Hacettepe Journal of Mathematics & Statistics

Hacet. J. Math. Stat.
Volume 52 (3) (2023), 828 – 840
DOI : 10.15672/hujms.1145607

Research Article

The beta Liu-type estimator:simulation and application

Ali Erkoç*[1], Esra Ertan[2], Zakariya Yahya Algamal[3,4], Kadri Ulaş Akay[2]

[1] *Department of Statistics, Faculty of Science and Letters, Mimar Sinan Fine Arts University, Istanbul, Turkey*
[2] *Department of Mathematics, Science Faculty, University of Istanbul, Istanbul, Turkey*
[3] *Department of Statistics and Informatics, University of Mosul, Mosul, Iraq*
[4] *College of Engineering, University of Warith Al-Anbiyaa, Karbala, Iraq*

Abstract

The Beta Regression Model (BRM) is commonly used while analyzing data where the dependent variable is restricted to the interval $[0, 1]$ for example proportion or probability. The Maximum Likelihood Estimator (MLE) is used to estimate the regression coefficients of BRMs. But in the presence of multicollinearity, MLE is very sensitive to high correlation among the explanatory variables. For this reason, we introduce a new biased estimator called the Beta Liu-Type Estimator (BLTE) to overcome the multicollinearity problem in the case that dependent variable follows a Beta distribution. The proposed estimator is a general estimator which includes other biased estimators, such as the Ridge Estimator, Liu Estimator, and the estimators with two biasing parameters as special cases in BRM. The performance of the proposed new estimator is compared to the MLE and other biased estimators in terms of the Estimated Mean Squared Error (EMSE) criterion by conducting a simulation study. Finally, a numerical example is given to show the benefit of the proposed estimator over existing estimators.

Mathematics Subject Classification (2020). 62J07, 62F10

Keywords. Beta regression model, Liu-type estimator, maximum likelihood estimator, multicollinearity

1. Introduction

The BRM is similar to a Binomial Generalized Linear Model (GLM) but provides some flexibility in particular when the trials are not independent and when the standard binomial model might be too strict. It has been used commonly in many areas, primarily engineering, medical sciences, physical sciences and social sciences. This model is used to examine the effect of some explanatory variables over a non-normal response variable. But in the case of BRM, the response variable is restricted to the interval [0,1] such as rates, proportions, percentages, probability and fractions. Firstly, the BRM was defined by [16] by relating the mean function of its response variable to a set of linear predictors

*Corresponding Author.
Email addresses: ali.erkoc@msgsu.edu.tr (A. Erkoç), eertan@istanbul.edu.tr (E. Ertan), zakariya.algamal@uomosul.edu.iq (Z.Y. Algamal), kulas@istanbul.edu.tr (K.U. Akay)
Received: 19.07.2022; Accepted: 10.01.2023

through a link function. This model includes a precision parameter, the inverse of which is called a dispersion scale [1,3].

Suppose that $y_1, y_2, \ldots, y_n$ are the observations of the random variable that follows a beta distribution. The beta probability density function is defined as

$$f(y;a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} y^{a-1}(1-y)^{b-1}, 0<y<1, \tag{1.1}$$

where $\Gamma(.)$ is the gamma function and $a, b > 0$. The mean and variance of beta probability distribution are given as follows:

$$E(y) = \frac{a}{a+b}, \quad \mathrm{Var}(y) = \frac{ab}{(a+b)^2(a+b+1)}. \tag{1.2}$$

Ferrari and Cribari-Neto [16] recommended a different parameterization by using $\frac{a}{a+b} = \mu$ and $a+b=\phi$, ϕ is called as the precision parameter. From these equalities a and b found as, $a=\mu\phi$, $b=\phi(1-\mu)$. Eq. (1.1) can be expressed through new parameterization as

$$f(y;\mu,\phi) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)} y^{\mu\phi-1}(1-y)^{(1-\mu)\phi-1}, 0<y<1, \tag{1.3}$$

where $0<\mu<1$ and $\phi>0$. The precision parameter ϕ can be written as $\phi = \frac{1-\sigma^2}{\sigma^2}$. With these transformations, the mean and variance of y are redefined respectively as

$$E(y) = \mu, \quad \mathrm{Var}(y) = \mu(1-\mu)\sigma^2. \tag{1.4}$$

The model allows the mean function to depend on linear predictors by using the following link function $g(.)$

$$g(\mu_i) = \log\left(\frac{\mu_i}{1-\mu_i}\right) = \mathbf{x_i'}\boldsymbol{\beta} = \eta_i, \tag{1.5}$$

where $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_p)'$ is an $p\times 1$ unknown parameters vector and $\mathbf{x_i} = (x_{i1}, x_{i2}, \ldots, x_{ip})'$ is the vector of p regressors and η_i is the linear predictor. This link function is strictly monotonic and twice differentiable [1,2,5,8,13,25]. Different link functions may be used for fitting the BRM as logit, probit, log-log, complementary log-log, and Cauchy link functions. However, Eq. (1.5) is a commonly used link function which is suggested by [16].

For the estimation of the BRM parameters, the MLE method is used by [13]. The log-likelihood function of the BRM is given by

$$\begin{aligned} L\left(\mu_i, \sigma_i^2; y_i\right) = \sum_{i=1}^{n} \Bigg\{ & \log\Gamma\left(\frac{1-\sigma_i^2}{\sigma_i^2}\right) - \log\Gamma\left(\mu_i\left(\frac{1-\sigma_i^2}{\sigma_i^2}\right)\right) - \log\Gamma\left((1-\mu_i)\left(\frac{1-\sigma_i^2}{\sigma_i^2}\right)\right) \\ & + \left(\mu_i\left(\frac{1-\sigma_i^2}{\sigma_i^2}\right) - 1\right)\log(y_i) + \left((1-\mu_i)\left(\frac{1-\sigma_i^2}{\sigma_i^2}\right) - 1\right)\log(1-y_i) \Bigg\}. \end{aligned} \tag{1.6}$$

The score function $S(\boldsymbol{\beta})$ can be find by differentiating the log-likelihood function in Eq. (1.6) with respect to $\boldsymbol{\beta}$

$$S(\boldsymbol{\beta}) = \phi\mathbf{x}'\mathbf{T}\left(\mathbf{y}^* - \boldsymbol{\mu}^*\right), \tag{1.7}$$

where $\mathbf{T} = \mathrm{diag}\left(\frac{1}{g(\mu_1)}, \frac{1}{g(\mu_2)}, \ldots, \frac{1}{g(\mu_n)}\right)$, $\mathbf{y}^* = (y_1^*, y_2^*, \ldots, y_n^*)'$, $\boldsymbol{\mu}^* = (\mu_1^*, \mu_2^*, \ldots, \mu_n^*)'$, $y_i^* = \log\left(\frac{y_i}{1-y_i}\right)$ and $\mu_i^* = \psi\left(\mu_i\left(\frac{1-\sigma_i^2}{\sigma_i^2}\right)\right) - \psi\left((1-\mu_i)\left(\frac{1-\sigma_i^2}{\sigma_i^2}\right)\right)$ where $\psi(.)$ denoting the digamma function. To obtain the estimated vector of $\boldsymbol{\beta}$, the Iterative Reweighted Least Square (IWLS) method or the Fisher Scoring method can be used [14,15]. By using these methods, the MLE of $\boldsymbol{\beta}$ is obtained as

$$\hat{\boldsymbol{\beta}}_{MLE} = \left(\mathbf{X}'\hat{\mathbf{W}}\mathbf{X}\right)^{-1}\mathbf{X}'\hat{\mathbf{W}}\mathbf{z}, \tag{1.8}$$

where $\mathbf{z} = \hat{\boldsymbol{\eta}} + \hat{\mathbf{W}}^{-1}\hat{\mathbf{T}}\left(\mathbf{y}^* - \boldsymbol{\mu}^*\right)$ and $\hat{\mathbf{W}} = \operatorname{diag}\left(\hat{w}_1, \ldots, \hat{w}_n\right)$ with

$$\hat{w}_i = \left(\frac{1-\hat{\sigma}_i^2}{\hat{\sigma}_i^2}\right)\left\{\psi'\left(\hat{\mu}_i\left(\frac{1-\hat{\sigma}_i^2}{\hat{\sigma}_i^2}\right)\right) + \psi'\left((1-\hat{\mu}_i)\left(\frac{1-\hat{\sigma}_i^2}{\hat{\sigma}_i^2}\right)\right)\right\}\frac{1}{\left\{g'\left(\hat{\mu}_i\right)\right\}^2}. \tag{1.9}$$

In Eq. (1.9), $\hat{\mathbf{W}}$ and $\hat{\mathbf{T}}$ are the estimated matrices of $\mathbf{W}$ and $\mathbf{T}$ from maximum likelihood estimation respectively [16]. The asymptotic covariance matrix of the MLE equals to

$$\operatorname{Cov}\left(\hat{\boldsymbol{\beta}}_{MLE}\right) = \frac{1}{\phi}\left(\mathbf{X}'\hat{\mathbf{W}}\mathbf{X}\right)^{-1}. \tag{1.10}$$

When the explanatory variables are correlated, the multicollinearity problem appears. In the presence of multicollinearity, $\mathbf{X}'\hat{\mathbf{W}}\mathbf{X}$ matrix is ill-conditioned. One of the disadvantages of using the MLE is that the variance of parameters becomes inflated when the collinearity among explanatory variables is severe. To prevent the undesirable effects of multicollinearity, many researchers have chosen to generalize the biased estimators used for linear regression models to apply on BRMs. For more detailed information about these proposed biased estimators in GLMs and BRMs, the articles [1–4,6–10,12,17,18,20,21,26,27] can be reviewed.

Firstly, to overcome the problem of the BRM, Abonazel and Taha [3] and Qasim et al. [26] introduced the Beta Ridge Estimator (BRE) which is the generalization of Hoerl and Kennard [17] as an alternative to the MLE. The BRE is defined as

$$\hat{\boldsymbol{\beta}}_{BRE} = \left(\mathbf{X}'\hat{\mathbf{W}}\mathbf{X} + k\mathbf{I}\right)^{-1}\mathbf{X}'\hat{\mathbf{W}}\mathbf{z}, \tag{1.11}$$

where $k > 0$ and when $k = 0$, $\hat{\boldsymbol{\beta}}_{BRE} = \hat{\boldsymbol{\beta}}_{MLE}$.

Then, Karlsson et al. [18] introduced another estimation method called Beta Liu Estimator (BLE) for the BRM as

$$\hat{\boldsymbol{\beta}}_{BLE} = \left(\mathbf{X}'\hat{\mathbf{W}}\mathbf{X} + \mathbf{I}\right)^{-1}\left(\mathbf{X}'\hat{\mathbf{W}}\mathbf{X} + d\mathbf{I}\right)\hat{\boldsymbol{\beta}}_{MLE}, \tag{1.12}$$

where d is the Liu parameter and $0 < d < 1$. The BLE is the generalization of the Liu estimator defined by [20] for the linear regression models.

On the other hand, Liu [21] proposed the Liu-type estimator for the linear regression model to overcome the multicollinearity problem. Algamal and Abonazel [8] adapted this estimator to BRM and called the Liu-Type Beta Regression (LTBR) estimator as

$$\hat{\boldsymbol{\beta}}_{LTBR} = \left(\mathbf{X}'\hat{\mathbf{W}}\mathbf{X} + k\mathbf{I}\right)^{-1}(\mathbf{X}\hat{\mathbf{W}}\mathbf{X} - d\mathbf{I})\hat{\boldsymbol{\beta}}_{MLE}, \tag{1.13}$$

where $k > 0$ and $-\infty < d < \infty$.

In addition, the two-parameter beta regression (TPBR) estimator for the BRM is obtained by [1] to combat multicollinearity as follows:

$$\hat{\boldsymbol{\beta}}_{TPBR} = \left(\mathbf{X}'\hat{\mathbf{W}}\mathbf{X} + k\mathbf{I}\right)^{-1}\left(\mathbf{X}'\hat{\mathbf{W}}\mathbf{X} + kd\mathbf{I}\right)\hat{\boldsymbol{\beta}}_{MLE}, \tag{1.14}$$

where $k > 0$ and $0 < d < 1$.

Abonazel et al. [2] introduced the beta version of the two-parameter estimator of [11] as follows:

$$\hat{\boldsymbol{\beta}}_{BDK} = \left(\mathbf{X}'\hat{\mathbf{W}}\mathbf{X} + k(1+d)\mathbf{I}\right)^{-1}\left(\mathbf{X}'\hat{\mathbf{W}}\mathbf{X} - k(1+d)\mathbf{I}\right)\hat{\boldsymbol{\beta}}_{MLE}, \tag{1.15}$$

where $k > 0$ and $0 < d < 1$.

The aim of this study is to introduce a new Beta Liu-type estimator for the BRMs. Show the superiority of this estimator from other estimators in overcoming the multicollinearity problem. More on Liu-type estimators for different models, we refer to [19], [23] and [22] among others.

In this paper, a new biased estimator named the beta Liu-type estimator is proposed and some of its statistical properties are given in Section 2. In Section 3, the approaches used to determine the biasing parameters for proposed biased estimators are summarized. Furthermore, several methods are proposed to determine the biasing parameters. A Monte Carlo simulation studies are executed in Section 4 to show the superiority of this estimator over the other biased estimators. In Section 5, a real data application is provided to illustrate the performances of the proposed estimators. Finally, conclusions of the study are given in Section 6.

2. The beta Liu-type estimator

Ertan and Akay [12] proposed a new general Liu-type estimator to reduce the effects of multicollinearity in logistic regression models. We implemented this estimator in BRMs. The beta Liu-type estimator (BLTE) is defined as

$$\hat{\boldsymbol{\beta}}_{\text{BLTE}} = \left(\mathbf{X}'\hat{\mathbf{W}}\mathbf{X} + k\mathbf{I}\right)^{-1}\left(\mathbf{X}'\hat{\mathbf{W}}\mathbf{X} + f(k)\mathbf{I}\right)\hat{\boldsymbol{\beta}}^{*}, k > 0, \tag{2.1}$$

where $\hat{\boldsymbol{\beta}}^{*}$ is any estimator of $\boldsymbol{\beta}$, k is a biasing parameter and $f(k)$ is a continuous function of the biasing parameter k. When we selected $f(k)$ as a linear function of the biasing parameter k such as $f(k) = ak + b$ where $a, b \in R$, the BLTE becomes a general estimator which includes the other biased estimators as special cases which can be summarized as follows:

$\hat{\boldsymbol{\beta}}_{BLTE} = \hat{\boldsymbol{\beta}}_{MLE}$, for $\hat{\boldsymbol{\beta}}^{*} = \hat{\boldsymbol{\beta}}_{MLE}$ and $f(k) = k$ where $a = 1$ and $b = 0$.

$\hat{\boldsymbol{\beta}}_{BLTE} = \hat{\boldsymbol{\beta}}_{BRE}$, for $\hat{\boldsymbol{\beta}}^{*} = \hat{\boldsymbol{\beta}}_{MLE}$ and $f(k) = 0$ where $a = 0$ and $b = 0$.

$\hat{\boldsymbol{\beta}}_{BLTE} = \hat{\boldsymbol{\beta}}_{BLE}$, for $\hat{\boldsymbol{\beta}}^{*} = \hat{\boldsymbol{\beta}}_{MLE}$ and $f(1) = a + b$ where $a + b$ corresponds to the biasing parameter d.

$\hat{\boldsymbol{\beta}}_{BLTE} = \hat{\boldsymbol{\beta}}_{LTBR}$, for $\hat{\boldsymbol{\beta}}^{*} = \hat{\boldsymbol{\beta}}_{MLE}$ and $f(k) = -b$ where b corresponds to the biasing parameter d.

$\hat{\boldsymbol{\beta}}_{BLTE} = \hat{\boldsymbol{\beta}}_{TPBR}$, for $\hat{\boldsymbol{\beta}}^{*} = \hat{\boldsymbol{\beta}}_{MLE}$ and $f(k) = ak$ where a corresponds to the biasing parameter d.

The Matrix Mean Squared Error (MMSE) and Scaler Mean Squared Error (SMSE) of an estimator $\hat{\boldsymbol{\beta}}$ are defined as

$$\text{MMSE}(\hat{\boldsymbol{\beta}}) = \text{Cov}(\hat{\boldsymbol{\beta}}) + \text{bias}(\hat{\boldsymbol{\beta}})\,\text{bias}(\hat{\boldsymbol{\beta}})', \tag{2.2}$$

$$\text{SMSE}(\hat{\boldsymbol{\beta}}) = trace(\text{MMSE}(\hat{\boldsymbol{\beta}})). \tag{2.3}$$

For the convenience of comparisons, we assume that $\lambda_1, \lambda_2, \ldots, \lambda_p \geq 0$ are eigenvalues of $\mathbf{X}'\hat{\mathbf{W}}\mathbf{X}$ matrix and $\mathbf{Q}$ is the matrix whose columns are the eigenvectors of $\mathbf{X}'\hat{\mathbf{W}}\mathbf{X}$ matrix. Let $\boldsymbol{\Lambda} = \text{diag}\left(\lambda_1, \ldots, \lambda_p\right) = \mathbf{Q}'\mathbf{X}'\hat{\mathbf{W}}\mathbf{X}\mathbf{Q}$ and $\boldsymbol{\alpha} = \mathbf{Q}'\boldsymbol{\beta}$.

By using Eq. (2.2) and Eq. (2.3) the MMSE of Eqs. (1.8), (1.11)–(1.15) and (2.1) are obtained as follows:

$$\text{MMSE}\left(\hat{\boldsymbol{\beta}}_{MLE}\right) = \frac{1}{\phi}\mathbf{Q}\boldsymbol{\Lambda}^{-1}\mathbf{Q}' \tag{2.4}$$

$$MMSE\left(\hat{\boldsymbol{\beta}}_{BRE}\right) = \frac{1}{\phi}\left(\mathbf{Q}\boldsymbol{\Lambda}_{\mathbf{k}}^{-1}\boldsymbol{\Lambda}\boldsymbol{\Lambda}_{\mathbf{k}}^{-1}\mathbf{Q}'\right) + k^2\mathbf{Q}\boldsymbol{\Lambda}_{\mathbf{k}}^{-1}\boldsymbol{\alpha}\boldsymbol{\alpha}'\boldsymbol{\Lambda}_{\mathbf{k}}^{-1}\mathbf{Q}', \tag{2.5}$$

where $\boldsymbol{\Lambda}_k = \text{diag}\left(\lambda_1 + k, \ldots, \lambda_p + k\right)$.

$$\text{MMSE}\left(\hat{\boldsymbol{\beta}}_{BLE}\right) = \frac{1}{\phi}\left(\mathbf{Q}\boldsymbol{\Lambda}_1^{-1}\boldsymbol{\Lambda}_d\boldsymbol{\Lambda}^{-1}\boldsymbol{\Lambda}_d\boldsymbol{\Lambda}_1^{-1}\mathbf{Q}'\right) + (d-1)^2\mathbf{Q}\boldsymbol{\Lambda}_1^{-1}\boldsymbol{\alpha}\boldsymbol{\alpha}'\boldsymbol{\Lambda}_1^{-1}\mathbf{Q}', \tag{2.6}$$

where $\boldsymbol{\Lambda}_1 = \text{diag}\left(\lambda_1+1,\ldots,\lambda_p+1\right)$ and $\boldsymbol{\Lambda}_d = \text{diag}\left(\lambda_1+d,\ldots,\lambda_p+d\right)$.

$$\text{MMSE}\left(\hat{\boldsymbol{\beta}}_{LTBR}\right) = \frac{1}{\phi}\left(\mathbf{Q}\boldsymbol{\Lambda}_k^{-1}\boldsymbol{\Lambda}_{-d}\boldsymbol{\Lambda}^{-1}\boldsymbol{\Lambda}_{-d}\boldsymbol{\Lambda}_k^{-1}\mathbf{Q}'\right) + (d+k)^2\mathbf{Q}\boldsymbol{\Lambda}_k^{-1}\boldsymbol{\alpha}\boldsymbol{\alpha}'\boldsymbol{\Lambda}_k^{-1}\mathbf{Q}', \tag{2.7}$$

where $\boldsymbol{\Lambda}_{-d} = \text{diag}\left(\lambda_1-d,\ldots,\lambda_p-d\right)$.

$$\text{MMSE}\left(\hat{\boldsymbol{\beta}}_{\text{TPRR}}\right) = \frac{1}{\phi}\left(\mathbf{Q}\boldsymbol{\Lambda}_k^{-1}\boldsymbol{\Lambda}_{kd}\boldsymbol{\Lambda}^{-1}\boldsymbol{\Lambda}_{kd}\boldsymbol{\Lambda}_k^{-1}\mathbf{Q}'\right) + k^2(d-1)^2\mathbf{Q}\boldsymbol{\Lambda}_k^{-1}\boldsymbol{\alpha}\boldsymbol{\alpha}'\boldsymbol{\Lambda}_k^{-1}\mathbf{Q}', \tag{2.8}$$

where $\boldsymbol{\Lambda}_{kd} = \text{diag}\left(\lambda_1+kd,\ldots,\lambda_p+kd\right)$.

$$\begin{aligned}\text{MMSE}\left(\hat{\boldsymbol{\beta}}_{BDK}\right) &= \frac{1}{\phi}\left(\mathbf{Q}\boldsymbol{\Lambda}_{k(1+d)}^{-1}\boldsymbol{\Lambda}_{-k(1+d)}\boldsymbol{\Lambda}^{-1}\boldsymbol{\Lambda}_{-k(1+d)}\boldsymbol{\Lambda}_{k(1+d)}^{-1}\mathbf{Q}'\right)\\ &+\left(\mathbf{Q}\boldsymbol{\Lambda}_{k(1+d)}^{-1}\boldsymbol{\Lambda}_{-k(1+d)}\mathbf{Q}'-\mathbf{I}\right)\alpha\boldsymbol{\alpha}'\left(\mathbf{Q}\boldsymbol{\Lambda}_{k(1+d)}^{-1}\boldsymbol{\Lambda}_{-k(1+d)}\mathbf{Q}'-\mathbf{I}\right)',\end{aligned} \tag{2.9}$$

where

$$\boldsymbol{\Lambda}_{k(1+d)} = \text{diag}\left(\lambda_1+k(1+d),\ldots,\lambda_p+k(1+d)\right)$$

and

$$\boldsymbol{\Lambda}_{-k(1+d)} = \text{diag}\left(\lambda_1-k(1+d),\ldots,\lambda_p-k(1+d)\right).$$

$$\text{MMSE}\left(\hat{\boldsymbol{\beta}}_{BLTE}\right) = \frac{1}{\phi}\left(\mathbf{Q}\boldsymbol{\Lambda}_k^{-1}\boldsymbol{\Lambda}_{f(k)}\boldsymbol{\Lambda}^{-1}\boldsymbol{\Lambda}_{f(k)}\boldsymbol{\Lambda}_k^{-1}\mathbf{Q}'\right) + (f(k)-k)^2\mathbf{Q}\boldsymbol{\Lambda}_k^{-1}\boldsymbol{\alpha}\boldsymbol{\alpha}'\boldsymbol{\Lambda}_k^{-1}\mathbf{Q}', \tag{2.10}$$

where $\boldsymbol{\Lambda}_{f(k)} = \text{diag}\left(\lambda_1+f(k),\ldots,\lambda_p+f(k)\right)$.

Similarly, the SMSE of Eqs. (1.8), (1.11)–(1.15) and (2.1) are defined as

$$\text{SMSE}\left(\hat{\boldsymbol{\beta}}_{MLE}\right) = \frac{1}{\phi}\sum_{j=1}^{p}\frac{1}{\lambda_j}, \tag{2.11}$$

$$\text{SMSE}\left(\hat{\boldsymbol{\beta}}_{BRE}\right) = \frac{1}{\phi}\sum_{j=1}^{p}\frac{\lambda_j}{(\lambda_j+k)^2} + k^2\sum_{j=1}^{p}\frac{\alpha_j^2}{(\lambda_j+k)^2}, \tag{2.12}$$

$$\text{SMSE}\left(\hat{\boldsymbol{\beta}}_{BLE}\right) = \frac{1}{\phi}\sum_{j=1}^{p}\frac{(\lambda_j+d)^2}{\lambda_j\left(\lambda_j+1\right)^2} + (d-1)^2\sum_{j=1}^{p}\frac{\alpha_j^2}{\left(\lambda_j+1\right)^2}, \tag{2.13}$$

$$\text{SMSE}\left(\hat{\boldsymbol{\beta}}_{\text{LTBR}}\right) = \frac{1}{\phi}\sum_{j=1}^{p}\frac{(\lambda_j-d)^2}{\lambda_j\left(\lambda_j+k\right)^2} + (d+k)^2\sum_{j=1}^{p}\frac{\alpha_j^2}{\left(\lambda_j+k\right)^2}, \tag{2.14}$$

$$\text{SMSE}\left(\hat{\boldsymbol{\beta}}_{\text{TPBR}}\right) = \frac{1}{\phi}\sum_{j=1}^{p}\frac{(\lambda_j+kd)^2}{\lambda_j\left(\lambda_j+k\right)^2} + k^2(d-1)^2\sum_{j=1}^{p}\frac{\alpha_j^2}{\left(\lambda_j+k\right)^2}, \tag{2.15}$$

$$\text{SMSE}\left(\hat{\boldsymbol{\beta}}_{BDK}\right) = \frac{1}{\phi}\sum_{j=1}^{p}\frac{(\lambda_j-k(1+d))^2}{\lambda_j\left(\lambda_j+k(1+d)\right)^2} + 4k^2(1+d)^2\sum_{j=1}^{p}\frac{\alpha_j^2}{\left(\lambda_j+k(1+d)\right)^2}, \tag{2.16}$$

$$\text{SMSE}\left(\hat{\boldsymbol{\beta}}_{BLTE}\right) = \frac{1}{\phi}\sum_{j=1}^{p}\frac{(\lambda_j+f(k))^2}{\lambda_j\left(\lambda_j+k\right)^2} + (f(k)-k)^2\sum_{j=1}^{p}\frac{\alpha_j^2}{\left(\lambda_j+k\right)^2}. \tag{2.17}$$

In these equalities, the first term is an asymptotic variance, and the second term is a squared bias.

3. Determination of $f(k)$ function

There is no strict analytical rule for estimating the biasing parameters. So, it is an important problem to find acceptable estimates of these parameters. Several approaches have been recommended for choosing the best biasing parameters of k and d. Several methods for estimating the value of k in BRE have been extensions of the methods proposed in linear regression models. Abonazel and Taha [3] suggested the following new estimation of biasing parameter k

$$k_{BRE1} = \frac{\lambda_{\min}}{\hat{\phi}\hat{\alpha}^2_{\min}}, \tag{3.1}$$

where $\hat{\boldsymbol{\alpha}} = (\hat{\alpha}_1, \dots \hat{\alpha}_p)' = \mathbf{Q}'\hat{\boldsymbol{\beta}}_{BML}, \hat{\phi} = \frac{\hat{\mu}(1-\hat{\mu})}{\hat{\sigma}^2}$ and $\lambda_{\min}$ is the minimum eigenvalue. In addition, Qasim et al. [25] also proposed another estimator of the biasing parameter k as follows:

$$k_{BRE2} = \frac{\sum_{j=1}^{p} \lambda_j \alpha_j^2}{p\left(\frac{1}{\phi}\right)}. \tag{3.2}$$

For BLE, Karlsson et al. [18] used the following method to estimate the biasing parameter d:

$$d_{BLE} = \max\left(0, \min\left(\left\{\frac{\hat{\beta}_j^2 - 1}{\hat{\lambda}_j^{-1} + \hat{\beta}_j^2}\right\}_{j=1}^{p}\right)\right). \tag{3.3}$$

To obtain the value of k for the LTBR estimator, Algamal and Abonazel [8] used the k parameter of [17] after tuning their formula based on the optimal k of BRMs as proposed by [26] as follows:

$$k_{LTBR} = \frac{1}{\phi \sum_{j=1}^{p} \alpha_j^2}. \tag{3.4}$$

Additionally, Algamal and Abonazel [8] used the following biasing parameter d for LTBR estimator:

$$d_{LTBR} = \frac{\sum_{j=1}^{p}\left(\frac{\frac{1}{\phi} - k_{LTBR}\alpha_j^2}{(\lambda_j + k_{LTBR})^2}\right)}{\sum_{j=1}^{p}\left(\frac{\frac{1}{\phi} + \lambda_j\alpha_j^2}{\lambda_j(\lambda_j + k_{LTBR})^2}\right)}. \tag{3.5}$$

In order to estimate TPBR parameters with minimum SMSE values, Abonazel et al. [1] proposed the biasing parameters d and k which are given in Eqs. (3.6) and (3.7), respectively.

$$d_{\text{TPBR}} = \frac{1}{2}\min\left(\frac{\lambda_j \hat{\alpha}_j^2}{\frac{1}{\phi} + \lambda_j \hat{\alpha}_j^2}\right)_{j=1}^{p}, \tag{3.6}$$

$$k_{\text{TPBR}} = \frac{1}{p}\sum_{j=1}^{p}\left(\frac{\lambda_j}{\phi\left(\lambda_j \hat{\alpha}_j^2 \left(1 - d_{\text{TPBR}}\right) - \frac{d_{TPBR}}{\phi}\right)}\right), \tag{3.7}$$

where $0 < d_{\text{TPBR}} < 1$ and $k_{\text{TPBR}} > 0$.

Recently, Abonazel et al. [2] suggested a new biasing parameter k of BDK estimator based on the work of Dawoud and Kibria [11], as follows:

$$k_{BDK} = \frac{1}{p}\left(\sum_{j=1}^{p} \frac{1}{\hat{\phi}\left(1 + d_{BOK}\right)\left(\frac{1}{\hat{\phi}\lambda_j} + 2\hat{\alpha}_j^2\right)}\right)^{\frac{1}{p}}, \tag{3.8}$$

where $d_{BOK} = \min\left(\frac{\hat{\alpha}_j^2}{\frac{1}{\phi\lambda_j} + 2\hat{\alpha}_j^2}\right)_{j=1}^{p}$.

The performance of our new estimator BLTE is dependent to function $f(k)$. So that we have only one biasing parameter k The appropriate selection of $f(k)$ functions yield different biased estimators. We can give a method to find the optimal $f(k)$ function minimizing SMSE $\left(\hat{\boldsymbol{\beta}}_{\mathrm{BLTE}}\right)$ according to parameter k. To find the optimal $f(k)$ function, we take the derivative of $h(k) = \mathrm{SMSE}\left(\hat{\boldsymbol{\beta}}_{\mathrm{BLTE}}\right)$ with respect to k. Then, we get

$$h'(k) = \sum_{j=1}^{p}\left(\frac{\left(f'(k)\left(\lambda_j + k\right) - \left(f(k) + \lambda_j\right)\right)\left(\frac{2}{\phi}\left(\lambda_j + f(k)\right) + 2\lambda_j\alpha_j^2(f(k) - k)\right)}{\lambda_j\left(\lambda_j + k\right)^3}\right). \tag{3.9}$$

When the derivative given in Eq. (3.9) is set to 0, we have these two facts,

Fact 1.$f'(k)\left(\lambda_j + k\right) - \left(f(k) + \lambda_j\right) = 0$. From the solution of this differential equation, we obtain

$$f(k) = c_1 k + \left(c_1 - 1\right)\lambda_j, j = 1, \ldots, p, \tag{3.10}$$

where c_1 is the constant of integration.

Fact 2. $\frac{2}{\phi}\left(\lambda_j + f(k)\right) + 2\lambda_j\alpha_j^2(f(k) - k) = 0$. From this equation we obtain

$$f(k) = \frac{\lambda_j\alpha_j^2}{\frac{1}{\phi} + \lambda_j\alpha_j^2}k - \frac{\frac{\lambda_j}{\phi}}{\frac{1}{\phi} + \lambda_j\alpha_j^2}, j = 1, \ldots, p. \tag{3.11}$$

According to Eqs. (3.10) and (3.11), the selection of $f(k) = ak + b$ $(a, b \in R)$ as a linear function of the biasing parameter k is applicable. Note that, $f(k)$ defined in Eq. (3.11) is a solution of the differential equation that is given in Fact 1. The $f(k)$ in Eqs. (3.10) and (3.11) makes the SMSE $\left(\hat{\boldsymbol{\beta}}_{\mathrm{BLTE}}\right)$ function approximately minimum for a given value of j. So, the function $f(k)$ depends on the eigenvalues of $\mathbf{X}'\hat{\mathbf{W}}\mathbf{X}$, the biasing parameter k and the unknown parameter α. With the structure of $f(k) = ck + (c-1)\lambda_{min}$ where $c \in (0, 1)$, we used the following functions for the determination of $f(k)$ in this paper:

$$f_1(k) = \frac{\lambda_{\min}\alpha_{\min}^2}{p + \lambda_{\max}\alpha_{\max}^2}k + \left(\frac{\lambda_{\min}\alpha_{\min}^2}{p + \lambda_{\max}\alpha_{\max}^2} - 1\right)\lambda_{\min}, \tag{3.12}$$

$$f_2(k) = \frac{\lambda_{\min}\alpha_{\min}^2}{n\left(1 + p\lambda_{\max}\alpha_{\max}^2\right)}k + \left(\frac{\lambda_{\min}\alpha_{\min}^2}{n\left(1 + p\lambda_{\max}\alpha_{\max}^2\right)} - 1\right)\lambda_{\min}, \tag{3.13}$$

$$f_3(k) = \frac{\min\left(\lambda_j\alpha_j^2\right)}{n\max\left(\phi + \lambda_j\alpha_j^2\right)}k + \left(\frac{\min\left(\lambda_j\alpha_j^2\right)}{n\max\left(\phi + \lambda_j\alpha_j^2\right)} - 1\right)\lambda_{\min}, \tag{3.14}$$

where α_{min}^2 and α_{max}^2 is defined as the minimum and maximum value of $\alpha_j^2, j = 1, 2, \ldots, p$.

Similarly, λ_{min} and λ_{max} is defined as the minimum and maximum eigenvalue of $\mathbf{X}'\hat{\mathbf{W}}\mathbf{X}$ respectively. Based on the simulation studies, we can use the following estimators to estimate k in the BLTEs

$$\hat{k}_{BLTE} = \frac{\lambda_{\max} + p\lambda_{\min}}{1 + \sqrt{p}}, \tag{3.15}$$

$$\hat{k}_{BLTE} = \left(\frac{np\lambda_{\min}}{\alpha_{\max}^2}\right)^{\frac{1}{p}}, \tag{3.16}$$

$$\hat{k}_{BLTE} = \sqrt{\frac{\sum_{j=1}^{p} \lambda_j}{p}}, \tag{3.17}$$

where p indicates the number of explanatory variables. We should note that k in the BLTEs must be estimated in such a way that the conditioning of the $\mathbf{X}'\hat{\mathbf{W}}\mathbf{X}$ matrix is controlled.

4. Monte-Carlo simulation study

In this section, a Monte Carlo simulation experiment is conducted to examine the performance of our new proposed estimator under different scenarios.

4.1. The design of the experiment

We designed our experiment according to the following conditions:

(1) The response variable is generated from the Beta distribution as $y_i \sim \text{Beta}\,(\mu_i, \phi)$ where $\mu_i = \frac{\exp(\mathbf{x}_i'\boldsymbol{\beta})}{(1+\exp(\mathbf{x}_i'\boldsymbol{\beta}))}, i = 1, 2, \ldots, n, \mathbf{x}_i'$ is the i th row of $\mathbf{X}, \phi$ is the precision parameter and $\boldsymbol{\beta}$ indicates the unknown regression coefficients vector.

(2) $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)$ parameter vector is chosen as $\sum_{j=1}^{p} \beta_j^2 = 1$ and $\beta_1 = \beta_2 = \ldots = \beta_p$.

(3) The correlated explanatory variables are generated as $x_{ij} = (1-\rho^2)^{\frac{1}{2}} z_{ij} + \rho z_{ip}, i = 1, 2, \ldots, n, j = 1, 2, \ldots, p$ where ρ represents the correlation between the explanatory variables and z_{ij} are independent standard normal pseudorandom numbers.

(4) The precision parameter ϕ in the simulation is chosen as $0.5, 1.5, 5$.

(5) Sample size n is taken as $50, 100$ and 200.

(6) The correlation ρ between the explanatory variables taken as $0.90, 0.95$ and 0.99.

(7) The number of explanatory variables p selected as $4, 8$ and 12.

We used the EMSE criterion for comparison, which are computed as $\text{EMSE}(\hat{\boldsymbol{\beta}}) = \frac{1}{N}\sum_{r=1}^{N} \left(\hat{\boldsymbol{\beta}}_r - \boldsymbol{\beta}\right)' \left(\hat{\boldsymbol{\beta}}_r - \boldsymbol{\beta}\right)$ where $\hat{\boldsymbol{\beta}}_r$ is the estimated value vector at the rth experiment of the simulation and $\boldsymbol{\beta}$ is the real parameter vector. The number of replications N is chosen as 2000.

In the simulation study, as the proposed estimator compared with the other estimators, the best biasing parameters suggested for each estimator were considered. To estimate the biasing parameter k in BRE, we used the best estimation of k as given in Eqs. (3.1) and (3.2) which are recommended by [3] and [25], respectively. For BLE, we used the best estimate of d given in Eq. (3.3) by [18]. Based on the results given by [8], we used the best estimation of d as defined in Eq. (3.5) in LTBR. Also k_{LTBR} is computed from Eq. (3.4) which is proposed by [26]. For TPBR, the biasing parameters d and k are estimated by Eq. (3.6) and Eq. (3.7) respectively from study of [1]. Abonazel et al. [2] suggested the biasing parameter k of BDK by Eq. (3.8). d_{BOK} is taken by [11].

The obtained results are reported in Table 1, Table 2 and Table 3, together with the following estimates of k and $f(k)$ functions:

BLTE1: $\hat{k}_{BLTE} = \frac{\lambda_{\max}+p\lambda_{\min}}{1+\sqrt{p}}$ and $f_1(k) = \frac{\lambda_{\min}\alpha_{\min}^2}{p+\lambda_{\max}\alpha_{\max}^2}k + \left(\frac{\lambda_{\min}\alpha_{\min}^2}{p+\lambda_{\max}\alpha_{\max}^2} - 1\right)\lambda_{\min}$

BLTE2: $\hat{k}_{\text{BLTE}} = \left(\frac{np\lambda_{\min}}{\alpha_{\max}^2}\right)^{\frac{1}{p}}$ and $f_2(k) = \frac{\lambda_{\min}\alpha_{\min}^2}{n(1+p\lambda_{\max}\alpha_{\max}^2)}k + \left(\frac{\lambda_{\min}\alpha_{\min}^2}{n(1+p\lambda_{\max}\alpha_{\max}^2)} - 1\right)\lambda_{\min}$

BLTE3: $\hat{k}_{BLTE} = \sqrt{\frac{\sum_{j=1}^{p}\lambda_j}{p}}$ and $f_3(k) = \frac{\min(\lambda_j\alpha_j^2)}{n\max(\phi+\lambda_j\alpha_j^2)}k + \left(\frac{\min(\lambda_j\alpha_j^2)}{n\max(\phi+\lambda_j\alpha_j^2)} - 1\right)\lambda_{\min}.$

Table 1. The EMSE values of the estimators when $p = 4$.

n	ρ	ϕ	MLE	BRE1	BRE2	BLE	LTBR	TPBR	BDK	BLTE1	BLTE2	BLTE3
50	0.9	0.5	21.6311	3.6003	1.0716	1.1643	9.0118	1.4029	10.192	**0.7263**	**0.7098**	**0.7260**
50	0.9	1.5	18.4436	2.4824	0.9561	0.8178	5.5637	0.8992	11.1834	**0.6166**	**0.6238**	**0.6163**
50	0.9	5	6.6955	0.8501	0.899	0.7504	1.1668	0.8183	3.698	**0.6307**	**0.6632**	**0.6202**
50	0.95	0.5	43.9383	7.4231	0.9356	1.4029	18.1467	1.6364	29.7873	**0.7005**	**0.7051**	**0.7015**
50	0.95	1.5	48.8748	7.0033	0.7345	0.8127	13.4467	0.8631	39.909	**0.5472**	**0.5536**	**0.5467**
50	0.95	5	17.8857	1.3314	0.7831	0.6798	2.8619	0.7658	13.9309	**0.6201**	**0.6245**	**0.6155**
50	0.99	0.5	175.3199	32.5725	0.7033	2.7417	68.0766	1.5107	158.5961	**0.6319**	**0.6668**	**0.6339**
50	0.99	1.5	261.0659	37.8245	0.5874	1.4108	71.7951	0.7093	250.9298	**0.5398**	**0.5369**	**0.5393**
50	0.99	5	74.4915	2.8775	0.6471	0.639	7.73	0.7037	69.9839	**0.6008**	**0.5668**	**0.5884**
100	0.9	0.5	27.7197	4.1028	0.9794	1.2741	10.8069	1.5193	15.0786	**0.6619**	**0.6668**	**0.6621**
100	0.9	1.5	14.9681	1.7692	0.9901	0.8283	4.2902	0.9092	8.1072	**0.6041**	**0.6409**	**0.6025**
100	0.9	5	7.6032	0.8617	0.89	0.7487	1.2302	0.8169	4.4699	**0.6368**	**0.6791**	**0.6258**
100	0.95	0.5	59.5852	9.3054	0.7901	1.3429	22.5727	1.3973	45.0624	**0.6165**	**0.6218**	**0.6176**
100	0.95	1.5	39.3112	4.405	0.7473	0.7717	9.9867	0.8192	30.921	**0.5383**	**0.5501**	**0.5363**
100	0.95	5	14.3426	1.0337	0.7939	0.6886	1.9992	0.7772	10.6244	**0.6176**	**0.6339**	**0.6078**
100	0.99	0.5	201.1309	31.0208	0.6711	2.2665	74.9197	1.1643	185.1741	**0.6116**	**0.6322**	**0.6125**
100	0.99	1.5	120.4205	15.2105	0.6078	0.9044	30.3626	0.7097	111.4048	**0.525**	**0.5197**	**0.5238**
100	0.99	5	48.7294	2.0461	0.6597	0.6298	5.3294	0.7099	44.3548	**0.5933**	**0.574**	**0.5806**
200	0.9	0.5	29.7046	4.3444	0.9168	1.289	11.064	1.4069	17.2374	**0.6313**	**0.6408**	**0.6317**
200	0.9	1.5	14.9788	1.7006	0.942	0.7809	3.9369	0.8776	8.1203	**0.5662**	**0.6316**	**0.5629**
200	0.9	5	8.8419	0.8684	0.8608	0.7394	1.4213	0.8094	5.4083	**0.6286**	**0.6837**	**0.6169**
200	0.95	0.5	52.7961	7.6841	0.8287	1.2225	19.7501	1.2782	39.0181	**0.6445**	**0.6453**	**0.6459**
200	0.95	1.5	32.8096	3.3747	0.7976	0.7786	8.4448	0.8461	24.4116	**0.5542**	**0.5822**	**0.5518**
200	0.95	5	12.8091	0.9417	0.7838	0.6732	1.6808	0.7676	9.2515	**0.6024**	**0.6399**	**0.5893**
200	0.99	0.5	219.4048	30.067	0.7079	2.3876	80.3514	1.1753	203.2388	**0.6436**	**0.6497**	**0.6438**
200	0.99	1.5	129.3059	12.6203	0.6081	1.0003	31.0791	0.7142	119.9668	**0.5331**	**0.5257**	**0.5308**
200	0.99	5	59.2135	2.3877	0.6399	0.6216	5.9221	0.6946	54.9485	**0.5892**	**0.5719**	**0.5764**

4.2. Simulation results

For the simulation study, we used R-software. The EMSE for all the combinations of n, p, ρ and ϕ are summarized in Table 1-2-3. The best three values of the EMSE that we obtained at the simulation are shown in bold. According to the simulation, we conclude the following results from the Table 1-2-3:

(1) In all the combinations of n, p, ρ and ϕ (total 81 scenarios), the estimators we suggested had smaller EMSE values than the existing estimators that we compared.
(2) As it can be seen from Table 1-2-3, generally, BLTE1 has the best EMSE value in many combinations. When the number of independent variables in the model was relatively high, BLTE1 has the smallest SMSE value in many cases. When the number of variables is relatively small, BLTE3 has the best EMSE value.
(3) In the relatively lower correlation, the BLTE1 had a smaller SMSE, while the BLTE2 and BLTE3 had a smaller EMSE in the highly correlated models.
(4) While the BLTE1 had smaller EMSE values at small ϕ values, in general, BLTE2 and BLTE3 had better EMSEs at higher ϕ values.
(5) In general, it was observed that the EMSE values of BRE2 and our proposed estimators tended to decrease, while the EMSE values of the other estimators tend to increase in the case of high correlation.

Table 2. The EMSE values of the estimators when $p = 8$.

n	ρ	ϕ	**MLE**	**BRE1**	**BRE2**	**BLE**	**LTBR**	**TPBR**	**BDK**	**BLTE1**	**BLTE2**	**BLTE3**
50	0.9	0.5	93.8278	11.2662	1.1812	1.1674	40.4812	2.05	64.71	**0.6514**	**0.7142**	**0.7913**
50	0.9	1.5	61.9516	4.315	1.1672	0.7503	17.1439	0.9192	44.7066	**0.52**	**0.5388**	**0.5652**
50	0.9	5	17.1674	0.8987	1.2349	0.904	2.5837	0.8186	9.3804	**0.6232**	**0.6371**	**0.6649**
50	0.95	0.5	136.5851	17.0555	1.0508	1.0397	63.5861	1.7514	106.1478	**0.6569**	**0.7733**	**0.7987**
50	0.95	1.5	76.5216	6.6043	0.9974	0.6817	21.2708	0.941	58.699	**0.4837**	**0.5068**	**0.5232**
50	0.95	5	27.1776	0.9881	1.0105	0.7687	3.7101	0.757	18.1247	**0.5687**	**0.5792**	**0.5902**
50	0.99	0.5	763.301	129.6957	0.5593	1.3888	312.7882	1.7598	728.1236	**0.4836**	**0.5294**	**0.4955**
50	0.99	1.5	479.6771	52.6478	0.5322	0.595	129.0926	0.7633	459.0657	**0.4473**	**0.4417**	**0.4352**
50	0.99	5	249.6432	10.8579	0.6105	0.5603	39.7353	0.6455	238.4401	**0.5564**	**0.553**	**0.5378**
100	0.9	0.5	50.9754	3.6385	1.3679	1.324	18.8995	1.3868	27.2484	**0.6508**	**0.6521**	**0.7596**
100	0.9	1.5	39.136	2.1362	1.2934	0.8491	9.57	0.8521	23.9414	**0.5276**	**0.5311**	**0.5641**
100	0.9	5	15.0022	0.8301	1.2009	0.8971	1.7139	0.8051	7.4092	**0.5831**	**0.5782**	**0.5766**
100	0.95	0.5	152.7753	15.5225	0.8331	0.9196	56.8754	1.5586	121.4166	**0.5296**	**0.5668**	**0.584**
100	0.95	1.5	75.093	5.2406	0.9504	0.6762	17.7288	0.8988	57.5188	**0.4731**	**0.4743**	**0.482**
100	0.95	5	29.2438	1.0334	0.9738	0.742	3.3904	0.7545	20.1553	**0.5663**	**0.5642**	**0.5712**
100	0.99	0.5	774.0129	114.2618	0.5404	1.2806	292.3458	1.4223	740.0923	**0.4818**	**0.5055**	**0.4864**
100	0.99	1.5	411.2774	39.235	0.5149	0.5392	92.8988	0.7507	391.0896	**0.4327**	**0.4122**	**0.4135**
100	0.99	5	179.9447	6.0944	0.6117	0.5839	23.3464	0.6517	168.6955	**0.5467**	**0.5273**	**0.5241**
200	0.9	0.5	68.1933	3.9271	1.1693	1.1879	22.695	1.3301	41.6812	**0.5647**	**0.5666**	**0.5889**
200	0.9	1.5	41.2022	1.9844	1.2116	0.814	8.8423	0.8552	25.5893	**0.4846**	**0.4874**	**0.4829**
200	0.9	5	18.7176	0.8346	1.0804	0.8095	1.8454	0.7905	10.6093	**0.5575**	**0.5502**	**0.538**
200	0.95	0.5	117.1055	8.2409	0.8906	0.9559	39.0465	1.4058	87.7555	**0.5137**	**0.5176**	**0.5307**
200	0.95	1.5	69.3549	3.6225	0.9341	0.6707	14.1324	0.8124	52.2297	**0.4586**	**0.455**	**0.4479**
200	0.95	5	29.854	0.9163	0.9438	0.7206	2.6801	0.7565	20.9637	**0.5633**	**0.5505**	**0.5427**
200	0.99	0.5	552.9092	52.5662	0.5955	0.8861	179.1318	1.2149	519.5047	**0.5074**	**0.4998**	**0.4987**
200	0.99	1.5	414.0522	33.7499	0.5048	0.5169	85.3423	0.6915	393.8937	**0.4281**	**0.4034**	**0.405**
200	0.99	5	150.4005	2.6752	0.626	0.5648	11.3286	0.6521	139.6566	**0.5519**	**0.5173**	**0.5182**

5. Real data application

To further examine and to show the practicality of our new proposed estimators, we apply the proposed estimators to the Australian Institute of Sport Data (AIS). The AIS data set has also been used by [24]. Also the data is included in R library sn. Here, percentage body fat (y_i), hematocrit in percent ($\boldsymbol{x_1}$) and hemoglobin concentration in gram per deciliter ($\boldsymbol{x_2}$) are selected for rowing athletes. The correlation of independent variables is equal to 0.963. Table 4 shows the results of these data for the logit link function. According to the data set, for used link functions, when the variance value of all estimators is calculated (last row in the Table 4), our proposed estimators have been quite successful. Specifically BLTE1 and BLTE3 outperformed all other predictors. Since BLTE2 has approximately similar variance value with BLE, it has proven to be successful compared to other estimators.

In addition, the bootstrap sampling method is used to calculate the SMSE values of the relevant estimators. For this reason, 10000 bootstrap samples have been created. For each of these samples, the parameter estimates of the proposed and existing biased estimators are calculated. The mean of the MLE estimates is considered as the real parameter and the calculated SMSE values are given in Table 4.

Table 3. The EMSE values of the estimators when $p = 12$.

n	ρ	ϕ	**MLE**	**BRE1**	**BRE2**	**BLE**	**LTBR**	**TPBR**	**BDK**	**BLTE1**	**BLTE2**	**BLTE3**
50	0.9	0.5	126.7153	9.512	1.7133	1.3808	55.5568	1.9904	84.4768	**0.6313**	**0.7801**	**0.9935**
50	0.9	1.5	100.8498	6.0594	1.5691	0.888	32.6452	0.9716	75.0141	**0.5265**	**0.6424**	**0.7486**
50	0.9	5	30.597	0.917	1.4962	1.0164	4.8008	0.7925	18.0153	**0.5991**	**0.6834**	**0.7388**
50	0.95	0.5	216.5942	19.6176	1.3518	1.0943	95.9499	2.1848	170.5494	**0.5837**	**0.739**	**0.8552**
50	0.95	1.5	172.3998	14.9221	1.2862	0.7565	65.154	1.0246	144.4295	**0.5251**	**0.6813**	**0.7326**
50	0.95	5	73.5961	1.8245	1.0166	0.7468	13.592	0.7115	58.0129	**0.5293**	**0.623**	**0.6288**
50	0.99	0.5	1417.9849	266.4884	0.5429	1.1105	675.2927	2.8763	1364.0739	**0.421**	**0.4999**	**0.4719**
50	0.99	1.5	782.4249	113.6931	0.543	0.4709	276.8606	1.0396	750.8841	**0.3928**	**0.4274**	**0.4094**
50	0.99	5	289.7669	8.4269	0.6639	0.5683	46.1859	0.6438	271.8784	**0.5334**	**0.5406**	**0.5265**
100	0.9	0.5	95.9181	4.9376	1.7149	1.6055	38.1135	1.4696	56.9282	**0.6755**	**0.7923**	**1.0414**
100	0.9	1.5	59.0279	2.0447	1.7635	1.0464	14.888	0.858	36.1809	**0.5067**	**0.5465**	**0.6515**
100	0.9	5	36.781	0.8521	1.2634	0.9018	4.0019	0.7679	22.9186	**0.5357**	**0.5442**	**0.5663**
100	0.95	0.5	213.4928	15.3818	1.0739	1.0119	77.6324	1.8604	167.9322	**0.4965**	**0.5626**	**0.627**
100	0.95	1.5	133.0176	7.5678	1.0691	0.6688	35.2159	0.837	106.4537	**0.4237**	**0.4667**	**0.5138**
100	0.95	5	53.3172	1.193	1.1257	0.8094	7.9374	0.7429	38.6889	**0.5542**	**0.6035**	**0.6316**
100	0.99	0.5	1127.7097	149.4226	0.5309	0.6823	443.8274	1.9407	1075.9271	**0.4294**	**0.4692**	**0.459**
100	0.99	1.5	740.3852	64.6641	0.5291	0.4711	188.2876	0.848	709.434	**0.4079**	**0.4001**	**0.3975**
100	0.99	5	254.5988	5.5096	0.6397	0.553	30.7731	0.6383	237.2881	**0.5204**	**0.5025**	**0.4995**
200	0.9	0.5	91.9788	3.1418	1.528	1.534	29.7031	1.2606	53.6285	**0.5367**	**0.5734**	**0.6844**
200	0.9	1.5	57.0663	1.6173	1.6664	1.0217	12.2218	0.8329	34.1209	**0.4597**	**0.4692**	**0.5249**
200	0.9	5	26.3901	0.8229	1.4629	1.0154	2.6239	0.7961	13.8923	**0.5581**	**0.5575**	**0.5819**
200	0.95	0.5	184.8183	8.7547	1.0716	1.0746	60.0984	1.5148	140.575	**0.4957**	**0.5181**	**0.561**
200	0.95	1.5	109.8284	3.9556	1.1764	0.7619	22.7946	0.8434	83.4023	**0.4407**	**0.4427**	**0.4663**
200	0.95	5	54.5301	0.9384	1.0891	0.7911	4.9765	0.7373	39.5982	**0.5446**	**0.5319**	**0.5393**
200	0.99	0.5	926.5518	79.5587	0.5613	0.6892	305.3659	1.5674	876.8201	**0.4475**	**0.4483**	**0.4489**
200	0.99	1.5	572.5614	37.8564	0.5309	0.4565	122.2598	0.8043	541.8656	**0.3949**	**0.3743**	**0.374**
200	0.99	5	253.2151	4.2294	0.6361	0.557	22.9285	0.6428	236.0445	**0.5272**	**0.4972**	**0.4968**

Table 4. The estimated coefficients and SMSE values for the Gasoline yield data.

	MLE	**BRE1**	**BRE2**	**BLE**	**LTBR**	**TPBR**	**BDK**	**BLTE1**	**BLTE2**	**BLTE3**
$\hat{\beta}_1$	-1.7157	-1.7156	-1.6766	-1.4037	-1.7128	-1.7103	-1.6443	-1.2208	-1.5305	-1.3222
$\hat{\beta}_2$	0.701	0.6852	-0.5322	-0.1178	0.1024	-0.0895	-1.7768	-0.0707	-0.2531	-0.1026
$\hat{\beta}_3$	-2.354	-2.3374	-0.708	-0.183	-1.7259	-1.5051	0.8863	-0.0674	-0.2448	-0.0976
$var\left(\hat{\beta}\right)$	3.2013	3.1346	0.0515	0.0068	1.19517	0.74370	2.3549	0.0025	0.0096	0.0035
$SMSE\left(\hat{\beta}\right)$	3.2013	3.13462	0.0522	0.0827	1.19519	0.74373	2.3572	0.2112	0.0327	0.1298

6. Conclusion

In this paper, the BLTE is proposed to combat the multicollinearity problem in the BRMs. It is given the properties of the new biased estimators and shown its superiority than the other estimators we compared. According to Monte Carlo simulation studies, BLTEs have better performance than BML, BRE, BLE, LTBR, TPBR and BDK estimators, in terms of EMSE. Especially, BLTE1 provided superiority in lower correlation and lower precision values while BLTE2 and BLTE3 outperformed superiority in high correlation and high precision values. A numerical example is given to evaluate the performance of our proposed estimator. The obtained results are consistent with the simulation results. Therefore, we recommend BLTEs to researchers to use in their studies.

Acknowledgment. The authors wish to thank the referees for several suggestions which helped to improve the quality of the presentation.

References

[1] M.R. Abonazel, Z.Y. Algamal, F.A. Awwad, and I.M. Taha, *A new two-parameter estimator for beta regression model: method, simulation, and application*, Front. Appl. Math. Stat. **7**, 780322, 1–10, 2022.

[2] M.R. Abonazel, I. Dawoud, F.A. Awwad, and A.F. Lukman, *DawoudKibria estimator for Beta regression model: simulation and application*, Front. Appl. Math. Stat. **8**, 775068, 1-12, 2022.

[3] M.R. Abonazel and I.M. Taha, *Beta ridge regression estimators: simulation and application*, Comm. Statist. Simulation Comput., Doi: 10.1080/03610918.2021.1960373, 2021.

[4] K.U. Akay and E. Ertan, *A new Liu-type estimator in Poisson regression models*, Hacet. J. Math. Stat. **51** (5), 1484-1503, 2022.

[5] M.N. Akram, M. Amin, A. Elhassanein, and M.A. Ullah, *A new modified ridge-type estimator for the beta regression model: simulation and application*, AIMS Math. **7** (1), 10351057.

[6] Z.Y. Algamal, *Diagnostic in poisson regression models*, Electron. J. Appl. Stat. Anal. **5** (2), 178-186, 2012.

[7] Z.Y. Algamal, *Performance of ridge estimator in inverse Gaussian regression model*, Comm. Statist. Theory Methods **48** (15), 3836-3849, 2019.

[8] Z.Y. Algamal and M.R. Abonazel, *Developing a Liu-type estimator in beta regression model*, Concurr. Comp.-Pract. E. **34**, e6685, 1-11, 2021.

[9] Z.Y. Algamal and M.M. Alanaz, *Proposed methods in estimating the ridge regression parameter in Poisson regression model*, Electron. J. Appl. Stat. Anal. **11** (2), 506-515, 2018.

[10] Z.Y. Algamal and Y. Asar, *Liu-type estimator for the gamma regression model*, Comm. Statist. Simulation Comput. **49** (8), 2035-2048, 2020.

[11] I. Dawoud and B.M.G. Kibria, *A new biased estimator to combat the multicollinearity of the gaussian linear regression model*, Stats **3** (4), 526-541, 2020.

[12] E. Ertan and K.U. Akay, *A new Liu-type estimator in binary logistic regression models*, Comm. Statist. Theory Methods **51** (13), 4370-4394, 2022.

[13] P.L. Espinheira, S.L.P. Ferrari, and F. Cribari-Neto, *On beta regression residuals*, J. Appl. Stat. **35** (4), 407-419, 2008.

[14] P.L. Espinheira, L.C.M. Silva, and A.D.O. Silva, *Prediction measures in beta regression models*, arXiv: 1501.04830 [stat.AP].

[15] P.L. Espinheira, L.C.M. Silva, A.D.O. Silva, and R. Ospina, *Model selection criteria on beta regression for machine learning*, Mach. Learn. Knowl. Extr. **1** (1), 427-449, 2019.

[16] S. Ferrari and F. Cribari-Neto, *Beta regression for modeling rates and proportions*, J. Appl. Stat. **31** (7), 799815, 2004.

[17] A.E. Hoerl and R.W. Kennard, *Ridge regression: Biased estimation for nonorthogonal problems*, Technometrics, **12** (1), 55-67, 1970.

[18] P. Karlsson, K. Månsson and B.M.G. Kibria, *A Liu estimator for the beta regression model and its application to chemical data*, J. Chemom. **34** (10), 2020.

[19] B.M.G. Kibria, *Some Liu and ridge-type estimators and their properties under the Ill-conditioned Gaussian linear regression model*, J. Stat. Comput. Simul. **82** (1), 1-17, 2012.

[20] K. Liu, *A new class of biased estimate in linear regression*, Comm. Statist. Theory Methods **22** (2), 393-402, 1993.

[21] K. Liu, *Using Liu-type estimator to combat collinearity*, Comm. Statist. Theory Methods **32** (5), 1009-1020, 2003.

[22] K. Månsson and B.M.G. Kibria, *Estimating the unrestricted and restricted Liu estimators for the Poisson regression model: method and application*, Comput. Econ. **58** (2), 311-326, 2021.

[23] K. Månsson, B.M.G. Kibria, and G. Shukur, *Some Liu Type Estimators for the dynamic OLS estimator: with an application to the carbon dioxide Kuznets curve for Turkey*, Commun. Stat. Case Stud. Data Anal. Appl. **3** (3-4), 55-61, 2017.

[24] S. Pirmohammadi and H. Bidram, *On the Liu estimator in the beta and Kumaraswamy regression models: A comparative study*, Comm. Statist. Theory Methods **51** (24), 8553-8578, 2021.

[25] M. Qasim, M. Amin and T. Omer, *Performance of some new Liu parameters for the linear regression model*, Comm. Statist. Theory Methods **49** (17), 41784196, 2020.

[26] M. Qasim, K. Månsson, and B.M.G. Kibria, *On some beta ridge regression estimators: method, simulation and application*, J. Stat. Comput. Simul. **91** (9), 1699-1712, 2021.

[27] N.K. Rashad and Z.Y. Algamal, *A new ridge estimator for the Poisson regression model*, Iran. J. Sci. Technol. Trans. A: Sci. **43** (6), 2921-2928, 2019.