AIMS Mathematics

AIMS Mathematics, 7(1): 1035–1057.
DOI: 10.3934/math.2022062
Received: 02 July 2021
Accepted: 13 October 2021
Published: 19 October 2021

http://www.aimspress.com/journal/Math

Research article

A new modified ridge-type estimator for the beta regression model: simulation and application

Muhammad Nauman Akram[1], Muhammad Amin[1], Ahmed Elhassanein[2,3,]* and Muhammad Aman Ullah[4]

[1] Department of Statistics, University of Sargodha, Sargodha, Pakistan
[2] Department of Mathematics, College of Science, University of Bisha, Bisha, Saudi Arabia
[3] Department of Mathematics, Damanhour University, Damanhour, Egypt
[4] Department of Statistics, Bahauddin Zakariya University, Multan, Pakistan

* **Correspondence:** Email: el_hassanein@yahoo.com; Tel: +201064941983.

Abstract: The beta regression model has become a popular tool for assessing the relationships among chemical characteristics. In the BRM, when the explanatory variables are highly correlated, then the maximum likelihood estimator (MLE) does not provide reliable results. So, in this study, we propose a new modified beta ridge-type (MBRT) estimator for the BRM to reduce the effect of multicollinearity and improve the estimation. Initially, we show analytically that the new estimator outperforms the MLE as well as the other two well-known biased estimators i.e., beta ridge regression estimator (BRRE) and beta Liu estimator (BLE) using the matrix mean squared error (MMSE) and mean squared error (MSE) criteria. The performance of the MBRT estimator is assessed using a simulation study and an empirical application. Findings demonstrate that our proposed MBRT estimator outperforms the MLE, BRRE and BLE in fitting the BRM with correlated explanatory variables.

Keywords: beta regression; BRRE; BLE; MBRTE; multicollinearity; MSE
Mathematics Subject Classification: 62F10, 62J07

1. Introduction

Regression analysis is one of the most important tools which has several applications in chemometrics and other fields [39–40,52–54]. There are various types of regression models available in the literature i.e., linear model, non-linear model, generalized linear models and nonparametric regression model [36,39]. The choice of the suitable regression model is the most important task for obtaining accurate and reliable results [37,39]. The choice of suitable regression depends on the nature and distribution of the response variable [37,38].

In many research areas, there are situations in which the response variable is restricted to the interval $[0,1]$, such as rates and proportions. The classical solution is to transform the response variable so that it is mapped from the interval $[0,1]$ on the real line($\mathbb{R}$). In this situation (which is common when analyzing chemical, environmental, or biological data), the logistic regression model is considered, where the log-odds are used as the response variable in a linear regression model. Another example of a common transformation is the inverse of a suitable cumulative density function, which also leads to the response variable $Y \in [0,1]$ being mapped onto the real line. A well-known example of the latter is to use the probit model. But such kinds of solutions have several demerits; for example, the model parameters cannot be easily interpreted in terms of the original response variable. Another shortcoming is that proportion measures typically show asymmetry, and hence any inference based on the normality assumption can be deceptive, especially for a small sample.

As a remedy to the aforementioned problems, Ferrari and Cribari-Neto [49] proposed a beta regression model (BRM) for a continuous response variable (Y) with support in the open interval $[0,1]$. This model is based on the assumption that the response variable follows the beta distribution. Further, the model can also accommodate asymmetries and heteroscedasticity. Later developments have expanded the model so that it is also possible to include covariates for modeling dispersion [45]. Generally, the maximum likelihood estimator (MLE) is used to estimate the unknown regression coefficients of the BRM.

It is a common assumption in the multiple BRM that the regressors are not linearly correlated with one another. Though, in routine, explanatory variables may be inter-correlated, which causes the problem of multicollinearity [48]. In the presence of multicollinearity, the variance of the MLE becomes overstated, and the inference based on this estimator may not be reliable. Another consequence of multicollinearity is the wider confidence interval and probability of type-II error increases in hypothesis testing of unknown parameters [18]. However, many biased estimators have been introduced to combat multicollinearity in linear regression modeling (LRM), such as Stein estimator [11], ridge regression estimator [2] (RRE), improved ridge estimator [7], contraction estimator [23], modified RRE [6]. Liu estimator [21] (LE), Liu-type estimator [22], mixed ridge estimator [51], and modified Liu-type estimator [25]. Among them, RRE is the most common and attractive method initially proposed by Hoerl and Kennard [2]. This method was further developed and applied to chemical data by Vigneau et al. [12] There are several methods to estimate the shrinkage parameter (see, e.g., [9,28,39,41]). Unlike the LRM, the effect of multicollinearity on the generalized linear model (GLM) has been prolonged by Segrestedt [8]. For instance, Månsson and Shukur [16] introduced some ridge parameters for the logistic regression model, Månsson and Shukur [17] established a Poisson RRE. While Amin et al. [29] proposed a James-Stein estimator for the Poisson regression model. Amin et al. [39] examined the performance of the inverse Gaussian ridge regression estimator. Later Amin et al. [30] recommended some methods for estimating the shrinkage parameter based on the Gamma RRE. Recently Qasim et al. [31] introduced the RRE for the BRM.

Liu [21] introduced a new estimator, subsequently known as Liu estimator (LE). The primary objective of the LE is that the biasing Liu parameter d is the linear function of the estimates instead of a non-linear function as in ridge parameter k. This leads to a more stable shrinkage vector of estimated coefficients. Therefore, due to the linear function of d, researchers have used the more robust LE instead of the conventional RRE. Regarding the vast literature on LE for the LRM, we refer our readers primarily to Liu [21]. This method has also been further developed and applied to chemical data by Alheety and Kibria [25,32], Kibria [10], Li and Yang [51], and Akram et al. [33]. However, the literature on the LE for the GLM is very limited for example, Månsson et al. [18] suggested some shrinkage parameters for the Poisson Liu regression estimator. Månsson et al. [19] introduced a LE for the logit regression model. Månsson [20] recommended some Liu parameters for the negative binomial regression model. Qasim et al. [34] developed and adopted some new shrinkage parameters for LE in the gamma regression model. Recently, Karlsson et al. [47] introduced the LE for the BRM. Another estimator efficiently tackles the problem of multicollinearity is called the modified ridge-type estimator (MRTE). The very limited literature on the MRTE is available. We refer to the following studies: Lukman et al. [3] introduced the MRTE for the LRM and Lukman et al. [24] introduced some new ridge parameters for the LRM. Further, Lukman et al. [4] proposed a modified ridge-type logistic estimator. The available literature shows that no such study is available for the BRM. The extent to which these types of methods have been studied is due to the importance of the situation when the regressors are non-orthogonal since this is the most common situation when analyzing chemical, environmental, or biological data. So, this article aims to adopt the MRTE for the BRM named MBRTE and derived its theoretical properties to determine its effectiveness. In addition, we also provide the theoretical comparison of the proposed MBRT with other estimation methods i.e., MLE, RRE, and LE in a sense of matrix mean squared error (MMSE) and mean squared error (MSE) criteria.

The paper unfolds as follows: we define the model of interest and estimation procedures of BRM in Section 2. The theoretical comparison of the proposed estimator with other estimation methods is also outlined in this section. The selection of the appropriate biasing parameters is presented in Section 3, whereas the layout of the Monte Carlo simulation and its findings is discussed in Section 4. An empirical application is outlined in Section 5. Finally, this paper ends with some concluding remarks.

2. Methodology

2.1. *The beta regression model*

Suppose that $y_1, y_2, \dots, y_n$ are the observations of the random variable, Y follows a beta distribution with parameters $a, b > 0$, symbolized as $\beta_e(a, b)$ and the beta probability density function is stated as

$$f(y; a, b) = \frac{\Gamma(a+b)}{\Gamma(a)\,\Gamma(b)} y^{a-1}(1-y)^{b-1}, y \in (0,1), \tag{1}$$

where $\Gamma(.)$ is the gamma function and $a, b > 0$. The mean and variance of Y are $E(Y) = \frac{a}{a+b}$ and $Var(Y) = \frac{ab}{(a+b)^2(a+b+1)}$, respectively. Ferrari and Cribari-Neto [49] defined the parameterization for developing a regression model of beta distributed responses based on Eq (1). By supposing that $\mu = \frac{a}{a+b}$ and $\varphi = a + b$, and re-parameterize the parameters of Eq (1) by defining $a = \mu\varphi$ and $b = \varphi - \varphi\mu$, the density function of Y can be expressed through new parameterization as

$$f(y;\mu,\varphi)=\frac{\Gamma(\varphi)}{\Gamma(\mu\varphi)\Gamma(\varphi-\varphi\mu)}y^{\mu\varphi-1}(1-y)^{(\varphi-\varphi\mu-1)},y\in(0,1), \tag{2}$$

where $Y\sim\beta_e(\mu,\varphi)$ and limited in the interval $(0,1)$, $\mu(0<\mu<1)$ is denoted the mean of the response variable and $\varphi(\varphi>0)$ is the precision parameter. The mean and variance of Y are defined on the new parameterization, i.e., $E(Y)=\mu$ and $Var(Y)=\mu(1-\mu)/(1+\delta)$. However, the reciprocal of φ is called the dispersion parameter $(\delta=\varphi^{-1})$. The variance of the response decreases as φ is increased for fixed μ. The BRM is obtained by assuming that $y_i\sim\beta_e(\mu_i,\varphi), i=1,\dots,n$ and the link function is defined as

$$g(\mu_i)=\eta_i=x_i^t\beta,$$

where x_i is the ith row of X which is an $n\times p$ data matrix with p non-stochastic explanatory variables, $\beta=\left(\beta_1,\beta_2,\dots,\beta_p\right)^t$ is a $p\times 1$ vector of unknown regression coefficients, η_i is the linear predictor and $g(.)$ is the link function of the BRM, and it is strictly monotonic and twice differentiable such that $g(.):(0,1)\to\mathbb{R}$. Since different link functions may be used for fitting the BRM, for instance; logit, probit, log-log, complementary log-log, and Cauchy link functions. Among these, the commonly used link function is the logit link, i.e., $g(\mu_i)=log\left(\frac{\mu_i}{1-\mu_i}\right)$, which was suggested by Ferrari and Cribari-Neto [49]. The mean function of the response variable with logit link function is defined as

$$\mu_i=\frac{exp\,(x_i^t\beta)}{1+exp\,(x_i^t\beta)}, \tag{3}$$

where μ_i is the mean function of the response variable. Since η depends on β and μ is a function of η, the means $\mu_1,\dots,\mu_n$ are the functions of β. The log-likelihood function of Eq (2) is given by

$$\begin{aligned} l(\beta)&=\sum_{i=1}^{n} l_i\,(\mu_i,\varphi)\\ &=\sum_{i=1}^{n}[log\Gamma(\varphi)-log\Gamma(\mu_i\varphi)-log\Gamma(\varphi-\mu_i\varphi)\\ &\quad+(\varphi\mu_i-1)\log y_i+(\varphi-\mu_i\varphi-1)\log(1-y_i)]. \end{aligned} \tag{4}$$

Let $\hat{\beta}$ be the estimated value of the MLE of β. We are only interested to estimate the value of the parameter vector β. Generally, the MLE is used to estimate the unknown regression coefficients $\beta_j, j=1,\dots p$. The score function $\mathcal{S}(\beta)$ can be found as

$$\mathcal{S}(\beta)=\frac{\partial l(\beta)}{\partial\beta_j}=\sum_{i=1}^{n}\left(\frac{\partial l(\beta)}{\partial\theta_i}\frac{\partial\theta_i}{\partial\mu_i}\frac{\partial\mu_i}{\partial\eta_i}\frac{\partial\eta_i}{\partial\beta_j}\right).$$

Then the score function becomes

$$\mathcal{S}(\beta)=\sum_{i=1}^{n}\left(\frac{y_i-\mu_i}{\varphi\,Var(\mu_i)}\frac{\partial\mu_i}{\partial\eta_i}x_{ij}\right), \tag{5}$$

where $Var(\mu_i)=\mu_i(1-\mu_i)$ and x_{ij} is the ith values for jth covariates $x_j, j=1,\dots,p$. Since Eq (5) is non-linear in β_j, so, the solution of Eq (5) which can be obtained using Fisher's scoring iterative procedure as

$$\beta_{m+1}=\beta_m+\left\{\frac{\partial^2 l(\beta)}{\partial\beta_j\beta_k}(\beta_m)\right\}^{-1}\mathcal{S}(\beta_m), \tag{6}$$

where $\mathcal{S}(\beta_m)$ is the first derivative of Eq (4) and $m=0,1,2,\dots$ are the iterations which are carried out until convergence and the second derivative of Eq (4) with respect to β_k becomes a Hessian matrix as

$$\begin{aligned}\frac{\partial^2 l(\beta)}{\partial\beta_j\beta_k}&=\sum_{i=1}^{n}\varphi^{-1}\left(\frac{\partial l(\beta)}{\partial\beta_k}\right)\left(\frac{y_i-\mu_i}{Var(\mu_i)}\frac{\partial\mu_i}{\partial\eta_i}x_{ij}\right)\\&=\sum_{i=1}^{n}\varphi^{-1}\left[\left\{Var(\mu_i)^{-1}\left(\frac{\partial\mu_i}{\partial\eta_i}\right)^2\right\}\right.\\&-(\mu_i-y_i)\\&\left.\times\left\{Var(\mu_i)^{-2}\left(\frac{\partial\mu_i}{\partial\eta_i}\right)^2\frac{\partial Var(\mu_i)}{\partial\mu_i}-Var(\mu_i)^{-1}\left(\frac{\partial^2 Var(\mu_i)}{\partial\eta_i^2}\right)\right\}\right]x_{ji}x_{ki}.\end{aligned} \tag{7}$$

Both $\mathcal{S}(\beta_m)$ and $\frac{\partial^2 l(\beta)}{\partial\beta_j\beta_k}(\beta_m)$ are computed at β_m. On some simplifications, the final form of estimation algorithm is then linked to as iterative reweighted least squares as

$$\beta_m=(X^t\mathcal{V}_mX)^{-1}X^t\mathcal{V}_m z_m, \tag{8}$$

where $\mathcal{V}_m=diag[\{g(\mu_i)\}^2Var(\mu_i)\varphi]^{-1}$, $g(.)$ is the logit link function and $z_m=\log(\mu_i)+\frac{y_i-\mu_i}{\mu_i(1-\mu_i)}$. However, we conclude that β_m converge to $\hat{\beta}_{MLE}$ as $m\to\infty$, so final form of the MLE, we obtain

$$\hat{\beta}_{MLE}=(X^t\mathcal{V}X)^{-1}X^t\mathcal{V}z, \tag{9}$$

where $\mathcal{V}$ and z are estimated at the final iteration. Equations (8) and (9) is a feasible estimator for the estimation of unknown regression parameters, for more details please see Roozbeh and Arashi [42]. Both $\mathcal{V}$ and z are evaluated by using Fisher scoring iterative procedure. One can find the MMSE and MSE by considering $\alpha=\xi^t\hat{\beta}_{MLE}$ and $\Lambda=diag(\lambda_1,\lambda_2,\dots,\lambda_p)$ which is further equivalent to $\xi(X^t\mathcal{V}X)\xi^t$, where ξ represents the orthogonal matrix whose columns are the eigenvectors of $X^t\mathcal{V}X$; i.e., $\xi=(\xi_1,\dots,\xi_p)$, where ξ_j is the jth eigenvectors of $X^t\mathcal{V}X$ and $\lambda_1\geq\lambda_2\geq,\dots,\lambda_p\geq 0$ are the eigenvalues of the matrix $X^t\mathcal{V}X$whereas $\alpha_j\ \forall\ j=1,\dots,p$ is the jth element of $\xi^t\hat{\beta}_{MLE}$. Then, the covariance and MMSE of the $\hat{\beta}_{MLE}$ are respectively defined by

$$Cov(\hat{\beta}_{MLE}) = \hat{\delta}(X^t \mathcal{V} X)^{-1},$$

$$MMSE(\hat{\beta}_{MLE}) = \xi \Lambda^{-1} \xi^t \tag{10}$$

Therefore, the scalar MSE of the $\hat{\beta}_{MLE}$ is given as

$$MSE(\hat{\beta}_{MLE}) = E(\hat{\beta}_{MLE} - \beta)^t (\hat{\beta}_{MLE} - \beta) = \delta[tr(\xi \Lambda^{-1} \xi^t)] = \hat{\delta} \sum_{j=1}^{p} \frac{1}{\lambda_j}, \tag{11}$$

where λ_j is the jth eigenvalue of $X^t \mathcal{V} X$. The matrix $X^t \mathcal{V} X$ is ill-conditioned when the regressors are correlated that leads to some eigenvalues being small, and the estimated MSE of MLE is inflated. The multicollinearity problem is a severe issue in applied research that leads to high variance, wider confidence interval and unstable parameter estimates. Ferrari and Cribari-Neto [49] proposed a BRM for modelling rates and proportions. The MLE estimates of the BRM, and it is complicated to draw an inference based on the MLE for the BRM in the existence of multicollinearity. To circumvent this issue, we introduce a shrinkage estimator in the next section.

2.2. A beta ridge regression estimator

Qasim et al. [31] proposed a beta ridge regression estimator (BRRE) which is the generalization of Hoerl and Kennard [2] and is defined as

$$\hat{\beta}_{BRRE} = R_k \, \hat{\beta}_{MLE}, \tag{12}$$

where $R_k = (X^t \mathcal{V} X + kI_p)^{-1}(X^t \mathcal{V} X), k\ (k > 0)$is the shrinkage ridge parameter whereas I_p is an identity matrix of order $p \times p$. If $k \to 0,$ then $\hat{\beta}_{BRRE} = \hat{\beta}_{MLE}$. The bias vector and covariance matrix of Eq (12) can be defined as

$$Bias(\hat{\beta}_{BRRE}) = -k\xi \Lambda_k^{-1} \beta. \tag{13}$$

$$Cov(\hat{\beta}_{BRRE}) = \hat{\delta} \xi \Lambda_k^{-1} \Lambda \Lambda_k^{-1} \xi^t. \tag{14}$$

$$\begin{aligned} MMSE(\hat{\beta}_{BRRE}) &= R_k (X^t \mathcal{V} X)^{-1} R_k^t + Bias(\hat{\beta}_{BRRE}) Bias(\hat{\beta}_{BRRE})^t \\ &= \hat{\delta} \xi \Lambda_k^{-1} \Lambda \Lambda_k^{-1} \xi^t + k^2 \xi \Lambda_k^{-1} \beta \beta^t \Lambda_k^{-1} \xi^t. \end{aligned} \tag{15}$$

where $\Lambda_k = diag(\lambda_1 + k, \lambda_2 + k, \dots, \lambda_p + k)$ and$\Lambda = diag(\lambda_1, \lambda_2, \dots, \lambda_p) = \xi (X^t \mathcal{V} X) \xi^t$, where ξ is the orthogonal matrix whose columns are the eigenvectors of $X^t \mathcal{V} X$. Hence, the scalar MSE of the BRRE is generally obtained by applying the $tr\ (.)$ operator on Eq (15), which can be defined as

$$MSE(\hat{\beta}_{BRRE}) = tr\{MMSE(\hat{\beta}_{BRRE})\} = \hat{\delta}\sum_{j=1}^{p}\frac{\lambda_j}{(\lambda_j+k)^2} + k^2\sum_{j=1}^{p}\frac{\alpha_j^2}{(\lambda_j+k)^2}, \tag{16}$$

where $\alpha = \xi^t\hat{\beta}_{MLE}$ and k (k > 0) is the ridge parameter suggested by Hoerl and Kennard [14] and for the BRRE it is computed by taking the derivative of Eq (16) with respect to k and equate to zero, we have

$$k = \frac{\hat{\delta}}{\sum_{j=1}^{p}\hat{\alpha}_j^2}. \tag{17}$$

2.3. A beta Liu regression estimator

Karlsson et al. [47] introduced another estimation method called beta Liu estimator (BLE) for the BRM as

$$\hat{\beta}_{BLE} = L_d\,\hat{\beta}_{MLE}, \tag{18}$$

where $L_d = (X^t\mathcal{V}X + I)^{-1}(X^t\mathcal{V}X + dI)$ and d [0,1] is the Liu parameter. If $d \to 1$, then $\hat{\beta}_{BRR} = \hat{\beta}_{MLE}$. The bias vector and covariance matrix of Eq (18) can be respectively defined as

$$Bias(\hat{\beta}_{BLE}) = \xi(d-1)\Lambda_I^{-1}\beta. \tag{19}$$

$$Cov(\hat{\beta}_{BLE}) = \hat{\delta}\xi\Lambda_I^{-1}\Lambda_d\Lambda^{-1}\Lambda_d\Lambda_I^{-1}\xi^t. \tag{20}$$

Thus, the MMSE and scalar MSE of the BLE are respectively given as

$$\begin{aligned} MMSE(\hat{\beta}_{BLE}) &= L_d(X^t\mathcal{V}X)^{-1}L_d^t + Bias(\hat{\beta}_{BLE})Bias(\hat{\beta}_{BLE})^t \\ &= \hat{\delta}\xi\Lambda_I^{-1}\Lambda_d\Lambda^{-1}\Lambda_d\Lambda_I^{-1}\xi^t + (d-1)^2\xi\Lambda_I^{-1}\beta\beta^t\Lambda_I^{-1}\xi^t, \end{aligned} \tag{21}$$

where $\Lambda_I = diag(\lambda_1 + I, \lambda_2 + I, \dots, \lambda_p + I)$, $\Lambda_d = diag(\lambda_1 + d, \lambda_2 + d, \dots, \lambda_p + d)$. Finally, the scalar MSE of the BLE can be defined as

$$MSE(\hat{\beta}_{BLE}) = tr\{MMSE(\hat{\beta}_{BLE})\} = \hat{\delta}\sum_{j=1}^{p}\frac{(\lambda_j+d)^2}{\lambda_j(\lambda_j+1)^2} + (d-1)^2\sum_{j=1}^{p}\frac{\alpha_j^2}{(\lambda_j+1)^2}, \tag{22}$$

The optimum value of the above expression can be obtained by taking the derivative of Eq (22) with respect to d and setting the entire expression equal to zero. We obtain

$$d = \frac{\sum_{j=1}^{p} \frac{\alpha_j^2 - \hat{\delta}}{(\lambda_j + 1)^2}}{\sum_{j=1}^{p} \frac{\hat{\delta} + \lambda_j \alpha_j^2}{\lambda_j (\lambda_j + 1)^2}}. \tag{23}$$

2.4. Construction of the proposed estimator

For the LRM, Lukman et al. [3] proposed the following MRTE by augmenting the RRE which is given as

$$\hat{\beta}_{MRTE}(k,d) = \left(X^t X + k(1+d)I_p\right)^{-1} X^t X \hat{\beta}_{OLS} = M_{kd} \hat{\beta}_{OLS}, \tag{24}$$

where $M_{kd} = \left(X^t X + k(1+d)I_p\right)^{-1} X^t X$, while $\hat{\beta}_{OLS} = (X^t X)^{-1} X^t Y$.

In this study, we propose a MRTE for the BRM called MBRT estimator (MBRTE) defined as

$$\hat{\beta}_{\mathrm{MBRTE}}(k,d) = \left(X^t \mathcal{V} X + k(1+d)I_p\right)^{-1} X^t \mathcal{V} X \hat{\beta}_{MLE} = \widetilde{M}_{kd} \hat{\beta}_{MLE}, \tag{25}$$

where $\widetilde{M}_{kd} = \left(X^t \mathcal{V} X + k(1+d)I_p\right)^{-1} X^t \mathcal{V} X$, k and d are the two biasing parameters of the MBRTE whereas I_p is the identity matrix of order $p \times p$. The bias vector and covariance matrix of MBRTE are respectively given as

$$Bias\left[\hat{\beta}_{\mathrm{MBRTE}}(k,d)\right] = \xi\left(\widetilde{M}_{kd} - I\right)\beta. \tag{26}$$

$$Cov\left[\hat{\beta}_{\mathrm{MBRTE}}(k,d)\right] = \hat{\delta}\xi \widetilde{M}_{kd} \Lambda \widetilde{M}_{kd}^t \xi^t. \tag{27}$$

Thus, the MMSE and scalar MSE of the MBRTE are respectively given as

$$MMSE\left[\hat{\beta}_{\mathrm{MBRTE}}(k,d)\right] = \hat{\delta}\xi \widetilde{M}_{kd} \Lambda \widetilde{M}_{kd}^t \xi^t + \left(\widetilde{M}_{kd} - I\right)\beta\beta^t\left(\widetilde{M}_{kd} - I\right)^t. \tag{28}$$

The scalar MSE of MBRTE can be written as

$$MSE\left[\hat{\beta}_{\mathrm{MBRTE}}(k,d)\right] = \hat{\delta} \sum_{j=1}^{p} \frac{\lambda_j}{\left[\lambda_j + k(1+d)\right]^2} + \sum_{j=1}^{p} \left[\frac{\lambda_j}{\{\lambda_j + k(1+d)\}} - 1\right]^2. \tag{29}$$

Or Eq (29) can be further reduced as

$$MSE\left[\hat{\beta}_{\mathrm{MBRTE}}(k,d)\right] = \hat{\delta} \sum_{j=1}^{p} \frac{\lambda_j}{\left[\lambda_j + k(1+d)\right]^2} + k^2(1+d)^2 \sum_{j=1}^{p} \frac{\alpha_j^2}{\left[\lambda_j + k(1+d)\right]^2}. \tag{30}$$

2.5. Theoretical comparisons based on MMSE and scalar MSE's

Lemma 2.1. Let M be a positive definite (pd) matrix, α be a vector of nonzero constants and c be a positive constant. Then $cM - \alpha\alpha^t > 0$ if and only if $\alpha^t M\alpha < c$ [13].
Lemma 2.2. Given two estimators of υ, $\hat{\upsilon}_1 = B_1 y$ and $\hat{\upsilon}_2 = B_2 y$. Suppose that $E = Cov(\hat{\upsilon}_1) - Cov(\hat{\upsilon}_2) > 0$, where $Cov(\hat{\upsilon}_j)$ represents the covariance matrices of $\hat{\upsilon}_j$ $(j = 1, 2)$. Therefore, $\nabla(\hat{\upsilon}_1 - \hat{\upsilon}_2) = MMSE(\hat{\upsilon}_1) - MMSE(\hat{\upsilon}_2) \geq 0$ if and only if $b_1^t(\vartheta E + b_1 b_1^t)^{-1} b_1 < 1$, where $MMSE(\hat{\upsilon}_j) = Cov(\hat{\upsilon}_1) + b_1 b_1^t$ represents the covariance matrix and the bias vector of $\hat{\upsilon}_j$, respectively [14].

Theorem 2.1. Under the BRM, consider $k > 0, 0 < d < 1$ then $MMSE(\hat{\beta}_{MLE}) - MMSE[\hat{\beta}_{MBRTE}(k, d)] > 0$ if and only if $\beta^t(\widetilde{M}_{kd} - I)^t[\hat{\delta}\{(X^t\mathcal{V}X)^{-1} - \widetilde{M}_{kd}(X^t\mathcal{V}X)^{-1}\widetilde{M}_{kd}^t\}]^{-1}(\widetilde{M}_{kd} - I)\beta < 1$.

Proof. The proof of Theorem 2.1 can be seen in Appendix A.

Theorem 2.2. Under the BRM, consider $k > 0, 0 < d < 1$ then $MMSE(\hat{\beta}_{BRRE}) - MMSE[\hat{\beta}_{MBRTE}(k, d)] > 0$ if and only if $\beta^t(\widetilde{M}_{kd} - I)^t[\hat{\delta}\{F_k(X^t\mathcal{V}X)^{-1}F_k^t - \widetilde{M}_{kd}(X^t\mathcal{V}X)^{-1}\widetilde{M}_{kd}^t\}]^{-1}(\widetilde{M}_{kd} - I)\beta \leq 1$.

Proof. The proof of Theorem 2.2 can be seen in Appendix B.

Theorem 2.3. Under the BRM, consider $k > 0, 0 < d < 1$ then $MMSE(\hat{\beta}_{BLE}) - MMSE[\hat{\beta}_{MBRTE}(k, d)] > 0$ if and only if $\beta^t(\widetilde{M}_{kd} - I)^t[\hat{\delta}\{L_d(X^t\mathcal{V}X)^{-1}L_d^t - \widetilde{M}_{kd}(X^t\mathcal{V}X)^{-1}\widetilde{M}_{kd}^t\}]^{-1}(\widetilde{M}_{kd} - I)\beta \leq 1$.

Proof. The proof of Theorem 2.3 can be seen in Appendix C.

2.6. Selection of the biasing parameters

Following the suggestion of the RRE and many other estimators proposed by different authors such as Hoerl and Kennard [2], and Qasim et al. [34]. There is a need to find an appropriate parameter for a practical purpose. The optimal values of k and d are determined for the proposed estimator. In determining the optimal value of k, d is fixed. The optimal value of k can be selected by differentiating Eq (30) and equate to zero, we have

$$\frac{\partial MSE\left[\hat{\beta}_{MBRTE}(k,d)\right]}{\partial k} = -2\hat{\delta}(1+d)\sum_{j=1}^{p}\frac{\lambda_j}{\left(\lambda_j+k(1+d)\right)^3} - 2k^2(1+d)^3\sum_{j=1}^{p}\frac{\alpha_j^2}{\left(\lambda_j+k(1+d)\right)^3} + 2k(1+d)^2\sum_{j=1}^{p}\frac{\alpha_j^2}{\left(\lambda_j+k(1+d)\right)^2}. \quad (31)$$

Let $\frac{\partial MSE\left[\hat{\beta}_{MBRTE}(k,d)\right]}{\partial k} = 0$. On simplifying, it becomes

$$k_{opt} = \frac{\hat{\delta}}{(1+d^*)\alpha_j^2}. \quad (32)$$

For practical purpose, ϑ and α_j^2 are replaced with their unbiased estimates i.e., $\hat{\delta}$ and $\hat{\alpha}_j^2$, respectively. Consequently, Eq (32) becomes

$$\hat{k}_{opt} = \frac{\hat{\delta}}{(1+d^*)\hat{\alpha}_j^2}. \quad (33)$$

Furthermore, the optimal value for d can be derived by differentiating Eq (29) with respect to d for fixed k, we get

$$\frac{\partial MSE\left[\hat{\beta}_{MBRTE}(k,d)\right]}{\partial d} = -2\hat{\delta}k\sum_{j=1}^{p}\frac{\lambda_j}{\left(\lambda_j+k(1+d)\right)^3} - 2k^2(1+d)\sum_{j=1}^{p}\frac{\alpha_j^2}{\left(\lambda_j+k(1+d)\right)^2} + 2k^3(1+d)^2\sum_{j=1}^{p}\frac{\alpha_j^2}{\left(\lambda_j+k(1+d)\right)^3}. \quad (34)$$

Let $\frac{\partial MSE\left[\hat{\beta}_{MBRTE}(k,d)\right]}{\partial d} = 0$. On simplifying, it becomes

$$d_{opt} = \frac{\delta}{k^*\alpha_j^2} - 1. \quad (35)$$

For practical purpose, δ and α_j^2 are replaced with their unbiased estimates i.e., $\hat{\delta}$ and $\hat{\alpha}_j^2$, respectively. Consequently, Eq (35) becomes

$$\hat{d}_{opt} = \frac{\hat{\delta}}{k^{*}\hat{\alpha}_j^2} - 1. \quad (36)$$

The selection of the estimators of the parameters d and k is obtained iteratively as follows:

Step 1: Obtain an initial estimate of d using $\hat{d} = \frac{\sum_{j=1}^{p}\frac{\hat{\alpha}_j^2-\hat{\delta}}{(\lambda_j+1)^2}}{\sum_{j=1}^{p}\frac{\hat{\delta}+\lambda_j\hat{\alpha}_j^2}{\lambda_j(\lambda_j+1)^2}}$.

Step 2: Obtain $\hat{k}_{opt}$ using Eq (33) by based on $\hat{d}$ as computed in step 1.

Step 3: Estimate $\hat{d}_{opt}$ using Eq (36) by utilizing the value of $\hat{k}_{opt}$ obtained from step 2.

Step 4: If the value obtained from Step 1 is not between 0 and 1, use $\hat{d}_{opt} = \hat{d}$.

3. Results Monte Carlo simulation

In this section, a brief discussion is presented about the generation of data associated with different factors that play a pivotal role in the design of the simulation experiment. In addition, the assessment criterion is also presented to investigate the performance of the proposed estimator and compared with other competitive estimators.

3.1. Simulation layout

The numerical results were obtained using the following BRM which is defined as

$$log\left(\frac{\mu_i}{1-\mu_i}\right) = \left(\beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip}\right)^{-1}, i = 1, \ldots, n; j = 1, \ldots, p, \quad (37)$$

where x_{ij} is the covariates and β_j represents the regression coefficients. The restriction on slope parameter values is $\sum_{j=1}^{p}\beta_j^2 = 1$, (see, e.g., Kibria [9] for more details). The intercept value is considered to be zero. The BRM is defined by using the logit link function in Eq (36). The n observations for the response variable are generated as $y_i \sim \beta_e(\mu_i, \varphi), i = 1, \ldots, n$ in each Monte Carlo replication. The correlated regressors are generated by following McDonald and Galarneau [15] as follows

$$x_{ij} = (1-\rho^2)^{1/2} z_{ij} + \rho z_{i(j+1)}, i = 1, \ldots, n; j = 1, \ldots, p. \quad (38)$$

where z_{ij} is the independent standard normal pseudo-random numbers and ρ is the degree of correlation among two regressors. The performance of the proposed estimators is assessed under different conditions such as the degree of correlation which is taken to be $\rho = 0.80, 0.90, 0.95, 0.99$, sample size $(n = 25, 50, 100, 200)$, number of explanatory variables $(p = 4, 8, 16)$, and dispersion parameter $(\varphi = 0.01, 0.10, 10, 20)$. For a combination of the various values of n, p, ρ, φ, the generated data are repeated 2000 times and the average MSE are respectively defined as

$$MSE(\hat{\beta}) = \frac{\sum_{i=1}^{R}(\hat{\beta}_i - \beta)^t(\hat{\beta}_i - \beta)}{R}, \tag{39}$$

where $(\hat{\beta}_i - \beta)$ is the difference between the estimated and true parameter vectors and R is the number replications. All the computations are performed using the R Programming Language with the support of betareg() R package.

3.2. Results and discussion

The estimated MSEs of the BRM with different estimation methods are reported in Tables 1–3. For the evaluation purposes, we consider different factors to notice the performance of the proposed MBRTE. The general comments on the findings of the simulation study are discussed below.

From the results reported in Tables 1–3, one can be noticed that the performance of the proposed MBRTE is quite satisfactory in a sense of minimum MSE as compared to other estimation methods under study.

Results also demonstrate that multicollinearity has a severe impact on the estimated MSE's of the estimators under study. While this impact is somehow lower in our proposed estimator which signifies that MBRT shows a robust behaviour in the presence of high but imperfect multicollinearity. It should be noted that the MLE is the most negatively affected estimator when the regressors are correlated with one another. In addition, our proposed estimator outperforms the traditional MLE, BRRE, BLE in the BRM to reduce the effect of collinearity among regressors.

Increasing the value of sample size (n) makes a decrease in the MSE values of the BRM for all the estimators under study. Meanwhile, the sample size has an indirect impact on the estimated MSE's.

Increasing the number of explanatory variables makes an increase in the simulated MSE values of the BRM estimators. Again, the MLE is the severely negatively affected estimator in this situation. If we examine the performance of the estimators concerning the regressors, then we conclude from the results that the proposed MBRTE still shows a reliable estimation method as compared to other estimation methods.

Hence, the findings of the simulation is compatible with the theoretical results. In summary, we suggest using MBRTE in the presence of moderate to strong multicollinearity due to its significantly decreasing the simulated MSE's.

Table 1. Estimated MSE values when $p = 4$.

ρ	n	MLE	BRRE	BLE	MBRTE	MLE	BRRE	BLE	MBRTE
		When $\varphi = 0.01$				**When $\varphi = 0.10$**			
0.8	25	1461.865	954.685	634.852	0.294	1510.681	956.411	653.440	0.454
	50	144.960	93.578	62.518	0.243	158.128	96.121	65.545	0.379
	100	29.752	19.903	13.375	0.207	29.763	19.578	13.202	0.363
	200	15.530	10.685	7.241	0.195	15.483	10.592	7.080	0.292
0.9	25	1538.564	958.662	680.459	0.320	1597.525	978.733	657.294	0.316

Continued on next page

ρ	n	MLE	BRRE	BLE	MBRTE	MLE	BRRE	BLE	MBRTE
		When $\varphi = 0.01$				**When $\varphi = 0.10$**			
	50	151.126	96.773	65.062	0.256	159.851	105.768	71.996	0.298
	100	32.122	20.981	13.979	0.214	30.887	20.033	13.430	0.285
	200	15.903	10.774	7.364	0.199	15.506	10.751	7.123	0.284
0.95	25	1548.607	1002.796	675.240	0.350	1506.279	1002.934	706.873	0.258
	50	156.657	102.254	70.254	0.276	151.471	109.378	72.372	0.244
	100	32.444	21.292	14.297	0.225	30.028	21.152	13.858	0.222
	200	16.456	11.171	7.485	0.206	16.846	11.175	7.445	0.206
0.99	25	1601.282	1035.429	688.990	0.470	1661.684	1083.714	713.653	0.202
	50	167.473	106.368	70.490	0.370	177.806	111.048	77.710	0.202
	100	32.975	22.498	14.840	0.304	34.010	21.314	14.517	0.201
	200	17.264	11.591	7.733	0.266	17.254	11.584	7.952	0.201
		When $\varphi = 10$				**When $\varphi = 20$**			
0.8	25	1520.454	974.694	665.840	0.295	1525.708	955.363	665.220	0.298
	50	150.927	96.626	64.800	0.238	147.848	95.274	65.179	0.244
	100	30.169	20.328	13.809	0.200	31.129	20.866	13.810	0.203
	200	15.285	10.390	6.989	0.194	15.476	10.594	7.111	0.199
0.9	25	1544.126	1003.403	676.376	0.330	1534.105	982.868	674.019	0.319
	50	155.875	102.261	69.402	0.259	156.330	98.842	68.386	0.252
	100	30.985	20.808	13.914	0.209	31.314	21.217	14.159	0.213
	200	16.451	11.177	7.527	0.197	16.013	10.998	7.294	0.201
0.95	25	1553.573	1008.130	682.785	0.352	1595.336	997.089	676.624	0.356
	50	157.302	103.901	70.245	0.286	160.072	102.678	69.935	0.283
	100	31.053	21.161	14.154	0.217	31.402	21.370	14.418	0.227
	200	16.689	11.288	7.620	0.212	16.136	11.052	7.473	0.208
0.99	25	1637.592	1026.718	692.474	0.493	1613.593	1083.808	713.157	0.458
	50	166.097	106.411	70.685	0.378	162.618	109.395	72.122	0.386
	100	34.318	21.937	14.985	0.289	35.886	22.973	15.594	0.301
	200	17.431	11.777	7.978	0.273	17.829	11.508	7.637	0.268

Table 2. Estimated MSE values when $p = 8$.

ρ	n	MLE	BRRE	BLE	MBRTE	MLE	BRRE	BLE	MBRTE
		When $\varphi = 0.01$				***When*** $\varphi = 0.10$			
0.8	25	3493.471	1655.005	2475.141	0.170	3410.860	1602.835	2419.461	0.173
	50	336.830	155.522	237.711	0.155	340.091	158.464	241.797	0.155
	100	69.429	32.637	50.238	0.135	70.570	33.249	51.332	0.139
	200	36.001	17.116	26.659	0.132	35.350	16.662	26.028	0.138
0.9	25	3787.412	1716.299	2453.250	0.178	3684.274	1682.698	2513.316	0.179
	50	381.654	178.889	256.158	0.157	378.671	177.312	262.299	0.156
	100	75.838	35.292	53.452	0.140	78.268	37.173	56.124	0.141
	200	38.895	18.335	28.095	0.137	38.254	17.819	27.304	0.133
0.95	25	3940.006	1744.503	2575.470	0.228	4156.820	1857.568	2625.028	0.237
	50	399.610	180.186	265.770	0.191	407.611	184.123	263.323	0.195
	100	82.505	37.450	55.708	0.156	83.855	38.087	56.724	0.154
	200	43.203	19.902	30.391	0.155	42.470	19.486	29.652	0.151
0.99	25	6353.596	2516.975	3323.980	0.588	6134.180	2409.218	3183.864	0.712
	50	627.016	244.927	334.904	0.552	632.958	254.474	342.081	0.523
	100	126.134	49.502	72.028	0.464	127.439	49.602	72.580	0.456
	200	63.745	25.255	38.347	0.449	66.473	27.221	40.485	0.405
		When $\varphi = 10$				***When*** $\varphi = 20$			
0.8	25	3584.167	1719.952	2580.859	0.172	3518.382	1675.885	2513.824	0.177
	50	339.155	156.971	240.639	0.150	344.651	162.652	245.150	0.151
	100	68.101	31.617	49.133	0.135	71.344	34.410	52.359	0.135
	200	36.056	17.139	26.568	0.134	35.487	16.876	26.128	0.132
0.9	25	3775.467	1760.573	2585.961	0.181	3723.077	1712.880	2537.605	0.177
	50	368.266	169.179	251.905	0.160	385.446	182.013	259.664	0.155
	100	75.468	35.394	53.593	0.142	75.382	34.689	53.250	0.142
	200	39.241	18.606	28.485	0.129	39.040	18.436	28.159	0.139
0.95	25	4072.415	1849.218	2587.453	0.223	4114.300	1856.151	2622.516	0.233
	50	412.676	189.978	267.416	0.204	405.125	179.866	267.277	0.187
	100	84.916	39.503	57.971	0.157	82.663	37.725	55.966	0.154
	200	41.293	18.640	28.558	0.154	42.185	19.505	29.484	0.143
0.99	25	6223.062	2357.773	3200.499	0.636	6332.104	2503.887	3315.646	0.634
	50	627.069	244.600	334.676	0.607	626.370	244.832	333.947	0.517
	100	126.896	49.587	72.767	0.421	125.758	49.557	72.559	0.445
	200	64.634	25.171	38.447	0.384	65.828	25.831	39.293	0.426

Table 3. Estimated MSE values when $p = 16$.

ρ	n	MLE	BRRE	BLE	MBRTE	MLE	BRRE	BLE	MBRTE
		When $\boldsymbol{\varphi = 0.01}$				***When*** $\boldsymbol{\varphi = 0.10}$			
0.8	25	7709.10	5772.19	3826.34	0.116	7608.27	5657.97	3705.16	0.116
	50	756.45	561.99	365.46	0.104	758.58	566.37	367.00	0.110
	100	154.40	117.41	74.85	0.096	157.24	120.65	77.56	0.095
	200	78.58	60.52	37.83	0.099	77.89	60.07	37.45	0.093
0.9	25	8798.98	5955.73	4080.84	0.120	9155.90	6188.07	4266.57	0.118
	50	906.99	624.09	425.79	0.114	928.01	642.18	442.26	0.112
	100	183.29	130.32	85.38	0.105	188.26	134.94	88.76	0.108
	200	97.99	72.53	46.59	0.103	95.50	70.18	44.86	0.106
0.95	25	11053.46	6186.55	4677.31	0.318	11239.23	6421.11	4834.07	0.344
	50	1110.72	651.09	465.55	0.297	1140.54	671.61	485.58	0.288
	100	234.59	148.34	100.19	0.238	227.69	143.23	96.69	0.253
	200	115.19	75.06	48.13	0.241	117.76	77.61	50.45	0.253
0.99	25	77507.03	12216.19	14917.40	12.075	78376.06	11677.04	14778.76	10.081
	50	7725.77	1391.07	1607.66	10.040	7986.19	1400.34	1598.89	9.494
	100	1691.96	373.89	360.84	7.462	1560.49	317.31	333.73	7.677
	200	825.70	188.55	161.77	7.367	805.44	191.74	166.76	7.460
		When $\boldsymbol{\varphi = 10}$				***When*** $\boldsymbol{\varphi = 20}$			
0.8	25	7414.81	5490.12	3594.12	0.117	7551.93	3671.15	5588.18	0.115
	50	762.57	569.52	370.52	0.106	758.88	368.35	568.66	0.105
	100	152.59	115.52	73.31	0.094	158.55	78.16	121.62	0.092
	200	78.56	60.52	37.98	0.091	79.84	39.13	61.99	0.094
0.9	25	9095.27	6207.31	4304.71	0.120	9219.84	4391.77	6308.37	0.117
	50	907.03	622.82	421.97	0.113	913.80	430.76	628.25	0.109
	100	186.21	133.97	87.64	0.107	185.75	88.17	133.14	0.108
	200	96.11	70.92	45.28	0.106	96.56	45.25	70.95	0.103
0.95	25	11205.91	6328.50	4709.50	0.343	11204.95	4706.73	6293.87	0.352
	50	1089.65	627.50	455.02	0.297	1082.86	454.36	631.05	0.268
	100	227.05	142.40	96.40	0.235	226.47	95.13	142.05	0.248
	200	117.53	76.64	49.31	0.233	117.90	50.20	77.52	0.244
0.99	25	77080.04	12297.92	14439.78	8.513	74840.34	14994.86	12475.33	8.137
	50	7937.47	1569.76	1629.59	7.969	7850.99	1608.74	1414.84	7.723
	100	1595.81	347.58	341.86	7.823	1665.56	338.02	348.46	7.639
	200	814.78	190.18	172.79	7.759	817.89	167.14	186.33	6.533

4. Application: Heat treating test data

To evaluate the performance of the proposed method, we consider the heat treating test data. This dataset is taken from Montgomery and Runger [56]. This dataset consists of one response and five explanatory variables, where the response variable (y) represents the PITCH that denotes the product introduction to customer heart whose meaning is the quality of a sound governed by the rate of vibrations producing it or the level of something. Whereas the description of regressors include: x_1= furnace temperature (Temp), x_2=carbon concentration x_3=carburizing cycle (soakpct, soaktime), x_4=carbon concentration and x_5=duration of the defuse time (Difftime, Diffpct). As the response variable is in the form of ratio, so we use the BRM to model the response variable. Initially, it is necessary to examine the multicollinearity issue among the regressors. For this, condition index is generally used i.e., $CI = \sqrt{\frac{\max(\lambda_j)}{\min(\lambda_j)}}$ which is to be 68621.13 indicating the existence of severe multicollinearity among the regressors [1,43].

Table 4. Estimated Coefficients and MSE's of the considered example.

	MLE		BRRE		BLE		MBRTE	
	Estimates	SE	Estimates	SE	Estimates	SE	Estimates	SE
β_1	0.00197	0.00162	0.00198	0.00023	0.00209	0.00023	0.00306	0.00022
β_2	0.01967	0.00167	0.01963	0.00040	0.01912	0.00040	0.01493	0.00035
β_3	0.00148	0.00163	0.00145	0.00066	0.00118	0.00064	-0.00112	0.00047
β_4	0.00833	0.00144	0.00836	0.00092	0.00840	0.00087	0.00896	0.00051
β_5	-0.00335	0.00862	-0.00318	0.00333	-0.00310	0.00320	-0.00094	0.00027
MSE	0.000085	-----	0.000076	-----	0.000070	-----	0.000041	-----
CV	0.97642	-----	0.95188	-----	0.94874	-----	0.94873	-----

*Note: SE = Standard error; CV = Cross validation.

The estimated coefficients and their respective MSE's of the BRM's estimators are reported in Table 4. Moreover, the estimated coefficients of the MLE, BRRE, BLE, and MBRTE are computed respectively using Eqs (9), (12), (18) and (25). While the MSE's of the considered estimators are computed respectively using Eqs (11), (16), (22) and (30). On the contrary, the value of ridge parameter used in scalar MSE of the BRRE is computed using Eq (17) which is to found be $\hat{k} = 0.0044$, Liu parameter of the BLE is calculated using Eq (23) which is found to be $\hat{d} = 0.8948$, and the values of the biasing parameters for the MBRTE are computed respectively using Eqs (33) and (36) and are found to be $\hat{k} = 0.0033$ and $\hat{d} = 0.5135$. The results of Table 4 signifies that the MSE of the proposed MBRTE is substantially smaller than the MLE as well as other biased estimators. Results also validate that the MLE is the most negatively affected estimator when the regressors are highly correlated with one another due to its inflated behaviour of the MSE. However, BRRE and BLE show a smaller MSE when compared with MLE but these two biased estimators are not reliable when compared with our proposed MBRTE. Further, we also consider the standard error (SE) of the estimators considered in the study. Table 4 clearly demonstrates that the proposed MBRTE attains the smaller SEs while MLE attains the larger SEs of the regression estimates. Further, it is also clear that due to large SE and MSE, the MLE is not a better estimation method.

We use another criteria i.e., cross validation (CV) applied to the real life data set for the assessment of the proposed method. The findings of average validation error with reference to CV method are also given in Table 4. In CV criteria, we divide the train set into K = 5 equal size subsets or folds. Now for the individual value of k i.e., $k = 1, \ldots, 5$, we fit our prediction function on all points but those I the kth fold, and evaluate the validation error on the points in the kth fold. For more details on the CV, we refer the readers, please see [26, 43–45]. Further, the CV is considered to examine the predictive performance of the estimators comprehensively. Table 4 represents the CV values of all the estimators considered in the study. Results signifies that the performance of the proposed MBRTE is better as compared to the MLE, BRRE, and BLE. So both criteria i.e., MSE and CV shows that the proposed estimator performs consistently better as compared to the competitors.

Hence, both the simulation results and empirical application findings consistently support our proposed estimator and we recommend the practitioners to use our MBRTE instead of BRRE and BLE for estimating the unknown parameters of the BRM whenever the regressors are highly multicollinear.

5. Conclusions

In this paper, we proposed a new modified beta ridge-type (MBRT) estimator. We show analytically that the new estimator is superior to the standard MLE as well as other well-known biased estimators, i.e., BRRE, and BLE. A simulation study has been conducted to compare the performance of our proposed estimator with the available estimators. In the simulation experiment, we consider several factors to monitor the behaviour of the proposed estimator in every perspective. Based on the findings of the simulation, we found that our proposed estimator performed better than some existing estimators in a sense of smaller MSE and can be recommended. Finally, the benefit of the proposed estimator is shown by an empirical application where the proposed MBRTE for the BRM performed considerably better in terms of smaller MSE and CV as compared to the MLE as well as the other biased estimators. So, we suggest the researchers to use MBRTE whenever they want to fit the BRM and face the issue of multicollinearity.

Acknowledgments

We would like to thank the editor and the referees for their suggestions that improve the first edition of the paper.

Conflict of interest

There is no conflict of interests.

References

1. M. N. Akram, M. Amin, M. Amanullah, Two-parameter estimator for the inverse Gaussian regression model, *Commun. Stat. Simul. C.*, 2020. doi: 10.1080/03610918.2020. 1797797.
2. A. E. Hoerl, R. W. Kennard, Ridge regression: Biased estimation for nonorthogonal problems, *Technometrics*, **12** (1970), 55–67.
3. A. F. Lukman, K. Ayinde, S. Binuomote, O. A. Clement, Modified ridge-type estimator to combat multicollinearity: Application to chemical data, *J. Chemometr.*, **33** (2019), e3125. doi: 10.1002/cem.3125.
4. A. F. Lukman, A. Emmanuel, O. A. Clement, K. Ayinde, A Modified Ridge-Type Logistic Estimator, *Iran. J. Sci. Technol. Trans. Sci.*, **44** (2020), 437–443. doi: 10.1007/s40995-020-00845-z.
5. A. J. Lemonte, S. L. P. Ferrari, F. Cribari-Neto, Improved likelihood inference in Birnbaum-saunders regressions, *Comput. Stat. Data An.*, **54** (2010), 1307–131. doi: 10.1016/j.csda.2009.11.017.
6. B. F. Swindel. Good ridge estimators based on prior information, *Commun. Stat. Theor. M.*, **5** (1976), 1065–1075. doi: 10.1080/03610927608827423.
7. B. Singh, Y. P. Chaubey, On some improved ridge estimators, *Statistische Hefte*, **28** (1987), 53–67. doi: 10.1007/BF02932590.
8. B. Segerstedt, On ordinary ridge regression in generalized linear models, *Commun. Stat. Theor. M.*, **21** (1992), 2227–2246. doi: 10.1080/03610929208830909.
9. B. M. G. Kibria, Performance of some new ridge regression estimators, *Commun. Stat. Simul. C.*, **32** (2003), 419–435. doi: 10.1081/SAC-120017499.
10. B. M. G. Kibria, Some Liu and ridge-type estimators and their properties under the ill-conditioned Gaussian linear regression model, *J. Stat. Comput. Sim.*, **82** (2012), 1–17. doi: 10.1080/00949655.2010.519705.
11. C. Stein, Inadmissibility of the usual estimator for the mean of a multivariate normal distribution, In: *Volume 1 contribution to the theory of statistics*, Berkeley: University of California Press, 1956, 197–206. doi: 10.1525/9780520313880-018.
12. E. Vigneau, M. F. Devaux, E. M. Qannari, P. Robert, Principal component regression, ridge regression and ridge principal component regression in spectroscopy calibration, *J. Chemometr.*, **11** (1998), 239–249. doi: 10.1002/(SICI)1099-128X(199705)11:3<239::AID-CEM470>3.0.CO;2-A.
13. R. W. Farebrother, Further results on the mean square error of ridge regression, *J. R. Stat. Soc. B*, **38** (1976), 248–250.
14. G. Trenkler, H. Toutenburg, Mean squared error matrix comparisons between biased estimators-An overview of recent results, *Stat. Pap.*, **31** (1990), 165. doi: 10.1007/BF02924687.
15. G. C. McDonald, D. I. Galarneau, A Monte Carlo evaluation of some ridge-type estimators, *J. Am. Stat. Assoc.*, **70** (1975), 407–416.
16. K. Månsson, G. Shukur, On ridge parameters in logistic regression, *Commun. Stat. Theor. M.*, **40** (2011), 3366–3381. doi: 10.1080/03610926.2010.500111.
17. K. Månsson, G. Shukur, A Poisson ridge regression estimator, *Econ. Model.*, **28** (2011), 1475–1481. doi: 10.1016/j.econmod.2011.02.030.
18. K. Månsson, B. M. G. Kibria, P. Sjölander, G. Shukur, V. Sweden, *New Liu estimators for the poisson regression model: Model and application*, HUI Research, 2011.

19. K. Månsson, B. M. G. Kibria, G. Shukur, On Liu estimators for the logit regression model, *Econ. Model.*, **29** (2012), 1483–1488. doi: 10.1016/j.econmod.2011.11.015.
20. K. Månsson, Developing a Liu estimator for the negative binomial regression model: Method and application, *J. Stat. Comput. Sim.*, **83** (2013), 1773–1780. doi: 10.1080/00949655.2012.673127.
21. L. Kejian, A new class of biased estimate in linear regression, *Commun. Stat. Theor. M.*, **22** (1993), 393–402. doi:.1080/03610929308831027.
22. L. Kejian, Using Liu-type estimator to combat collinearity, *Commun. Stat. Theor. M.*, **32** (2003), 1009–1020. doi: 10.1081/STA-120019959.
23. L. S. Mayer, T. A. Willke, On biased estimation in linear models, *Technometrics*, **15** (1973), 497–508. doi: 10.1080/00401706.1973.10489076.
24. A. F. Lukman, K. Ayinde, Review and classifications of the ridge parameter estimation techniques, *Hacet. J. Math. Stat.*, **46** (2017), 953–967.
25. M. I. Alheety, B. M. G. Kibria, Modified Liu-type estimator based on (r-k) class estimator, *Commun. Stat. Theor. M.*, **42** (2013), 304–319. doi: 10.1080/03610926.2011.577552.
26. M. Amini, M. Roozbeh, Optimal partial ridge estimation in restricted semiparametric regression models, *J. Multivariate Anal.*, **136** (2015), 26–40. doi: 10.1016/j.jmva.2015.01.005.
27. M. Arashi, S. M. M. Tabatabaey, B. H. Bashtian, Shrinkage ridge estimators in linear regression, *Commun. Stat. Simul. C.*, **43** (2014), 871–904. doi: 10.1080/03610918.2012.718838.
28. M. Qasim, K. Månsson, M. Amin, B. M. G. Kibria, P. Sjolander, Biased adjusted Poisson ridge estimators-method and application, *Iran. J. Sci. Technol. Tran. Sci.*, **44** (2020), 1775–1789. doi: 10.1007/s40995-020-00974-5.
29. M. Amin, M. N. Akram, M. Amanullah, On the James-Stein estimator for the poisson regression model, *Commun. Stat. Simul. C.*, 2020, 1–13. doi: 10.1080/03610918.2020.1775851.
30. M. Amin, M. Qasim, M. Amanullah, S. Afzal, Performance of some ridge estimators for the gamma regression model, *Stat. Pap.*, **61** (2020), 997–1026. doi: 10.1007/s00362-017-0971-z.
31. M. Qasim, K. Månsson, B. M. G. Kibria, On some beta ridge regression estimators: method, simulation and application, *J. Stat. Comput. Sim.,* **91** (2021), 1699–1712. doi: 10.1080/00949655.2020.1867549.
32. M. I. Alheety, B. M. G. Kibria, On the Liu and almost unbiased Liu estimators in the presence of multicollinearity with heteroscedastic or correlated errors, *Surv. Math. Appl.*, **4** (2009), 155–167.
33. M. N. Akram, M. Amin, M. Qasim, A new Liu-type estimator for the inverse Gaussian regression model, *J. Stat. Comput. Sim.*, **90** (2020), 1153–1172. doi: 10.1080/00949655.2020.1718150.
34. M. Qasim, M. Amin, M. Amanullah, On the performance of some new Liu parameters for the gamma regression model, *J. Stat. Comput. Sim.*, **88** (2018), 3065–3080. doi: 10.1080/00949655.2018.1498502.
35. M. Qasim, B. M. G. Kibria, K. Månsson, P. Sjölander, A new Poisson Liu regression estimator: method and application, *J. Appl. Stat.*, **47** (2020), 2258–2271. doi: 10.1080/02664763.2019.1707485.
36. M. Amin, M. A. Ullah, G. M. Cordeiro, Influence diagnostics in the gamma regression model with adjusted deviance residuals, *Commun. Stat. Simul. C.*, **46** (2017), 6959–6973. doi: 10.1080/03610918.2016.1222420.
37. M. Amin, M. A. Ullah, M. Aslam, Empirical evaluation of the inverse Gaussian regression residuals for the assessment of influential points, *J. Chemometr.*, **30** (2016), 394–404. doi: 10.1002/cem.2805.

38. M. Amin, M. Faisal, M. N. Akram, Influence diagnostics in the inverse gaussian ridge regression model: Applications in chemometrics, *J. Chemometr.*, ***35*** (2021), e3342. doi: 10.1002/cem.3342.
39. M. Amin, M. Qasim, S. Afzal, K. Naveed, New ridge estimators in the inverse Gaussian regression: Monte Carlo simulation and application to chemical data, *Commun. Stat. Simul. C.*, 2020. doi: 10.1080/03610918.2020.1797794.
40. M. Meloun, J. Militký, Detection of single influential points in OLS regression model building, *Anal. Chim. Acta*, **439** (2001), 169–191. doi: 10.1016/S0003-2670(01)01040-6.
41. G. Muniz, B. M. G. Kibria, On some ridge regression estimators: An empirical comparisons, *Commun. Stat. Simul. Comput.,* **38** (2009), 621–630. doi: 10.1080/03610910802592838.
42. M. Roozbeh, M. Arashi, Feasible ridge estimator in partially linear models, *J. Multivariate Anal.*, **116** (2013), 35–44. doi: 10.1016/j.jmva.2012.11.006.
43. M. Roozbeh, G. Hesamian, M. G. Akbari, Ridge estimation in semi-parametric regression models under the stochastic restriction and correlated elliptically contoured errors, *J. Comput. Appl. Math.*, **378** (2020), 112940. doi: 10.1016/j.cam.2020.112940.
44. M. Roozbeh, Optimal QR-based estimation in partially linear regression models with correlated errors using GCV criterion, *Comput. Stat. Data Anal.*, **117** (2018), 45–61. doi: 10.1016/j.csda.2017.08.002.
45. M. Roozbeh, M. Arashi, N. A. Hamzah, Generalized cross-validation for simultaneous optimization of tuning parameters in ridge regression, *Iran. J. Sci. Technol. Trans. Sci.*, **44** (2020), 473–485. doi: 10.1007/s40995-020-00851-1.
46. N. H. Prater, Estimate gasoline yields from crudes, *Petrol. Refiner*, **35** (1956), 236–238.
47. P. Karlsson, K. Månsson, B. M. G. Kibria, A Liu estimator for the beta regression model and its application to chemical data, *J. Chemometr.*, **34** (2020), e3300. doi: 10.1002/cem.3300.
48. R. Frisch, *Statistical confluence analysis by means of complete regression systems*, Universitetets Økonomiske Institute, 1934.
49. S. Ferrari, F. Cribari-Neto, Beta regression for modeling rates and proportions, *J. Appl. Stat.*, **31** (2004), 799–815. doi: 10.1080/0266476042000214501.
50. A. B. Simas, W. Barreto-Souza, A. V. Rocha, Improved estimators for a general class of beta regression models, *Comput. Stat. Data Anal.*, 54 (2010), 348–366. doi: 10.1016/j.csda.2009.08.017.
51. Y. Li, H. Yang, A new Liu-type estimator in linear regression model, *Stat. Pap.*, **53** (2012), 427–437. doi: 10.1007/s00362-010-0349-y.
52. Y. Li, H. Yang, A new stochastic mixed ridge estimator in linear regression model, *Stat. Pap.*, **51** (2010), 315–323. doi: 10.1007/S00362-008-0169-5.
53. Z. Y. Algamal, M. H. Lee, A. M. Al-Fakih, M. Aziz, High-dimensional QSAR prediction of anticancer potency of imidazo[4,5-b]pyridine derivatives using adjusted adaptive LASSO, *J. Chemometr.*, **29** (2015), 547–556. doi: 10.1002/cem.2741.
54. Z. Y. Algamal, M. K. Qasim, M. H. Lee, T. H. M. Ali, High-dimensional QSAR/QSPR classification modeling based on improving pigeon optimization algorithm, *Chemometr. Intell. Lab.*, **206** (2020), 104170. doi: 10.1016/j.chemolab.2020.104170.
55. Z. Y. Algamal, M. K. Qasim, M. H. Lee, T. H. M. Ali, Improving grasshopper optimization algorithm for hyper-parameters estimation and feature selection in support vector regression, *Chemometr. Intell. Lab.*, **208** (2020), 104196. doi: 10.1016/j.chemolab.2020.104196.

56. D. C. Montgomery, G. C. Runger, *Applied statistics and probability for engineers*, John Wiley & Sons, 2014.

Appendix A: Proof of Theorem 1

The difference among the MMSE functions of MLE and MBRTE is obtained by

$$\begin{aligned}\nabla &= MMSE(\hat{\beta}_{MLE}) - MMSE[\hat{\beta}_{MBRTE}(k,d)] \\ &= \hat{\delta}\left((X^t\mathcal{V}X)^{-1} - \widetilde{M}_{kd}(X^t\mathcal{V}X)^{-1}\widetilde{M}_{kd}^t\right) - (\widetilde{M}_{kd} - I)\beta\beta^t(\widetilde{M}_{kd} - I)^t.\end{aligned} \tag{A1}$$

However, Eq (A1) can be written in terms of scalar MSE as

$$\begin{aligned}&MSE(\hat{\beta}_{MLE}) - MSE[\hat{\beta}_{MBRTE}(k,d)] \\ &\qquad = \xi\hat{\delta}diag\left\{\frac{1}{\lambda_j} - \frac{\lambda_j}{[\lambda_j + k(1+d)]^2}\right\}_{j=1}^p \xi^t - b_{MBRTE}^t b_{MBRTE}.\end{aligned} \tag{A2}$$

$$= \xi\hat{\delta}diag\left\{\frac{[\lambda_j + k(1+d)]^2 - \lambda_j}{\lambda_j[\lambda_j + k(1+d)]^2}\right\}_{j=1}^p \xi^t - b_{MBRTE}^t b_{MBRTE}. \tag{A3}$$

Eq (A3) can be further simplified as

$$\begin{aligned}&MSE(\hat{\beta}_{MLE}) - MSE[\hat{\beta}_{MBRTE}(k,d)] \\ &\qquad = \xi\hat{\delta}diag\left\{\frac{[\lambda_j + k(1+d)]^2 - \lambda_j}{\lambda_j[\lambda_j + k(1+d)]^2}\right\}_{j=1}^p \xi^t - b_{MBRTE}^t b_{MBRTE} \\ &\qquad = \xi\hat{\delta}diag\left\{\frac{2\lambda_j k(1+d) + k^2(1+d+d^2)}{\lambda_j[\lambda_j + k(1+d)]^2}\right\}_{j=1}^p \xi^t - b_{MBRTE}^t b_{MBRTE}.\end{aligned} \tag{A4}$$

We observed that $(X^t\mathcal{V}X)^{-1} - \widetilde{M}_{kd}(X^t\mathcal{V}X)^{-1}\widetilde{M}_{kd}^t$ is pd since $\left(X^t\mathcal{V}X + k(1+d)\right)^2 > (X^t\mathcal{V}X)^2 \forall\, k > 0$ and $0 < d < 1$. Further, $\hat{\delta}(X^t\mathcal{V}X)^{-1}\left(I_p - \widetilde{M}_{kd}\widetilde{M}_{kd}^t\right)$ is pd iff $2\lambda_j k(1+d) > k^2(1+d+d^2) \forall\, j = 1,2,\ldots,p$. Thus, using Lemma 1 and 2, we conclude that $\nabla > 0$ iff $(\widetilde{M}_{kd} - I)\beta\beta^t(\widetilde{M}_{kd} - I)^t < 1$, then the proof is completed.

Appendix B: Proof of Theorem 2

The difference among the MMSE functions of BRRE and MBRTE is obtained by

$$\nabla = MMSE(\hat{\beta}_{BRRE}) - MMSE[\hat{\beta}_{MBRTE}(k,d)]$$
$$= \hat{\delta}(E) + k^2 F_k \beta\beta^t F_k - (\widetilde{M}_{kd} - I)\beta\beta^t(\widetilde{M}_{kd} - I)^t. \tag{B1}$$

However, Eq (B1) can be written in terms of scalar MSE as

$$MSE(\hat{\beta}_{BRR}) - MSE[\hat{\beta}_{MBRTE}(k,d)]$$
$$= \xi\hat{\delta}diag\left\{\frac{\lambda_j}{(\lambda_j + k)^2} - \frac{\lambda_j}{[\lambda_j + k(1+d)]^2}\right\}_{j=1}^{p} \xi^t + b_{BRRE}^t b_{BRRE} - b_{MBRTE}^t b_{MBRTE}$$
$$= \xi\hat{\delta}diag\left\{\frac{\lambda_j[\lambda_j + k(1+d)]^2 - \lambda_j(\lambda_j + k)^2}{(\lambda_j + k)^2[\lambda_j + k(1+d)]^2}\right\}_{j=1}^{p} \xi^t + b_{BRRE}^t b_{BRRE} - b_{MBRTE}^t b_{MBRTE}. \tag{B2}$$

where $F_k = (X^t \mathcal{V} X + kI)^{-1}$ and $E = F_k(X^t\mathcal{V}X)F_k^t - \widetilde{M}_{kd}(X^t\mathcal{V}X)^{-1}\widetilde{M}_{kd}^t$. Hence, $\hat{\delta}(E) + k^2 F_k\beta\beta^t F_k$ is non-negative, it is enough to prove that $F_k(X^t\mathcal{V}X)F_k^t - \widetilde{M}_{kd}(X^t\mathcal{V}X)^{-1}\widetilde{M}_{kd}^t - (\widetilde{M}_{kd} - I)\beta\beta^t(\widetilde{M}_{kd} - I)^t$ is pd. Further, $F_k(X^t\mathcal{V}X)F_k^t - \widetilde{M}_{kd}(X^t\mathcal{V}X)^{-1}\widetilde{M}_{kd}^t$ is pd iff $\lambda_j[\lambda_j + k(1+d)]^2 > \lambda_j(\lambda_j + k)^2 \forall\, j = 1,2,\ldots,p$. Thus, using Lemma 1 and 2, we conclude that $\nabla > 0$ iff $\beta^t(\widetilde{M}_{kd} - I)^t[\hat{\delta}\{F_k(X^t\mathcal{V}X)^{-1}F_k^t - \widetilde{M}_{kd}(X^t\mathcal{V}X)^{-1}\widetilde{M}_{kd}^t\}]^{-1}(\widetilde{M}_{kd} - I)\beta \le 1$.

Appendix C: Proof of Theorem 3

The difference among the MMSE functions of BLE and MBRTE is obtained by

$$\nabla = MMSE(\hat{\beta}_{BLE}) - MMSE[\hat{\beta}_{MBRTE}(k,d)]$$
$$= \hat{\delta}[L_d(X^t\mathcal{V}X)^{-1}L_d^t - \widetilde{M}_{kd}(X^t\mathcal{V}X)^{-1}\widetilde{M}_{kd}^t] + (d-1)^2 A_d\beta\beta^t A_d^t - (\widetilde{M}_{kd} - I)\beta\beta^t(\widetilde{M}_{kd} - I)^t. \tag{C1}$$

where $A_d = (X^t\mathcal{V}X + I)^{-1}$. Eq (C1) can be further expanded as

$$\nabla = MMSE(\hat{\beta}_{BLE}) - MMSE[\hat{\beta}_{MBRTE}(k,d)]$$

$$= \hat{\delta}[L_d(X^t\mathcal{V}X)^{-1}L_d^t - \widetilde{M}_{kd}(X^t\mathcal{V}X)^{-1}\widetilde{M}_{kd}^t] + (d-1)^2 A_d \beta\beta^t A_d^t$$

$$- (\widetilde{M}_{kd} - I)\beta\beta^t(\widetilde{M}_{kd} - I)^t . \nabla = MMSE(\hat{\beta}_{BLE}) - MMSE[\hat{\beta}_{MBRTE}(k,d)] \quad \text{(C2)}$$

$$= \hat{\delta}[L_d(X^t\mathcal{V}X)^{-1}L_d^t - \widetilde{M}_{kd}(X^t\mathcal{V}X)^{-1}\widetilde{M}_{kd}^t] + (d-1)^2 A_d \beta\beta^t A_d^t$$

$$- (\widetilde{M}_{kd} - I)\beta\beta^t(\widetilde{M}_{kd} - I)^t .$$

However, Eq (C2) can be written in terms of scalar MSE as

$$MSE(\hat{\beta}_{BLE}) - MSE[\hat{\beta}_{MBRTE}(k,d)]$$

$$= \xi\hat{\delta} diag\left\{\frac{(\lambda_j + d)^2}{\lambda_j(\lambda_j+1)^2} - \frac{\lambda_j}{[\lambda_j + k(1+d)]^2}\right\}_{j=1}^{p} \xi^t + b_{BLE}^t b_{BLE}$$

$$- b_{MBRTE}^t b_{MBRTE} \quad \text{(C3)}$$

$$= \xi\hat{\delta} diag\left\{\frac{(\lambda_j + d)^2[\lambda_j + k(1+d)]^2 - \lambda_j^2(\lambda_j+1)^2}{\lambda_j(\lambda_j+1)^2[\lambda_j + k(1+d)]^2}\right\}_{j=1}^{p} \xi^t + b_{BLE}^t b_{BLE}$$

$$- b_{MBRTE}^t b_{MBRTE} .$$

Since $(d-1)^2 A_d \beta\beta^t A_d^t$ in Eq (C2) is nonnegative definite, it is enough to prove that $L_d(X^t\mathcal{V}X)^{-1}L_d^t - \widetilde{M}_{kd}(X^t\mathcal{V}X)^{-1}\widetilde{M}_{kd}^t - (\widetilde{M}_{kd} - I)\beta\beta^t(\widetilde{M}_{kd} - I)^t$ is pd. $\hat{\delta}\,(X^t\mathcal{V}X)^{-1}$ $[L_d L_d^t - \widetilde{M}_{kd}\widetilde{M}_{kd}^t]$is pd iff $(\lambda_j + d)^2[\lambda_j + k(1+d)]^2 > \lambda_j^2(\lambda_j+1)^2 \forall j = 1,2,\ldots,p$. Thus if $k > 0, 0 < d < 1$, then the theorem is ended by Lemma 1.