses, $\mathbf{X} = (x_1, \dots, x_n)$ is an $n \times p$ matrix, $\mathbf{f}(\mathbf{t}) = [f(t_1), \dots, f(t_2)]'$ is an $n \times 1$ vector of the unknown function values, and $\boldsymbol{\epsilon} = (\epsilon_1, \dots, \epsilon_n)'$ is an $n \times 1$ vector of random errors. The partial linear model (2) comprises a linear component, $\mathbf{X}\boldsymbol{\beta}$, that is easily comprehended, and a nonlinear component, $\mathbf{f}(\mathbf{t})$, which reflects complicated interactions that are difficult to predict. This mixed character makes the model an effective tool for describing events that do not follow strictly linear assumptions. To accurately estimate the parametric coefficients $\boldsymbol{\beta}$ in the model, the effect of the non-parametric function $\mathbf{f}(\mathbf{t})$ must be eliminated. This is accomplished using the difference method suggested by Yatchew (1997, 2000)[1,2], which is an excellent tool for accomplishing this task. This approach works by sorting the data according to the non-parametric variable t and then calculating weighted differences between adjacent observations. This procedure essentially cancels out the term $\mathbf{f}(\mathbf{t})$, assuming that the function is locally linear (or nearly constant) over small intervals of t , converting the model into an approximate linear model that may be estimated using traditional techniques. Despite this transformation, the design matrix \mathbf{X} associated with the parametric component is still subject to the multicollinearity problem among its independent variables. Multicollinearity is defined as the presence of nearly perfect linear dependence between the column vectors of the design matrix \mathbf{X} in a linear model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$. Here, \mathbf{y} is an $n \times 1$ vector of observations, \mathbf{X} is an $n \times p$ matrix of independent variables assumed to have full column rank p , $\boldsymbol{\beta}$ is a $p \times 1$ vector of unknown parameters, and $\boldsymbol{\epsilon}$ is an error vector such that $E(\boldsymbol{\epsilon}) = 0, E(\boldsymbol{\epsilon}\boldsymbol{\epsilon}') = \sigma^2\mathbf{I}_n$ (where \mathbf{I}_n is the identity matrix of order n). Multicollinearity can result in large confidence intervals for individual parameters and erroneous signs for coefficients. The condition number of the matrix \mathbf{X} indicates the presence of multicollinearity, but it does not show the structure of the linear dependence between its column vectors x_1, x_2, \dots, x_p . Analyzing the eigenvalues of the matrix $\mathbf{X}'\mathbf{X}$ is the most effective method for identifying multicollinearity. This paper presents two new estimators, the Difference-Based Unbiased Ridge Estimator (DB-URR) and the Difference-Based Modified Unbiased Ridge Estimator (DB-MUR), and explores their application within the framework of the partial linear model after using the difference transformation. The remainder of the paper is organized as follows: Section 2 presents the related works. Section 3 defines the partial linear model and the differencing estimator and provides comprehensive mathematical formulae for the estimators used in the comparison: DB-ORR, DB-URR, and DB-MUR. Section 4 examines the theoretical performance of these estimators based on the Mean Squared Error (MSE) matrix criterion. Section 5 presents the simulation study and a numerical example using real-world data. Finally, Section 6 offers the key results and conclusions.

2. Related Works

Estimating the partial linear model often involves handling the non-parametric component effectively. For instance, Wang et al. (2007) [3] addressed this by using a unique class of difference sequences incorporating higher-order differences to enhance efficiency while estimating the linear component. However, their study primarily focused on efficiency and did not address the severe complications that arise when the explanatory variables are highly correlated. Indeed, multicollinearity causes the transformed design matrix $\tilde{\mathbf{X}}\mathbf{X}$ to be ill-conditioned, leading to an inflation in the variance of ordinary least squares (OLS) estimators and a significant rise in the mean squared error (MSE). To account for the significant variability of OLS, numerous biased estimators have been devised. The most well-known is the Ordinary Ridge Regression in Linear Model (ORR) estimator, developed

by Hoerl and Kennard (1970)[4], which decreases variance by adding a minor bias parameter (known as the Ridge parameter k) to the diagonal elements of the \mathbf{XX} matrix. In the context of the partial linear model, this Ridge logic was applied to the difference-transformed model; Tabakan and Akdeniz (2010)[5] proposed the difference-based Ridge estimator (DB-ORR) to decrease the huge variance in the difference-based least squares estimator (DB-OLS). However, a major limitation of the DB-ORR estimator is that it remains clearly biased, which can distort statistical inferences. To address the bias issue in standard linear models, Crouse et al. (1995)[6] proposed a unique theoretical framework, the Unbiased Ridge Estimator (URR), to bridge the gap caused by the Ridge estimator's explicit bias. This estimator tries to harness the variance-reduction capabilities of the Ridge estimator while retaining its unbiased nature by combining it with a priori information denoted by \mathbf{J} . Based on this concept, Batah 2009 [7] introduced the Modified Unbiased Ridge (MUR) estimator. MUR is a shrinkage estimator, derived from the URR in the same way that the (ORR) is derived from (OLS). The motivation behind this modification was that the URR, despite being unbiased, has a limitation: it still suffers from variance inflation under severe multicollinearity. Therefore, the MUR uses an additional shrinkage matrix \mathbf{W} to improve the estimator's properties in terms of MSE and further reduce variance, even if this introduces some bias. Despite these advancements in standard linear models, there is a distinct gap in the literature regarding the application of URR and MUR estimators within the context of partial linear models. Therefore, to overcome the limitations of the biased DB-ORR and the high variance of DB-OLS, this paper bridges this gap by proposing the DB-URR and DB-MUR estimators, combining the strengths of difference-based transformations with advanced ridge regression methodologies.

3. Methodology

3.1 The model and differencing-based estimator.

To estimate β in model (2), eliminate the influence of the non-parametric function $\mathbf{f}(\mathbf{t})$. To do this, we employ the differencing approach proposed by Yatchew (1997, 2000,2003)[1, 2, 8]. This estimating technique is known as difference-based estimation. We have the weight vector $\mathbf{d} = (d_0, \dots, d_m)'$ of degree $m + 1$, where m denotes the order of the differences and d_0, \dots, d_m represent the weights of the differences that lower the variance in linear estimators, which means minimizing:

$$\min_{d_0, \dots, d_m} \delta = \sum_{i=1}^m \left(\sum_{j=0}^{m-i} d_j d_{i+j} \right)^2 \tag{3}$$

Here, δ is called the difference series of order m^{th} , and the weights must satisfy the following two conditions.

$$\sum_{j=0}^m d_j = 0 \quad , \quad \sum_{j=0}^m d_j^2 = 1 \tag{4}$$

The difference matrix \mathbf{D} of dimension $(n - m) \times n$ whose elements satisfy conditions (4), is given by

$$\mathbf{D} = \begin{pmatrix} d_0 & d_1 & \dots & d_m & 0 & 0 & \dots & 0 \\ 0 & d_0 & d_1 & \dots & d_m & 0 & \dots & 0 \\ \dots & \dots & & & & & & \\ \dots & \dots & & & & & & \\ 0 & 0 & \dots & d_1 & \dots & d_m & 0 & 0 \\ 0 & 0 & \dots & d_0 & d_1 & \dots & d_m & 0 \\ 0 & 0 & \dots & 0 & d_0 & d_1 & \dots & d_m \end{pmatrix} \tag{5}$$

The difference matrix \mathbf{D} performs a linear composition of adjacent observations after sorting the data in ascending order according to the variable t , which leads to the approximate removal of the function $\mathbf{f}(t)$ when applying the matrix \mathbf{D} to both sides of the partial linear equation (2) as follows:

$$\mathbf{Dy} = \mathbf{DX}\boldsymbol{\beta} + \mathbf{Df}(t) + \mathbf{D}\boldsymbol{\epsilon} \tag{6}$$

Since $\mathbf{Df}(t) \approx \mathbf{0}$, we obtain the following transformed model

$$\mathbf{Dy} \cong \mathbf{D}\mathbf{X}\boldsymbol{\beta} + \mathbf{D}\boldsymbol{\epsilon} \tag{7}$$

We can rewrite equation (7) as:

$$\tilde{\mathbf{y}} \cong \tilde{\mathbf{X}}\boldsymbol{\beta} + \tilde{\boldsymbol{\epsilon}},$$

where $\tilde{\mathbf{y}} = \mathbf{Dy}$, $\tilde{\mathbf{X}} = \mathbf{DX}$, $\tilde{\boldsymbol{\epsilon}} = \mathbf{D}\boldsymbol{\epsilon}$.

So, we can see that $\tilde{\boldsymbol{\epsilon}}$ is a $n - m$ vector of distributed with a zero mean $E(\tilde{\boldsymbol{\epsilon}}) = 0$ and a variance $E(\tilde{\boldsymbol{\epsilon}}\tilde{\boldsymbol{\epsilon}}') = \sigma^2 \mathbf{DD}'$. This transformation linearizes the model, allowing for the estimation of $\boldsymbol{\beta}$ using standard linear regression methods, notably the Ordinary Least Squares (OLS) approach. Using optimum weights, the OLS technique can provide the differencing estimator for the partial linear model suggested by Yatchew (1997)[1] for estimating $\boldsymbol{\beta}$, which takes the following form:

$$\hat{\boldsymbol{\beta}}_{DB-OLS} = (\tilde{\mathbf{X}}'\tilde{\mathbf{X}})^{-1}\tilde{\mathbf{X}}'\tilde{\mathbf{y}} \tag{8}$$

Once $\boldsymbol{\beta}$ has been estimated, a non-parametric method can be used to estimate $\mathbf{f}(t)$. To calculate the variance, we use the estimator of σ^2 proposed by Eubank et al. (1998)[9] and Klippel, K., and Eubank, R. L. (2007) [10], as follows:

$$\hat{\sigma}^2 = \frac{\tilde{\mathbf{y}}'(\mathbf{I} - \mathbf{P})\tilde{\mathbf{y}}}{tr(\mathbf{D}'(\mathbf{I} - \mathbf{P})\mathbf{D})} \tag{9}$$

where $tr(\cdot)$ denotes the trace of a square matrix and \mathbf{P} is the projection matrix

$$\mathbf{P} = \tilde{\mathbf{X}}(\tilde{\mathbf{X}}'\tilde{\mathbf{X}})^{-1}\tilde{\mathbf{X}}' \tag{10}$$

The DB-OLS estimator has a very large variance in the presence of multicollinearity, which causes instability. To address this issue, Tabakan and Akdeniz (2010)[5] proposed the DB-ORR estimator, which is the most common biased estimator for dealing with multicollinearity, and its formula is given as:

$$\begin{aligned} \hat{\boldsymbol{\beta}}_{DB-ORR} &= (\tilde{\mathbf{X}}'\tilde{\mathbf{X}} + \mathbf{kI})^{-1}\tilde{\mathbf{X}}'\tilde{\mathbf{y}} \\ &= \mathbf{C}_k\tilde{\mathbf{X}}'\tilde{\mathbf{y}} \end{aligned} \tag{11}$$

where $\mathbf{C}_k = (\tilde{\mathbf{X}}'\tilde{\mathbf{X}} + \mathbf{kI})^{-1}$ and \mathbf{I}_p is the $\mathbf{p} \times \mathbf{p}$ identity matrix, and k is the ridge parameter, (also known as the bias coefficient). The condition $k \geq 0$, lowers variance by introducing a small amount of bias.

3.2 Difference-based unbiased ridge estimator and Modified Unbiased Ridge estimator

The proposed estimators use the conventional Ridge methodology combined with J , which represents prior knowledge about the location of the coefficient vector $\boldsymbol{\beta}$. The first suggested estimator, DB-URR, is a novel estimator that improves upon the performance of the standard ridge estimator DB-ORR in the partial linear model. This estimator was developed within a unique theoretical framework that uses the ridge parameter for minimizing variance while remaining unbiased. The formula is as follows:

$$\begin{aligned} \hat{\boldsymbol{\beta}}_{DB-URR} &= (\tilde{\mathbf{X}}'\tilde{\mathbf{X}} + \mathbf{kI})^{-1}(\tilde{\mathbf{X}}'\tilde{\mathbf{y}} + \mathbf{kJ}) \\ &= \mathbf{C}_k(\tilde{\mathbf{X}}'\tilde{\mathbf{y}} + \mathbf{kJ}) \end{aligned} \tag{12}$$

where $\mathbf{C}_k = (\tilde{\mathbf{X}}'\tilde{\mathbf{X}} + \mathbf{kI})^{-1}$ and \mathbf{J} is a random vector representing prior information that follows a normal distribution with a mean $\boldsymbol{\beta}$ and a specified variance matrix $\left(\frac{\sigma^2}{k}\right)\mathbf{I}_p$. The second suggested estimator, DB-MUR, is a further refinement of the DB-URR estimator obtained by pre-multiplying it with a weighting matrix \mathbf{W} . It is defined by the formula:

$$\hat{\boldsymbol{\beta}}_{DB-MUR} = \mathbf{W}\hat{\boldsymbol{\beta}}_{DB-URR} \tag{13}$$

$$= \mathbf{W}(\tilde{\mathbf{X}}'\tilde{\mathbf{X}} + \mathbf{kI})^{-1}(\tilde{\mathbf{X}}'\tilde{\mathbf{y}} + \mathbf{kJ})$$

where $\mathbf{W} = [\mathbf{I}_p - \mathbf{k}(\tilde{\mathbf{X}}'\tilde{\mathbf{X}} + \mathbf{kI})^{-1}]$ is a shrinkage matrix. It seeks to lower the high variance in the DB-URR estimator while adding a small amount of bias. This adjustment enhances the shrinkage process by striking an optimal balance between bias and variance, resulting in better estimator qualities in terms of MSE.

4. Theoretical comparison in terms of MSE

Let us suppose \mathbf{b}^* is an estimator for $\boldsymbol{\beta}$ in the linear model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$. The mean square error Matrix (MSEM) of the estimator \mathbf{b}^* is defined as follows.

$$\text{MSEM}(\mathbf{b}^*, \boldsymbol{\beta}) = \mathbf{E}[(\mathbf{b}^* - \boldsymbol{\beta})(\mathbf{b}^* - \boldsymbol{\beta})'] \tag{14}$$

If we designate the variance and covariance matrix of the estimator \mathbf{b}^* by the symbol $\text{Var}(\mathbf{b}^*)$, the equation (14) is equivalent to

$$\text{MSE}(\mathbf{b}^*, \boldsymbol{\beta}) = \mathbf{V}(\mathbf{b}^*) + (\text{bias}(\mathbf{b}^*))(\text{bias}(\mathbf{b}^*))'$$

where $\text{bias}(\mathbf{b}^*) = \mathbf{E}(\mathbf{b}^*) - \boldsymbol{\beta}$. The scalar Mean Squared Error (MSE) is provided by the relation. $\text{SMSE}(\mathbf{b}^*, \boldsymbol{\beta}) = \mathbf{E}[(\mathbf{b}^* - \boldsymbol{\beta})(\mathbf{b}^* - \boldsymbol{\beta})'] = \text{tr}(\text{MSEM}(\mathbf{b}^*, \boldsymbol{\beta}))$

To make a comparison between any two estimators, we may need some of the lemmas below:

Lemma 4.1 Let $\hat{\boldsymbol{\beta}}_1$ and $\hat{\boldsymbol{\beta}}_2$ be the two estimators of $\boldsymbol{\beta}$. The estimator $\hat{\boldsymbol{\beta}}_2$ is said to be superior to $\hat{\boldsymbol{\beta}}_1$ in the sense of the MSEM criterion if [15]

$$\text{MSEM}(\hat{\boldsymbol{\beta}}_1) - \text{MSEM}(\hat{\boldsymbol{\beta}}_2) \geq 0$$

(i.e., the difference is a positive semi-definite matrix).

Lemma 4.2 (Farebrother, 1976) [16] Let \mathbf{A} be a symmetric positive definite $m \times m$ matrix, \mathbf{a} be an $m \times 1$ vector, and c be a positive number. Then the matrix $c\mathbf{A} - \mathbf{a}\mathbf{a}'$ is a non-negative definite matrix if and only if $\mathbf{a}'\mathbf{A}^{-1}\mathbf{a} \leq c$.

4.1 Difference-based unbiased ridge estimator and Modified Unbiased Ridge estimator

Tabakan and Akdeniz (2010)[5] defined the properties of the DB-ORR estimator as follows:

$$\text{bias}(\hat{\boldsymbol{\beta}}_{DB-ORR}) = -k\mathbf{C}_k\boldsymbol{\beta} \tag{15}$$

and

$$\text{Var}(\hat{\boldsymbol{\beta}}_{DB-ORR}) = \sigma^2\mathbf{C}_k\mathbf{C}_k \tag{16}$$

where $\mathbf{C} = (\mathbf{D}'\tilde{\mathbf{X}})'(\mathbf{D}'\tilde{\mathbf{X}})$ and $\mathbf{C}_k = (\tilde{\mathbf{X}}'\tilde{\mathbf{X}} + \mathbf{kI})^{-1}$

To facilitate the SMSE comparison, we use the spectral decomposition (canonical form). Let \mathbf{T} be an orthogonal matrix such that $\mathbf{T}\mathbf{T}' = \mathbf{I}_p$, $\mathbf{T}\tilde{\mathbf{X}}'\tilde{\mathbf{X}}\mathbf{T}' = \boldsymbol{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_p)$, and $\mathbf{T}'\mathbf{C}\mathbf{T} = \boldsymbol{\Lambda} = \text{diag}(c_1, \dots, c_p)$, where λ_i are the eigenvalues of $\tilde{\mathbf{X}}'\tilde{\mathbf{X}}$. We also define $\boldsymbol{\alpha} = \mathbf{T}'\boldsymbol{\beta}$. The Scalar Mean Squared Error (SMSE) for DB-ORR can be expressed in the canonical form as:

$$\begin{aligned} \text{SMSE}(\hat{\boldsymbol{\beta}}_{DB-ORR}) &= \text{tr}(\text{Var}(\hat{\boldsymbol{\beta}}_{DB-ORR})) + \text{bias}(\hat{\boldsymbol{\beta}}_{DB-ORR})\text{bias}(\hat{\boldsymbol{\beta}}_{DB-ORR})' \\ &= \text{tr}(\sigma^2\mathbf{C}_k\mathbf{C}_k) + k^2\mathbf{C}_k\boldsymbol{\beta}\boldsymbol{\beta}'\mathbf{C}_k \\ &= \sigma^2 \sum_{i=1}^p \frac{c_i}{(\lambda_i + k)^2} + k^2 \sum_{i=1}^p \frac{\alpha_i^2}{(\lambda_i + k)^2} \end{aligned} \tag{17}$$

Now, we will find out the properties of the new estimator DB-URR

Using equation (12), we obtain

$$\mathbf{E}(\hat{\boldsymbol{\beta}}_{DB-URR}) = \boldsymbol{\beta} \tag{18}$$

It is an unbiased estimator, meaning the bias is zero; thus the estimated mean squared error matrix (MSEM) of the (DB-URR) is exactly equal to its variance matrix.

$$\text{MSEM}(\hat{\boldsymbol{\beta}}_{DB-URR}) = \text{Var}(\hat{\boldsymbol{\beta}}_{DB-URR}) = \sigma^2\mathbf{C}_k[\mathbf{C} + k\mathbf{I}]\mathbf{C}_k \tag{19}$$

The SMSE of the DB-URR estimator is exactly equal to the trace of its variance matrix, which can be written in the canonical form as:

$$SMSE(\hat{\beta}_{DB-URR}) = tr(\sigma^2 \mathbf{C}_k[\mathbf{C} + k\mathbf{I}]\mathbf{C}_k) = \sigma^2 \sum_{i=1}^p \frac{c_i + k}{(\lambda_i + k)^2} \tag{20}$$

Now, we investigate the difference between the SMSEs of the two estimators to establish superiority.

$$\begin{aligned} \Delta &= SMSE(\hat{\beta}_{DB-ORR}) - SMSE(\hat{\beta}_{DB-URR}) \\ &= \sum_{i=1}^p \frac{\sigma^2 c_i + k^2 \alpha_i^2 - \sigma^2 (c_i + k)}{(\lambda_i + k)^2} \\ &= \sum_{i=1}^p \frac{k}{(\lambda_i + k)^2} [k\alpha_i^2 - \sigma^2] \end{aligned} \tag{21}$$

For the proposed estimator $\hat{\beta}_{DB-URR}$ to be superior to $\hat{\beta}_{DB-ORR}$, the difference Δ must be non-negative (i.e., $\Delta \geq 0$). This yields the following theorem:

Theorem 4.1 The proposed estimator $\hat{\beta}_{DB-URR}$ is superior to the estimator $\hat{\beta}_{DB-ORR}$ in terms of the Scalar Mean Squared Error (SMSE) if and only if

$$k\alpha_i^2 > \sigma^2, \quad \forall i = 1, \dots, p \tag{22}$$

4.2 Comparing the $\hat{\beta}_{DB-MUR}$ estimator with the $\hat{\beta}_{DB-ORR}$ estimator

Now, we investigate the properties of the second proposed estimator $\hat{\beta}_{DB-MUR}$ Using equation (13), we obtain the estimator's expectation

$$E(\hat{\beta}_{DB-MUR}) = E(\mathbf{W} \hat{\beta}_{DB-URR}) = \mathbf{W}\beta \tag{23}$$

and its bias is :

$$bias(\hat{\beta}_{DB-MUR}) = -k\mathbf{C}_k\beta \tag{24}$$

We note that the estimated bias of $\hat{\beta}_{DB-MUR}$ is exactly equal to the estimated bias $\hat{\beta}_{DB-ORR}$ The variance matrix of the DB-MUR estimator is known

$$Var(\hat{\beta}_{DB-MUR}) = \sigma^2 \mathbf{W}\mathbf{C}_k[\mathbf{C} + k\mathbf{I}]\mathbf{C}_k\mathbf{W}' \tag{25}$$

Therefore, the mean squared error matrix (MSEM) is formulated as:

$$MSEM(\hat{\beta}_{DB-MUR}) = Var(\hat{\beta}_{DB-MUR}) + (bias(\hat{\beta}_{DB-MUR}))(bias(\hat{\beta}_{DB-MUR}))'$$

Let $\mathbf{b}_1 = bias(\hat{\beta}_{DB-MUR}) = -k\mathbf{C}_k\beta$

$$\begin{aligned} &= Var(\hat{\beta}_{DB-MUR}) + (\mathbf{b}_1)(\mathbf{b}_1)' \\ &= \sigma^2 \mathbf{W}\mathbf{C}_k[\mathbf{C} + k\mathbf{I}]\mathbf{C}_k\mathbf{W}' + (-k\mathbf{C}_k\beta)(-k\mathbf{C}_k\beta)' \end{aligned} \tag{26}$$

$$MSE(\hat{\beta}_{DB-MUR}) = \sigma^2 \mathbf{W}\mathbf{C}_k[\mathbf{C} + k\mathbf{I}]\mathbf{C}_k\mathbf{W}' + k^2 \mathbf{C}_k\beta\beta'\mathbf{C}_k \tag{26}$$

Now let's look at the difference between the Mean Squared Error (MSEM) matrix

$$\begin{aligned} \Delta &= MSEM(\hat{\beta}_{DB-ORR}) - MSEM(\hat{\beta}_{DB-MUR}) \\ &= [Var(\hat{\beta}_{DB-ORR}) + (\mathbf{b})(\mathbf{b})'] - [Var(\hat{\beta}_{DB-MUR}) + (\mathbf{b}_1)(\mathbf{b}_1)'] \\ &= [Var(\hat{\beta}_{DB-ORR}) - Var(\hat{\beta}_{DB-MUR})] + [(\mathbf{b})(\mathbf{b})' - (\mathbf{b}_1)(\mathbf{b}_1)'] \end{aligned}$$

Since $\mathbf{b} = \mathbf{b}_1 = -k\mathbf{C}_k\beta$, the squared bias matrices cancel each other out. Thus, the MSEM difference reduces exclusively to the difference in their variance matrices:

$$\begin{aligned} \Delta &= Var(\hat{\beta}_{DB-ORR}) - Var(\hat{\beta}_{DB-MUR}) \\ &= \sigma^2 \mathbf{C}_k\mathbf{C}\mathbf{C}_k - \sigma^2 \mathbf{W}\mathbf{C}_k[\mathbf{C} + k\mathbf{I}]\mathbf{C}_k\mathbf{W}' \end{aligned} \tag{27}$$

To facilitate comparison and verify the positivity definiteness condition, we use spectral decomposition to transform these matrices into their canonical form. Equation (27) can be written using the canonical form as.

$$diag \left\{ \frac{\sigma^2 k}{(\lambda_i + k)^4} [2c_i \lambda_i + c_i k - \lambda_i^2] \right\}_{i=1}^p$$

This results in the matrix Δ being a positively definite (p.d.) matrix if and only if all its diagonal elements are non-negative. This simplifies to:

$$k > \frac{\lambda_i^2}{c_i} - 2\lambda_i$$

Theorem 4.2 The estimator of $\hat{\beta}_{DB-MUR}$ is superior than the estimator of $\hat{\beta}_{DB-ORR}$ in terms of the Mean Squared Error Matrix (MSEM) if and only if

$$k > \frac{\lambda_i^2}{c_i} - 2\lambda_i \quad \forall i = 1, \dots, p \tag{28}$$

4.3 Comparing the $\hat{\beta}_{DB-MUR}$ estimator with the $\hat{\beta}_{DB-URR}$ estimator

The difference between the variance matrices can be written from equations (19) and (25) as follows:

$$\begin{aligned} V &= Var(\hat{\beta}_{DB-URR}) - Var(\hat{\beta}_{DB-MUR}) \\ &= \sigma^2 C_k [C + kI] C_k - \sigma^2 W C_k [C + kI] C_k W' \end{aligned} \tag{29}$$

Since $W = I - k C_k$ is a shrinking matrix, its eigenvalues are strictly bounded between 0 and 1, making V a strictly positive definite matrix. Now let's look at the difference between the Mean Squared Error Matrix (MSEM):

$$\begin{aligned} \Delta &= MSEM(\hat{\beta}_{DB-URR}) - MSEM(\hat{\beta}_{DB-MUR}) \\ &= [Var(\hat{\beta}_{DB-URR})] - [Var(\hat{\beta}_{DB-MUR}) + (b_1)(b_1)'] \\ &= \sigma^2 C_k [C + kI] C_k - \sigma^2 W C_k [C + kI] C_k W' - k^2 C_k \beta \beta' C_k \\ &= k^2 C_k \left[\sigma^2 \left[\frac{1}{k^2} (C + kI) - \frac{1}{k^2} W(C + kI)W' \right] - \beta \beta' \right] C_k \\ &= k^2 C_k [\sigma^2 A_1 - \beta \beta'] C_k \end{aligned}$$

where $A = \frac{1}{k^2} [(C + kI) - W(C + kI)W']$. since A is a positive definite matrix and $\sigma^2 > 0$. (see, e.g.,Trenkler and Toutenburg 1990 [11]; Rao and Toutenburg 1995 [12]; Groß 2003 [13]).By Lemma 4.2 (Farebrother, 1976)[16], we obtain $\sigma^2 A - \beta \beta'$ and Δ is nonnegative definite matrix if and only if

$$\beta' A^{-1} \beta \leq \sigma^2$$

Theorem 4.3 The estimator of $\hat{\beta}_{DB-MUR}$ is superior than the estimator of $\hat{\beta}_{DB-URR}$ in terms of the Mean Squared Error Matrix (MSEM) if and only if

$$\beta' A^{-1} \beta \leq \sigma^2 \quad , \quad \forall i = 1, \dots, p \tag{30}$$

5. Simulation Study and numerical example

5.1 Simulation Study

In this section, the Estimated Mean Squared Error (EMSE) is employed as a criterion to compare the estimators and determine the optimal performance. The effects of sample size and correlation strength were considered, alongside the selection of various values for the ridge parameter (k). To meet the requirements for robust testing three different sample sizes were chosen: $n = 50, 100, 200$. To examine the estimators' consistency, small, medium, and large samples were used. Additionally, four levels of correlation were investigated: $\rho = 0.8, 0.9, 0.95, 0.99$. These numbers were intentionally chosen to represent high to severe levels of multicollinearity, which is the primary difficulty that these estimators are intended to address. with the number of variables set to ($p = 5$). The standard deviation of the error terms was set to three levels: $\sigma = 0.1, 1, 5$. Third-order difference coefficients ($m = 3$) were utilized, with values $d_0 = 0.8582, d_1 = -0.3832, d_2 = -0.2809, d_3 = -0.1942$. Optimal difference coefficients for $1 \leq m \leq 10$ can be found in Yatchew (2003)[8].The explanatory (independent) variables were generated using a Monte Carlo simulation method based on the equation proposed by McDonald and Galarneau (1975) [14] :

$$x_{ij} = (1 - \rho^2)^{1/2} w_{ij} + \rho w_{i(p+1)}, \quad i = 1, 2, \dots, n; j = 1, 2, \dots, p \quad (31)$$

where w_{ij} are independent pseudo-random numbers following the standard normal distribution, and ρ^2 represents the theoretical correlation between any two explanatory variables. The model is generated according to the following equation:

$$y_i = \sum_{j=1}^5 x_{ij} \beta_j + f(t_i) + \varepsilon_i \quad (32)$$

where $\beta = (1.5, 3, 2, -5, 4)'$, and $\varepsilon_i \sim N(0, \sigma^2)$. The non-parametric component is defined by the Doppler function $f(t_i) = \sqrt{t_i(1 - t_i)} \sin(2.1\pi/(t_i + 0.05))$, where $t_i = (i - 0.5)/n$ for $i = 1, 2, \dots, n$. The experiment was replicated 2000 times by generating new error terms while the matrix \mathbf{X} , the vector β , and the function \mathbf{f} remained fixed. This large number of replications ensures that the variations in EMSE are strictly due to the random component and guarantees statistically stable results. After generating 2000 samples, the estimated Mean Squared Error (EMSE) was calculated for each estimator. The EMSE is defined by the formula:

$$EMSE(\hat{\beta}) = \frac{1}{2000} \sum_j^{2000} \sum_i^5 (\hat{\beta}_{ij} - \beta_i)^2 \quad (33)$$

where $\hat{\beta}_{ij}$ denotes the estimate of the i parameter in the j replication, and β_i represents the true parameter values. Following the approach suggested by Crouse et al, the prior information vector \mathbf{J} was determined empirically. To maintain consistency with the difference-based structure of our model, we utilized the Difference-Based Ordinary Least Squares DBOLS estimates. Specifically, we calculated the arithmetic mean of the DBOLS coefficients and assigned this average value to all elements of the prior vector. Mathematically, \mathbf{J} is defined as: $\mathbf{J} = \left(\frac{1}{p} \sum_{i=1}^p \beta_{i,DBOLS} \right) \times \mathbf{1}_{p \times 1}$, where $\mathbf{1}_{p \times 1}$ is a vector of ones. It is also worth noting that when $k = 0$, the estimators become equivalent, such that $\hat{\beta}_{DB-OLS} = \hat{\beta}_{DB-ORR} = \hat{\beta}_{DB-URR} = \hat{\beta}_{DB-MUR}$. The simulation was conducted using the MATLAB R2025b software package.

		n=200 $\sigma=0.1$																
		$\rho=0.8$					$\rho=0.9$					$\rho=0.95$						
		0	0.001	0.005	0.015	0.03	0.04	0.05	0.07	0.08	0.09	0.1	0.3	0.5	0.7	0.9	1.5	5
DB-OLS	0.001157	0.001157	0.001157	0.001157	0.001157	0.001157	0.001157	0.001157	0.001157	0.001157	0.001157	0.001157	0.001157	0.001157	0.001157	0.001157	0.001157	
DB-ORR	0.001157	0.001155	0.001149	0.001148	0.001146	0.001143	0.001142	0.001128	0.001115	0.001104	0.001095	0.001079	0.001074	0.001072	0.001067	0.001064	0.001062	
DB-URR	0.001157	0.001155	0.001149	0.001148	0.001146	0.001143	0.001142	0.001127	0.001115	0.001103	0.001093	0.001078	0.001072	0.001067	0.001064	0.001062	0.001061	
DB-MUR	0.001157	0.001154	0.001142	0.001139	0.001136	0.00113	0.001128	0.001104	0.001086	0.001073	0.001066	0.001068	0.001077	0.001092	0.001112	0.002666	0.006371	
Best		DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-OLS	DB-OLS	DB-OLS	
k	0	0.001	0.005	0.006	0.007	0.009	0.01	0.02	0.03	0.04	0.05	0.07	0.08	0.09	0.1	0.3	0.5	
DB-OLS	0.002173	0.002173	0.002173	0.002173	0.002173	0.002173	0.002173	0.002173	0.002173	0.002173	0.002173	0.002173	0.002173	0.002173	0.002173	0.002173	0.002173	
DB-ORR	0.002173	0.002169	0.002163	0.002165	0.002166	0.002169	0.002173	0.002163	0.002153	0.002146	0.002139	0.002135	0.002103	0.002075	0.002030	0.002014	0.002022	
DB-URR	0.002173	0.002169	0.002163	0.002165	0.002166	0.002169	0.002173	0.002163	0.002153	0.002146	0.002139	0.002135	0.002103	0.002075	0.002030	0.002014	0.002022	
DB-MUR	0.002173	0.002165	0.002153	0.002149	0.002145	0.002138	0.002135	0.002102	0.002073	0.00205	0.002032	0.002011	0.002008	0.00201	0.002017	0.003185	0.006281	
Best		DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-URR	DB-URR	DB-URR	
k	0	0.001	0.005	0.006	0.007	0.009	0.01	0.02	0.03	0.04	0.05	0.07	0.08	0.09	0.1	0.3	0.5	
DB-OLS	0.004197	0.004197	0.004197	0.004197	0.004197	0.004197	0.004197	0.004197	0.004197	0.004197	0.004197	0.004197	0.004197	0.004197	0.004197	0.004197	0.004197	
DB-ORR	0.004197	0.004185	0.004141	0.00413	0.00412	0.004099	0.004089	0.00404	0.003931	0.00388	0.003848	0.00384	0.003864	0.003907	0.003969	0.009033	0.021141	
DB-URR	0.004197	0.004185	0.004141	0.00413	0.00412	0.004099	0.004089	0.00404	0.003931	0.00388	0.003848	0.00384	0.003864	0.003907	0.003969	0.009033	0.021141	
DB-MUR	0.004197	0.004174	0.004069	0.004051	0.004016	0.003999	0.003878	0.003831	0.003861	0.003965	0.004398	0.004726	0.005128	0.005603	0.030119	0.081696	0.158227	
Best		DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-URR	DB-URR	DB-URR	
k	0	0.001	0.005	0.006	0.007	0.009	0.01	0.02	0.03	0.04	0.05	0.07	0.08	0.09	0.1	0.3	0.5	
DB-OLS	0.021169	0.021169	0.021169	0.021169	0.021169	0.021169	0.021169	0.021169	0.021169	0.021169	0.021169	0.021169	0.021169	0.021169	0.021169	0.021169	0.021169	
DB-ORR	0.021169	0.021034	0.020542	0.020431	0.020324	0.020124	0.020031	0.019353	0.019129	0.019353	0.020019	0.022654	0.024612	0.02699	0.029782	0.164839	0.426614	
DB-URR	0.021169	0.021034	0.020539	0.020427	0.02032	0.020119	0.020026	0.019343	0.019115	0.019336	0.02	0.022632	0.024589	0.026987	0.164942	0.427008	0.786656	
DB-MUR	0.021169	0.020904	0.020028	0.019855	0.019701	0.019447	0.019347	0.019348	0.021135	0.024674	0.029931	0.045464	0.055676	0.067475	0.08083	0.631419	1.594097	
Best		DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-ORR	DB-ORR	DB-ORR	DB-ORR	DB-ORR	DB-ORR	
k	0	0.001	0.005	0.006	0.007	0.009	0.01	0.02	0.03	0.04	0.05	0.07	0.08	0.09	0.1	0.3	0.5	
DB-OLS	0.066102	0.066102	0.066102	0.066102	0.066102	0.066102	0.066102	0.066102	0.066102	0.066102	0.066102	0.066102	0.066102	0.066102	0.066102	0.066102	0.066102	
DB-ORR	0.066102	0.066099	0.066088	0.066085	0.066082	0.066076	0.066073	0.066045	0.066019	0.065994	0.065971	0.065874	0.065908	0.06589	0.065874	0.06583	0.067367	
DB-URR	0.066102	0.066099	0.066088	0.066085	0.066082	0.066076	0.066073	0.066045	0.066019	0.065994	0.065971	0.065874	0.065908	0.06589	0.065874	0.06583	0.067367	
DB-MUR	0.066102	0.066096	0.066073	0.066067	0.066062	0.066051	0.066045	0.065994	0.065948	0.065907	0.065872	0.065839	0.065801	0.065788	0.065781	0.066787	0.069953	
Best		DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-URR	DB-URR	DB-URR	
k	0	0.001	0.005	0.006	0.007	0.009	0.01	0.02	0.03	0.04	0.05	0.07	0.08	0.09	0.1	0.3	0.5	
DB-OLS	0.122548	0.122548	0.122548	0.122548	0.122548	0.122548	0.122548	0.122548	0.122548	0.122548	0.122548	0.122548	0.122548	0.122548	0.122548	0.122548	0.122548	
DB-ORR	0.122548	0.122538	0.122498	0.122489	0.122479	0.12246	0.122451	0.122359	0.122272	0.122191	0.122114	0.121975	0.121914	0.121857	0.121805	0.12374	0.172557	
DB-URR	0.122548	0.122538	0.122498	0.122489	0.122479	0.12246	0.122451	0.122358	0.122271	0.122189	0.122112	0.121973	0.121914	0.121857	0.121805	0.12374	0.172557	
DB-MUR	0.122548	0.122528	0.122432	0.122413	0.122377	0.122359	0.122319	0.122204	0.122112	0.121912	0.121804	0.121646	0.121597	0.121568	0.121559	0.123653	0.155716	
Best		DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-MUR	DB-ORR	DB-ORR	DB-ORR	

Table 3. EMSE values of the estimators for when n=200.

The simulation results presented in the tables above demonstrate that higher correlation levels lead to a significant deterioration in the performance of the DB-OLS estimator. This is evidenced by the increased Estimated Mean Squared Error (EMSE) values resulting from the exacerbated multicollinearity problem. In Table 1 $n = 50$, with $\sigma = 0.1$ and across all correlation levels, the DB-MUR estimator performs best at small values of k . However, for $\sigma = 1$ and $\sigma = 5$, the DB-MUR estimator is superior for nearly all values of k . Regarding Table 2 $n = 100$, with $\sigma = 0.1$ and for all correlation levels, the DB-MUR estimator is optimal at small values of k , while the DB-URR estimator performs better at intermediate values of k . For $\sigma = 1$ and $\sigma = 5$, the DB-MUR estimator consistently outperforms the others across almost all values of k . Similarly, in Table 3 $n = 200$, with $\sigma = 0.1$ and across all correlation levels, the DB-MUR estimator is preferred at small values of k , and the DB-URR estimator is superior at intermediate values of k . When $\sigma = 1$ and $\sigma = 5$, the DB-MUR estimator remains the best choice for nearly all values of k .

To provide further theoretical interpretations of these results, the consistent superiority of the $\hat{\beta}_{DB-MUR}$ estimator under severe multicollinearity and high variance levels (i.e., $\sigma = 1$ and $\sigma = 5$) can be explained by its modified shrinkage mechanism. In contrast to the standard $\hat{\beta}_{DB-ORR}$, the DB-MUR estimator successfully balances the bias-variance trade-off; it mitigates the inflated variance caused by strongly correlated variables without incurring an excessive bias penalty. Meanwhile, the robust performance of the $\hat{\beta}_{DB-URR}$ estimator at intermediate values of k perfectly reflects the theoretical bounds established in our theorems, confirming that the reduction in variance effectively outweighs the introduced bias.

5.2 A numerical example

This section presents a numerical example to illustrate the application of the theoretical concepts outlined in Section 3. Yatchew (2000) [2] generated the data, which was subsequently considered by Tabakan and Akdeniz (2010) [5]. The data originated from a 1993 survey of 81 municipal power distribution sites in Ontario, Canada. It is well-established that the partial linear model is a straightforward quasi-parametric extension of the Cobb-Douglas model. We investigate a simpler version of the Cobb-Douglas model for electricity distribution expenses

$$tc = f(cust) + \beta_1 wage + \beta_2 pcap + \beta_3 puc + \beta_4 kWh + \beta_5 life + \beta_6 lf + \beta_7 kmwire + \epsilon \tag{34}$$

where tc is the logarithm of total cost per customer, $cust$ is the logarithm of the number of customers, $wage$ is the logarithm of the wage rate, $pcap$ is the logarithm of the capital price, puc is a dummy variable for utility committees that provide additional services and may benefit from economies of scale, kWh is the logarithm of kilowatt-hours per customer, $life$ is the logarithm of the remaining life of the distribution assets, lf is the logarithm of the load factor, and kilometers represents the kilometers of distribution wire [5]. It is clear that the model (34) contains both parametric and nonparametric effects. In this part, we employ the method given in this study to estimate β , the differentiating procedure. In this paper, we look at the differencing coefficients defined in Section 5.1. We now denote the $(81 - 3) \times 81$ differentiating matrix D as follows

$$D = \begin{pmatrix} d_0 & d_1 & \dots & d_m & 0 & 0 & \dots & 0 \\ 0 & d_0 & d_1 & \dots & d_m & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & d_1 & \dots & d_m & 0 & 0 \\ 0 & 0 & \dots & d_0 & d_1 & \dots & d_m & 0 \\ 0 & 0 & \dots & 0 & d_0 & d_1 & \dots & d_m \end{pmatrix}$$

The matrix $\tilde{X}'\tilde{X}$ contains eigenvalues $\lambda_1 = 0.1133, \lambda_2 = 0.2341, \lambda_3 = 1.5387, \lambda_4 = 2.6188, \lambda_5 = 3.4198, \lambda_6 = 15.1399, \text{ and } \lambda_7 = 19.4871$, and a condition number $CN = \lambda_7/\lambda_1 = 171.9952$, indicating severe multicollinearity in the dataset. We want to compare the traces of the estimated

Scalar mean squared error (SMSE) matrices $\hat{\beta}_{DB-OLS}, \hat{\beta}_{DB-ORR}, \hat{\beta}_{DB-URR}, \hat{\beta}_{DB-MUR}$. Substituting β and σ^2 with their corresponding difference estimates, $\hat{\beta}_{DB-OLS}$ and $\hat{\sigma}^2$, and selecting the Ridge parameter k for a specified interval yields the estimated SMSE. The Trace of the estimated SMSE of the $\hat{\beta}_{DB-OLS}$ estimator is represented as

$$SMSE(\hat{\beta}_{DB-OLS}) = tr[\hat{\sigma}^2 UCU] \tag{35}$$

where $U = (\tilde{X}'\tilde{X})^{-1}$

The estimated SMSE of the $\hat{\beta}_{DB-ORR}$ estimator is represented as

$$SMSE(\hat{\beta}_{DB-ORR}) = tr[C_k[\hat{\sigma}^2 C + k^2 \hat{\beta}_{DB-OLS} \hat{\beta}_{DB-OLS}']C_k] \tag{36}$$

The estimated SMSE of the $\hat{\beta}_{DB-URR}$ estimator is represented as

$$SMSE(\hat{\beta}_{DB-URR}) = tr[\hat{\sigma}^2 C_k[C + kI]C_k] \tag{37}$$

The estimated SMSE of the estimator, $\hat{\beta}_{DB-MUR}$ is expressed as follows

$$SMSE(\hat{\beta}_{DB-MUR}) = tr[\hat{\sigma}^2 W C_k[C + kI]C_k W' + k^2 C_k \hat{\beta}_{DB-OLS} \hat{\beta}_{DB-OLS}' C_k] \tag{38}$$

Table 4. Estimated variance and SMSE values for the $\hat{\beta}_{DB-OLS}(k = 0)$ and $\hat{\beta}_{DB-ORR}$ estimators

Coef	k=0	k=0.001	k=0.005	k=0.006	k=0.007	k=0.009	k=0.01	k=0.02	k=0.03	k=0.04	k=0.05	k=0.07	k=0.08	k=0.09	k=0.1	k=0.3	k=0.5	k=0.7	k=0.9	k=1.5	k=5
Wage	0.622935	0.618396	0.600923	0.596718	0.592575	0.584469	0.580504	0.543774	0.511622	0.483209	0.457895	0.414671	0.396049	0.379058	0.363486	0.199411	0.1365	0.103161	0.082564	0.051002	0.015063
Pcap	0.545191	0.544512	0.541876	0.541237	0.540604	0.53936	0.538748	0.532967	0.527717	0.522899	0.51844	0.510381	0.506699	0.503209	0.499886	0.451939	0.419408	0.393841	0.372688	0.325402	0.209156
Puc	-0.07451	-0.07463	-0.0751	-0.07522	-0.07533	-0.07555	-0.07566	-0.07666	-0.07754	-0.07833	-0.07904	-0.08028	-0.08082	-0.08131	-0.08177	-0.087	-0.08932	-0.09064	-0.09144	-0.09231	-0.08628
kWh	0.00819	0.008988	0.01205	0.012784	0.013506	0.014914	0.015602	0.021917	0.027364	0.032112	0.036288	0.043296	0.046267	0.04895	0.051387	0.07553	0.083583	0.087038	0.088467	0.087867	0.067549
Lf	-0.6277	-0.62722	-0.62532	-0.62485	-0.62438	-0.62344	-0.62298	-0.61839	-0.61391	-0.60952	-0.59686	-0.59278	-0.58877	-0.58482	-0.51599	-0.46146	-0.41703	-0.38015	-0.29953	-0.13019	
Life	1.32687	1.314893	1.269036	1.258059	1.247266	1.226218	1.215953	1.121885	1.041118	0.971032	0.909651	0.807265	0.764135	0.725314	0.690191	0.348485	0.232	0.173546	0.138485	0.086035	0.026407
Kmwire	0.412594	0.412321	0.411232	0.410956	0.410688	0.410146	0.409874	0.407176	0.404503	0.401857	0.399239	0.394094	0.391566	0.389068	0.3866	0.342788	0.307326	0.27791	0.253067	0.197211	0.071891
VGr	0.367754	0.362914	0.344644	0.340331	0.336114	0.327956	0.324008	0.288793	0.259934	0.235924	0.215685	0.183573	0.170667	0.159368	0.149407	0.065939	0.043211	0.032731	0.025699	0.017266	0.005364
mSe	0.367754	0.36308	0.348507	0.345802	0.343437	0.33967	0.338238	0.337536	0.354917	0.383597	0.419203	0.500706	0.543446	0.586252	0.628565	1.233287	1.539012	1.724955	1.854077	2.093096	2.546244
R ²	0.652837	0.652833	0.652736	0.652694	0.652645	0.65253	0.652464	0.651553	0.650322	0.648907	0.647393	0.644264	0.642704	0.641167	0.639663	0.616386	0.599732	0.585423	0.572413	0.538848	0.431262

Table 5. Estimated variance and SMSE values for the $\hat{\beta}_{DB-OLS}(k = 0)$ and $\hat{\beta}_{DB-URR}$ estimators

Coef	k=0	k=0.001	k=0.005	k=0.006	k=0.007	k=0.009	k=0.01	k=0.02	k=0.03	k=0.04	k=0.05	k=0.07	k=0.08	k=0.09	k=0.1	k=0.3	k=0.5	k=0.7	k=0.9	k=1.5	k=5
Wage	0.622935	0.622935	0.622935	0.622935	0.622935	0.622935	0.622935	0.622935	0.622935	0.622935	0.622935	0.622935	0.622935	0.622935	0.622935	0.622935	0.622935	0.622935	0.622935	0.622935	0.622935
Pcap	0.545191	0.545191	0.545191	0.545191	0.545191	0.545191	0.545191	0.545191	0.545191	0.545191	0.545191	0.545191	0.545191	0.545191	0.545191	0.545191	0.545191	0.545191	0.545191	0.545191	0.545191
Puc	-0.07451	-0.07451	-0.07451	-0.07451	-0.07451	-0.07451	-0.07451	-0.07451	-0.07451	-0.07451	-0.07451	-0.07451	-0.07451	-0.07451	-0.07451	-0.07451	-0.07451	-0.07451	-0.07451	-0.07451	-0.07451
kWh	0.00819	0.00819	0.00819	0.00819	0.00819	0.00819	0.00819	0.00819	0.00819	0.00819	0.00819	0.00819	0.00819	0.00819	0.00819	0.00819	0.00819	0.00819	0.00819	0.00819	0.00819
Lf	-0.6277	-0.6277	-0.6277	-0.6277	-0.6277	-0.6277	-0.6277	-0.6277	-0.6277	-0.6277	-0.6277	-0.6277	-0.6277	-0.6277	-0.6277	-0.6277	-0.6277	-0.6277	-0.6277	-0.6277	-0.6277
Life	1.32687	1.32687	1.32687	1.32687	1.32687	1.32687	1.32687	1.32687	1.32687	1.32687	1.32687	1.32687	1.32687	1.32687	1.32687	1.32687	1.32687	1.32687	1.32687	1.32687	1.32687
Kmwire	0.412594	0.412594	0.412594	0.412594	0.412594	0.412594	0.412594	0.412594	0.412594	0.412594	0.412594	0.412594	0.412594	0.412594	0.412594	0.412594	0.412594	0.412594	0.412594	0.412594	0.412594
VGr	0.367754	0.365325	0.359976	0.353726	0.351511	0.347176	0.345056	0.325424	0.308251	0.293082	0.279571	0.256502	0.246561	0.237487	0.229167	0.140697	0.105551	0.086089	0.073499	0.052597	0.022261
mSe	0.367754	0.365325	0.359976	0.353726	0.351511	0.347176	0.345056	0.325424	0.308251	0.293082	0.279571	0.256502	0.246561	0.237487	0.229167	0.140697	0.105551	0.086089	0.073499	0.052597	0.022261
R ²	0.652837	0.652837	0.652837	0.652837	0.652837	0.652837	0.652837	0.652837	0.652837	0.652837	0.652837	0.652837	0.652837	0.652837	0.652837	0.652837	0.652837	0.652837	0.652837	0.652837	0.652837

Table 6. Estimated variance and SMSE values for the $\hat{\beta}_{DB-OLS}(k = 0)$ and $\hat{\beta}_{DB-MUR}$ estimators

Coef	k=0	k=0.001	k=0.005	k=0.006	k=0.007	k=0.009	k=0.01	k=0.02	k=0.03	k=0.04	k=0.05	k=0.07	k=0.08	k=0.09	k=0.1	k=0.3	k=0.5	k=0.7	k=0.9	k=1.5	k=5
Wage	0.622935	0.618396	0.600923	0.596718	0.592575	0.584469	0.580504	0.543774	0.511622	0.483209	0.457895	0.414671	0.396049	0.379058	0.363486	0.199411	0.1365	0.103161	0.082564	0.051002	0.015063
Pcap	0.545191	0.544512	0.541876	0.541237	0.540604	0.53936	0.538748	0.532967	0.527717	0.522899	0.51844	0.510381	0.506699	0.503209	0.499886	0.451939	0.419408	0.393841	0.372688	0.325402	0.209156
Puc	-0.07451	-0.07463	-0.0751	-0.07522	-0.07533	-0.07555	-0.07566	-0.07666	-0.07754	-0.07833	-0.07904	-0.08028	-0.08082	-0.08131	-0.08177	-0.087	-0.08932	-0.09064	-0.09144	-0.09231	-0.08628
kWh	0.00819	0.008988	0.01205	0.012784	0.013506	0.014914	0.015602	0.021917	0.027364	0.032112	0.036288	0.043296	0.046267	0.04895	0.051387	0.07553	0.083583	0.087038	0.088467	0.087867	0.067549
Lf	-0.6277	-0.62722	-0.62532	-0.62485	-0.62438	-0.62344	-0.62298	-0.61839	-0.61391	-0.60952	-0.59686	-0.59278	-0.58877	-0.58482	-0.51599	-0.46146	-0.41703	-0.38015	-0.29953	-0.13019	
Life	1.32687	1.314893	1.269036	1.258059	1.247266	1.226218	1.215953	1.121885	1.041118	0.971032	0.909651	0.807265	0.764135	0.725314	0.690191	0.348485	0.232	0.173546	0.138485	0.086035	0.026407
Kmwire	0.412594	0.412321	0.411232	0.410956	0.410688	0.410146	0.409874	0.407176	0.404503	0.401857	0.399239	0.394094	0.391566	0.389068	0.3866	0.342788	0.307326	0.27791	0.253067	0.197211	0.071891
VGr	0.367754	0.360523	0.33374	0.327538	0.321519	0.309998	0.304485	0.257053	0.220597	0.191977	0.169097	0.135214	0.122459	0.111713	0.02569	0.03887	0.025328	0.019244	0.015581	0.009742	0.002478
mSe	0.367754	0.366688	0.337603	0.333009	0.328842	0.321713	0.318715	0.305796	0.315579	0.33965	0.372616	0.452347	0.495239	0.538596	0.581727	1.206218	1.521129	1.711469	1.943059	2.085572	2.543359
R ²	0.652837	0.652833	0.652736	0.652694	0.652645	0.65253	0.652464	0.651553	0.650322	0.648907	0.647393	0.644264	0.642704	0.641167	0.639663	0.616386	0.599732	0.585423	0.572413	0.538848	0.431262

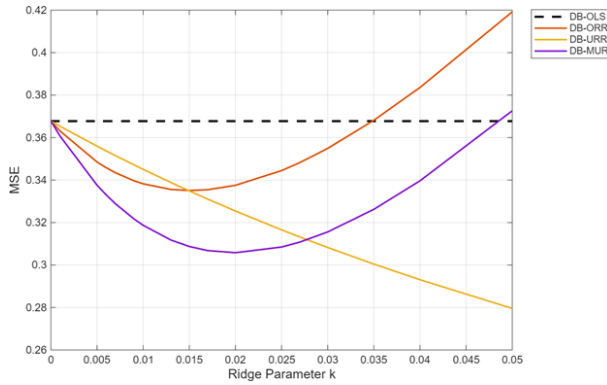


Fig. 1. Estimated SMSE values of the estimators $\hat{\beta}_{DB-OLS}(k = 0)$, $\hat{\beta}_{DB-ORR}$, and $\hat{\beta}_{DB-URR}$, and $\hat{\beta}_{DB-MUR}$ for varied $k > 0$

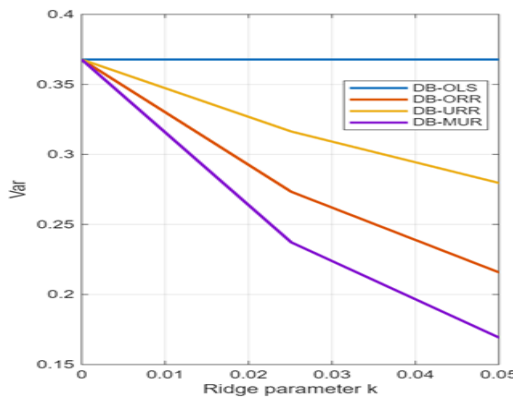


Fig. 2. Estimated variance value of the estimators $\hat{\beta}_{DB-OLS}(k = 0)$, $\hat{\beta}_{DB-ORR}$, and $\hat{\beta}_{DB-URR}$, and $\hat{\beta}_{DB-MUR}$ for varied $k > 0$

Fig. 1. displays the estimated SMSE values for the estimators $\hat{\beta}_{DB-OLS}$, $\hat{\beta}_{DB-ORR}$, $\hat{\beta}_{DB-URR}$, and $\hat{\beta}_{DB-MUR}$. The $\hat{\beta}_{DB-MUR}$ estimator outperforms the other estimators in the range $0 < k \leq 0,02$, whereas the $\hat{\beta}_{DB-URR}$ estimator demonstrates superior performance for $k \geq 0.03$. These findings are consistent with Theorems 4.1, 4.2, and 4.3. Furthermore, Fig. 2. illustrates the estimated variance values for the estimators $\hat{\beta}_{DB-OLS}$, $\hat{\beta}_{DB-ORR}$, $\hat{\beta}_{DB-URR}$ and $\hat{\beta}_{DB-MUR}$.

6. Conclusions

In this paper, we propose two new estimators called the Difference-Based Unbiased Ridge Estimator ($\hat{\beta}_{DB-URR}$) and the Difference-Based Modified Unbiased Ridge Estimator ($\hat{\beta}_{DB-MUR}$). These estimators are designed to be applied in the presence of multicollinearity within partial linear models. A comprehensive comparison was conducted in terms of the Mean Squared Error Matrix (MSEM) between the two new estimators and both the Difference-Based Ordinary Least Squares estimator ($\hat{\beta}_{DB-OLS}$) and the Difference-Based Ordinary Ridge Estimator ($\hat{\beta}_{DB-ORR}$). Furthermore, a simulation study and a numerical example are presented to demonstrate the performance of the proposed estimators based on the MSEM criterion. Theoretical results indicate that the $\hat{\beta}_{DB-MUR}$ estimator outperforms the other estimators when small values of the ridge parameter k are selected, whereas the $\hat{\beta}_{DB-URR}$ estimator proves superior for intermediate and large values of the ridge parameter k .

Acknowledgment

The authors would also like to thank the University of Anbar for its support, and the editorial board and reviewers for their valuable comments and suggestions, which greatly enhanced the quality and clarity of the overall manuscript.

References

1. A. J. E. I. Yatchew, "An elementary estimator of the partial linear model," vol. 57, no. 2, pp. 135-143, 1997.
2. A. J. J. o. a. E. Yatchew, "Scale economies in electricity distribution: A semiparametric analysis," vol. 15, no. 2, pp. 187-210, 2000.
3. L. Wang, L. Brown, and T. Cai, "A difference-based approach to semiparametric partial linear model. Department of Statistics, The Wharton School University of Pennsylvania," Pennsylvania, Technical report2007.
4. A. E. Hoerl and R. W. J. T. Kennard, "Ridge regression: Biased estimation for nonorthogonal problems," vol. 12, no. 1, pp. 55-67, 1970.
5. G. Tabakan and F. J. S. P. Akdeniz, "Difference-based ridge estimator of parameters in partial linear model," vol. 51, no. 2, pp. 357-368, 2010.
6. R. H. Crouse, C. Jin, R. J. C. i. S.-T. Hanumara, and Methods, "Unbiased ridge estimation with prior information and ridge trace," vol. 24, no. 9, pp. 2341-2354, 1995.
7. F. S. M. Batah, S. D. J. S. i. M. Gore, and i. Applications, "Ridge regression estimator: Combining unbiased and ordinary ridge regression methods of estimation," vol. 4, pp. 99-109, 2009.
8. A. Yatchew, *Semiparametric regression for the applied econometrician*. Cambridge University Press, 2003.
9. R. Eubank *et al.*, "Estimation in partially linear models," vol. 29, no. 1, pp. 27-34, 1998.
10. K. Klipple and R. J. F. i. h. o. D. P. M. M. A. o. t. o. o. h. r. Eubank, "Difference-based variance estimators for partially linear models," pp. 313-323, 2007.
11. G. Trenkler and H. J. S. p. Toutenburg, "Mean squared error matrix comparisons between biased estimators an overview of recent results," vol. 31, no. 1, pp. 165-179, 1990.
12. C. R. Rao and H. Toutenburg, "Linear models," in *Linear Models: Least Squares and Alternatives*: Springer, 1995, pp. 3-18.
13. J. Groß, *Linear regression*. Springer Science & Business Media, 2003.
14. G. C. McDonald and D. I. J. J. o. t. A. S. A. Galarneau, "A Monte Carlo evaluation of some ridge-type estimators," vol. 70, no. 350, pp. 407-416, 1975.
15. M. I. Alheety, H. Nayem, and B. G. J. S. Kibria, "An unbiased convex estimator depending on prior information for the classical linear regression model," vol. 8, no. 1, p. 16, 2025.
16. R. W. Farebrother, "Further results on the mean square error of ridge regression," *Journal of the Royal Statistical Society: Series B*, vol. 38, no. 3, pp. 248-250, 1976.