Hindawi
Journal of Mathematics
Volume 2024, Article ID 6694880, 12 pages
https://doi.org/10.1155/2024/6694880

Research Article

Reducing Bias in Beta Regression Models Using Jackknifed Liu-Type Estimators: Applications to Chemical Data

Solmaz Seifollahi,[1] **Hossein Bevrani**,[1] **and Olayan Albalawi**[2]

[1]*Department of Statistics, University of Tabriz, Tabriz, Iran*
[2]*Department of Statistics, Faculty of Science, University of Tabuk, Tabuk 47512, Saudi Arabia*

Correspondence should be addressed to Hossein Bevrani; bevrani@gmail.com

Received 26 July 2023; Revised 31 December 2023; Accepted 17 January 2024; Published 27 January 2024

Academic Editor: Niansheng Tang

In the field of chemical data modeling, it is common to encounter response variables that are constrained to the interval (0, 1). In such cases, the beta regression model is often a more suitable choice for modeling. However, like any regression model, collinearity can present a significant challenge. To address this issue, the Liu-type estimator has been used as an alternative to the maximum likelihood estimator, but it suffers from bias. In this paper, we introduce the Jackknifed Liu-type estimator and its modified version, which demonstrate improved bias reduction compared to the original Liu-type estimator. We assess the theoretical and numerical performance of these estimators through Monte Carlo simulations and real-data examples from the field of chemistry. Our findings highlight the significant improvements offered by the proposed estimators in terms of accuracy and reliability.

1. Introduction

Regression models are widely employed in various fields, including chemometrics, for modeling data (see [1–4], for example). Different types of regression models, such as linear, generalized linear, nonlinear, and nonparametric regression models, have been introduced. However, selecting the appropriate model is crucial to obtain reliable and precise results. The nature and distribution of the response variable should be carefully considered when choosing a regression model.

In certain areas of research, the possible values of the response variable are limited to the interval (0, 1), such as rates and proportions. To address this, the beta regression model (BRM) was introduced by Ferrari and Cribari-Neto [5]. However, the quality of parameter estimation has a significant impact on the use of BRMs. The maximum likelihood estimation (MLE) is commonly used for parameter estimation in BRMs, but if the independent variables are ill-conditioned, the results may be unsatisfactory. The variable selection methods, such as adjusted R^2 [6] or the swarm optimization method [7], can help to deal with this issue. However, this issue often necessitates the use of biased estimators, such as Stein-type estimators [8], ridge estimators [9, 10], modified ridge-type estimators [11], Liu estimators [12, 13], two-parameter estimators [14], Dawoud-Kibria estimators [15], and also Liu-type estimators [16, 17] which is of particular interest in this paper. Furthermore, there has been recent interest in a class of almost unbiased estimators for parameter estimation in regression models based on biased estimators. Notable examples of these estimators include the work of Ohtani [18], Amin et al. [19], Wu and Asar [20, 21], Varathan and Wijekoon [22], and Asar and Korkmaz [23]. On the other hand, the Jackknife approach has emerged as a viable method for reducing estimator bias. The Jackknife approach was initially developed by Quenouille [9] and Tukey [24] to significantly reduce the bias of estimators. Singh et al. [25] suggested an unbiased ridge estimator for linear regression models using this technique. Later, Batah et al. [26] suggested a modified Jackknifed ridge estimator in linear regression models and demonstrated its superiority over the generalized ridge estimator, Jackknifed ridge estimator, and LASSO [27]. In gamma regression models, Algamal [28, 29] employed the

Jackknife technique to mitigate the bias of the ridge estimator. Yildiz [30] and Chaubey et al. [31] both presented Jackknifed Liu-type estimators and conducted theoretical and numerical analyses to explore their properties. In this paper, we apply the Jackknife technique to minimize the bias of the Liu-type estimator in BRMs.

The paper is organized as follows: Section 2 provides an overview of the BRM and determines the MLE of the parameters. In Section 3, the Jackknife procedure is applied to the Liu-type estimator and its modified version is introduced. The properties of the proposed estimators such as bias, covariance, and the means squared error are determined in Section 4. A theoretical comparison of the estimators is presented in Section 5, followed by a Monte Carlo simulation experiment in Section 6 to evaluate their performance. Section 7 showcases two applications of the proposed estimator in chemometrics. Finally, Section 8 presents the conclusions of the study.

2. Beta Regression Model

BRMs are commonly employed for analyzing data that is expressed as proportions and rates, such as migration rates and unemployment rates. These models rely on the fundamental assumption that the response variable follows beta$(\mu\tau, (1-\mu)\tau)$; e.g.,

$$f(y\mid\mu,\tau)=\frac{\Gamma(\tau)}{\Gamma(\mu\tau)\Gamma((1-\mu)\tau)}y^{\mu\tau-1}(1-y)^{(1-\mu)\tau-1};\quad 0<\mu, y<1, \tau>0. \tag{1}$$

The model, introduced by Ferrari and Cribari-Neto [5], assumes a constant precision parameter, τ, over observations. Let $y_1, y_2, \ldots, y_n$ represent the response observations with the density defined in equation (1). The regression model can be expressed as follows:

$$g(\mu_i)=\mathbf{x}_i^T\boldsymbol{\beta}=\eta_i;\quad i: 1, 2, \ldots, n, \tag{2}$$

where $\boldsymbol{\beta}=(\beta_0,\beta_1,\beta_2,\ldots,\beta_p)^T\in\mathbb{R}^p$ and $\mathbf{x}_i=(1,x_{i1},x_{i2},\ldots,x_{ip})^T$ is the i th observation of the independent variables and the link function, $g(.)$ which maps (0,1) into $\mathbb{R}$ is a continuous and double differentiable function. To derive the MLE of parameters, $\boldsymbol{\beta}$, one can use the iterative re-weighted least-squares algorithm. Let $y^*=(y_1^*,y_2^*,\ldots,y_n^*)^T$, $\mu^*=(\mu_1^*,\mu_2^*,\ldots,\mu_n^*)^T$, $y_i^*=\text{logit}(y_i)$, and $\mu_i^*=\psi(\mu_i\tau)-\psi((1-\mu_i)\tau)$, where $\psi(.)$ denotes the digamma function. Therefore, the MLE in the BRM will be [9, 10]

$$\widehat{\beta}_{\text{BMLE}}=\left(\mathbf{X}^T\widehat{\mathbf{V}}\mathbf{X}\right)^{-1}\mathbf{X}^T\widehat{\mathbf{V}}\mathbf{Z}, \tag{3}$$

where

$$\begin{aligned}\hat{\mathbf{V}}&=\text{diag}\left(\hat{v}_1,\cdots,\hat{v}_n\right),\\ \hat{v}_i&=\tau\left\{\psi'\left(\hat{\mu}_i\tau\right)-\psi'\left(\left(1-\hat{\mu}_i\right)\tau\right)\right\}\frac{1}{\left\{g'\left(\hat{\mu}_i\right)\right\}^2},\\ \hat{\mathbf{Z}}&=\mathbf{X}^T\hat{\boldsymbol{\beta}}+\hat{\mathbf{V}}^{-1}\frac{\left(y^*-\mu^*\right)}{\left\{g'(\hat{\mu})\right\}^2}.\end{aligned} \tag{4}$$

The values of $\widehat{\mathbf{V}}$ and $\widehat{\mathbf{Z}}$ are evaluated at the final iteration. As n increases, the distribution of $\widehat{\boldsymbol{\beta}}_{\text{BMLE}}$ approaches normal distribution with the mean vector $\boldsymbol{\beta}$ and covariance matrix $1/\tau(\mathbf{X}^T\widehat{\mathbf{V}}\mathbf{X})^{-1}$. As a consequence, the scalar mean squared error (MSE) of $\widehat{\beta}_{\text{BMLE}}$ can be represented as

$$\text{MSE}\left(\widehat{\beta}_{\text{BMLE}}\right)=\frac{1}{\tau}\text{tr}\left(\mathbf{C}^{-1}\right)=\frac{1}{\tau}\sum_{k=1}^{p+1}\frac{1}{q_k}, \tag{5}$$

where q_k is the k th eigenvalue of $\mathbf{C}=\mathbf{X}^T\widehat{\mathbf{V}}\mathbf{X}$.

It should come as no surprise that the ill-condition of the matrix $\mathbf{X}^T\widehat{\mathbf{V}}\mathbf{X}$ negatively impacts both the variance of MLE and the accuracy of parameter estimates. To overcome this issue in estimating parameters, Qasim et al. [12] and Abonazel and Taha [10] proposed the beta ridge estimator (BRE) as follows:

$$\widehat{\beta}_{\text{BRE}}=\left(\mathbf{C}+\lambda I_p\right)^{-1}\mathbf{C}\widehat{\beta}_{\text{BMLE}},\quad \lambda>0. \tag{6}$$

The beta Liu estimator (BLE) is proposed by Karlsson et al. [32] as

$$\widehat{\beta}_{\text{BLE}}=\left(\mathbf{C}+I_p\right)^{-1}\left(\mathbf{C}+dI_p\right)\widehat{\beta}_{\text{BMLE}},\quad 0<d<1, \tag{7}$$

and also, the beta Liu-type estimator (BLTE) is proposed by Algamal and Abonazel [16] as

$$\begin{aligned}\widehat{\beta}_{\text{BLTE}}&=\left(\mathbf{C}+\lambda I_p\right)^{-1}\left(\mathbf{C}-dI_p\right)\mathbf{C}^{-1}\mathbf{X}^T\widehat{V}Z\\ &=\left(\mathbf{C}+\lambda I_p\right)^{-1}\left(\mathbf{C}-dI_p\right)\widehat{\beta}_{\text{BMLE}}\\ &=\left[I_p-(\lambda+d)\left(\mathbf{C}+\lambda I_p\right)^{-1}\right]\widehat{\beta}_{BMLE},\end{aligned} \tag{8}$$

where $\lambda>0$ and $d\in\mathbb{R}$. The beta Liu-type estimator (BLTE) encompasses the beta ridge estimator (BRE) and beta Liu estimator (BLE) as specific cases. Algamal and Abonazel [16] demonstrated that the BLTE outperforms both of these estimators.

3. Suggested Estimators

The bias of BLTE is given by

$$\begin{aligned}\mathbb{B}\left(\widehat{\beta}_{\text{BLTE}}\right)&=\mathbb{E}\left(\widehat{\beta}_{\text{BLTE}}\right)-\boldsymbol{\beta}\\ &=\left(\mathbf{C}+\lambda I_p\right)^{-1}\left(\mathbf{C}-dI_p\right)\boldsymbol{\beta}-\boldsymbol{\beta}\\ &=-(d+\lambda)\left(\mathbf{C}+\lambda I_p\right)^{-1}\boldsymbol{\beta}.\end{aligned} \tag{9}$$

The substantial bias exhibited by the BLTE is an undesirable characteristic for researchers. To mitigate this bias, the Jackknife approach was developed by Quenouille [9] and Tukey [24] to significantly reduce the bias of estimators. In this section, we will follow the approach of Singh et al. [25] and Batah et al. [26] to derive the Jackknifed form of the BLTE. In addition, we present a modified version of this estimator. If we delete the j th observation, $(\mathbf{x}_j, y_j)$, from the data then

$$\widehat{\beta}_{\mathrm{BLTE}(-j)} = \left(\mathbf{C}_{(-j)} + \lambda I_p\right)^{-1}\left(\mathbf{C}_{(-j)} - dI_p\right)\widehat{\beta}_{\mathrm{BMLE}_{(-j)}}. \quad (10)$$

After some algebraic manipulations, we have

$$\widehat{\beta}_{\mathrm{BLTE}(-j)} = \left[\mathscr{M} - \mathbf{x}_j\hat{V}_j\mathbf{x}_j^T\right]^{-1}\left(\mathbf{X}^T\hat{\mathbf{V}}\mathbf{Z} - \mathbf{x}_j^T\hat{V}_j z_j\right), \quad (11)$$

where

$$\mathscr{M}^{-1} = \left(\mathbf{C} + \lambda I_p\right)^{-1}\left(\mathbf{C} - dI_p\right)\mathbf{C}^{-1}. \quad (12)$$

Thus,

$$\begin{aligned}
\widehat{\beta}_{\mathrm{BLTE}(-j)} &= \left[\mathscr{M}^{-1} + \frac{\mathscr{M}^{-1}\mathbf{x}_j\widehat{V}_j\mathbf{x}_j^T\mathscr{M}^{-1}}{1 - \mathbf{x}_j^T\mathscr{M}^{-1}\widehat{V}_j\mathbf{x}_j}\right]\left(\mathbf{X}^T\hat{\mathbf{V}}\mathbf{Z} - \mathbf{x}_j^T\widehat{V}_j z_j\right) \\
&= \mathscr{M}^{-1}\mathbf{X}^T\hat{\mathbf{V}}\mathbf{Z} - \mathscr{M}^{-1}\mathbf{x}_j\widehat{V}_j z_j\left[1 + \frac{\mathbf{x}_j\mathscr{M}^{-1}\widehat{V}_j\mathbf{x}_j^T}{1 - \mathbf{x}_j^T\mathscr{M}^{-1}\widehat{V}_j\mathbf{x}_j}\right] \\
&\quad + \frac{\mathscr{M}^{-1}\mathbf{x}_j\widehat{V}_j\mathbf{x}_j^T\mathscr{M}^{-1}}{1 - \mathbf{x}_j^T\mathscr{M}^{-1}\widehat{V}_j\mathbf{x}_j}\mathscr{M}^{-1}\mathbf{X}^T\hat{\mathbf{V}}\mathbf{Z} \\
&= \widehat{\beta}_{\mathrm{BLTE}} - \frac{\mathbf{x}_j\mathscr{M}^{-1}\widehat{V}_j\mathbf{x}_j^T}{1 - \mathbf{x}_j^T\mathscr{M}^{-1}\widehat{V}_j\mathbf{x}_j} + \frac{\mathscr{M}^{-1}\mathbf{x}_j\widehat{V}_j\mathbf{x}_j^T\mathscr{M}^{-1}}{1 - \mathbf{x}_j^T\mathscr{M}^{-1}\widehat{V}_j\mathbf{x}_j}\widehat{\beta}_{\mathrm{BLTE}} \\
&= \widehat{\beta}_{\mathrm{BLTE}} - \frac{\mathscr{M}^{-1}\mathbf{x}_j\widehat{V}_j e_j}{1 - w_j},
\end{aligned} \quad (13)$$

where $e_j = z_j - \mathbf{x}_j^T\widehat{\beta}_{\mathrm{BLTE}}$ and $w_j = 1 - \mathbf{x}_j^T\mathscr{M}^{-1}\widehat{V}_j\mathbf{x}_j$. We consider weighted pseudovalues in the weighted Jackknife method as

$$u_j = \widehat{\beta}_{\mathrm{BLTE}} + n\left(1 - w_j\right)\left[\widehat{\beta}_{\mathrm{BLTE}} - \widehat{\beta}_{\mathrm{BLTE}(-j)}\right]. \quad (14)$$

Therefore, the weighted Jackknifed estimator will be

$$\widehat{\beta}_{\mathrm{BJLTE}} = \frac{1}{n}\sum_{j=1}^{n} u_j = \widehat{\beta}_{\mathrm{BLTE}} + \mathscr{M}^{-1}\sum_{j=1}^{n}\mathbf{x}_j\widehat{V}_j e_j, \quad (15)$$

and

$$\begin{aligned}
\sum_{j=1}^{n}\mathbf{x}_j\widehat{V}_j e_j &= \sum_{j=1}^{n}\mathbf{x}_j\widehat{V}_j\left[z_j - \mathbf{x}_j^T\widehat{\beta}_{\mathrm{BLTE}}\right] \\
&= \sum_{j=1}^{n}\mathbf{x}_j\widehat{V}_j z_j - \sum_{j=1}^{n}\mathbf{x}_j\widehat{V}_j\mathbf{x}_j^T\widehat{\beta}_{\mathrm{BLTE}} \\
&= \mathbf{X}^T\widehat{V}Z - \mathbf{C}\widehat{\beta}_{\mathrm{BLTE}}.
\end{aligned} \quad (16)$$

By replacing (16) in (15), the beta Jackknifed Liu-type estimator (BJLTE) is given by

$$\begin{aligned}
\hat{\boldsymbol{\beta}}_{\mathrm{BJLTE}} &= \hat{\boldsymbol{\beta}}_{\mathrm{BLTE}} + \mathscr{M}^{-1}\mathbf{X}^T\hat{\mathbf{V}}\mathbf{Z} - \mathscr{M}^{-1}\mathbf{C}\hat{\boldsymbol{\beta}}_{\mathrm{BLTE}} \\
&= \left[2I_p - \mathscr{M}^{-1}\mathbf{C}\right]\hat{\boldsymbol{\beta}}_{\mathrm{BLTE}} \\
&= \left[2I_p - \left(\mathbf{C} + \lambda I_p\right)^{-1}\left(\mathbf{C} - dI_p\right)\right]\hat{\boldsymbol{\beta}}_{\mathrm{BLTE}} \\
&= \left[I_p - (\lambda + d)^2\left(\mathbf{C} + \lambda I_p\right)^{-2}\right]\hat{\boldsymbol{\beta}}_{\mathrm{BMLE}}.
\end{aligned} \quad (17)$$

We also define a modified version of the BJLTE which is obtained by replacing $\widehat{\beta}_{\mathrm{BLTE}}$ instead of $\widehat{\beta}_{\mathrm{BMLE}}$, that is

$$\begin{aligned}
\hat{\boldsymbol{\beta}}_{\mathrm{BMJLTE}} &= \left[I_p - (\lambda + d)^2\left(\mathbf{C} + \lambda I_p\right)^{-2}\right]\hat{\boldsymbol{\beta}}_{\mathrm{BLTE}} \\
&= \left[I_p - (\lambda + d)^2\left(\mathbf{C} + \lambda I_p\right)^{-2}\right]\left[I_p - (\lambda + d)\left(\mathbf{C} + \lambda I_p\right)^{-1}\right]\hat{\boldsymbol{\beta}}_{\mathrm{BMLE}}.
\end{aligned} \quad (18)$$

4. Properties of the Estimators

To simplify the analysis and understand the properties of the estimators, we derive their canonical form. Consider q_i for $i = 1,2,\ldots,p+1$ as the eigenvalues of matrix $\mathbf{C}$ such that $\mathbf{C} = \mathbf{E}^T\mathbf{Q}\mathbf{E}$, where $\mathbf{Q} = \mathrm{diag}(q_1, q_2, \ldots, q_{p+1})$ and $\mathbf{E} = [e_1, e_2, \ldots, e_{p+1}]$ which the columns are the eigenvector of $\mathbf{C}$. So, the canonical forms can be used to express in the terms of $\mathbf{U} = \mathbf{X}\mathbf{E}$ and $\boldsymbol{\gamma} = \mathbf{E}^T\boldsymbol{\beta}$. Specifically, we denote the canonical form of BMLE as $\hat{\boldsymbol{\gamma}}_{\mathrm{BMLE}}$.

4.1. Properties of BLTE. The canonical form of BLTE of $\boldsymbol{\gamma}$ is given by

$$\hat{\boldsymbol{\gamma}}_{\mathrm{BLTE}} = \mathbf{A}_\lambda^{-1}\left(\mathbf{Q} - dI_p\right)\hat{\boldsymbol{\gamma}}_{\mathrm{BMLE}}; \quad \lambda > 0, d \in \mathbb{R}, \quad (19)$$

where $\mathbf{A}_\lambda = \mathbf{Q} + \lambda I_p$. The bias and covariance of $\hat{\boldsymbol{\gamma}}_{\mathrm{BLTE}}$ will be

$$\mathbb{B}\left(\hat{\boldsymbol{\gamma}}_{\mathrm{BLTE}}\right) = -(d + \lambda)\mathbf{A}_\lambda^{-1}\boldsymbol{\gamma}, \quad (20)$$

$$\mathrm{Cov}\left(\hat{\boldsymbol{\gamma}}_{\mathrm{BLTE}}\right) = \frac{1}{\tau}\mathbf{A}_\lambda^{-1}\left(\mathbf{Q} - dI_p\right)\mathbf{Q}^{-1}\left(\mathbf{Q} - dI_p\right)^T\mathbf{A}_\lambda^{-1}. \quad (21)$$

The matrix MSE (MMSE) and scalar MSE (SMSE) are given, respectively, by

$$\mathrm{MMSE}\left(\hat{\boldsymbol{\gamma}}_{\mathrm{BLTE}}\right) = \frac{1}{\tau}\mathbf{A}_{\lambda}^{-1}\left(\mathbf{Q} - dI_p\right)\mathbf{Q}^{-1}\left(\mathbf{Q} - dI_p\right)^T\mathbf{A}_{\lambda}^{-1} + (d+\lambda)^2\mathbf{A}_{\lambda}^{-1}\boldsymbol{\gamma}\boldsymbol{\gamma}^T\mathbf{A}_{\lambda}^{-1},$$
$$\mathrm{SMSE}\left(\hat{\boldsymbol{\gamma}}_{\mathrm{BLTE}}\right) = \frac{1}{\tau}\sum_{i=1}^{p+1}\frac{\left(q_i - d\right)^2}{q_i\left(q_i+\lambda\right)^2} + (d+\lambda)^2\sum_{i=1}^{p+1}\frac{\gamma_i^2}{\left(q_i+\lambda\right)^2}. \tag{22}$$

4.2. Properties of BJLTE. The canonical form of BJLTE of $\boldsymbol{\gamma}$ is given by

$$\hat{\boldsymbol{\gamma}}_{\mathrm{BJLTE}} = \mathbf{B}_2\hat{\boldsymbol{\gamma}}_{\mathrm{BMLE}}; \quad \lambda > 0, d \in \mathbb{R}, \tag{23}$$

where $\mathbf{B}_2 = I_p - (\lambda+d)^2\mathbf{A}_{\lambda}^{-2}$. The bias and covariance of $\hat{\boldsymbol{\gamma}}_{\mathrm{BJLTE}}$ are as follows:

$$\mathbb{B}\left(\hat{\boldsymbol{\gamma}}_{\mathrm{BJLTE}}\right) = -(d+\lambda)^2\mathbf{A}_{\lambda}^{-2}\boldsymbol{\gamma}, \tag{24}$$

$$\mathrm{Cov}\left(\hat{\boldsymbol{\gamma}}_{\mathrm{BJLTE}}\right) = \frac{1}{\tau}\mathbf{B}_2\mathbf{Q}^{-1}\mathbf{B}_2^T. \tag{25}$$

The MMSE and SMSE are given, respectively, by

$$\mathrm{MMSE}\left(\hat{\boldsymbol{\gamma}}_{\mathrm{BJLTE}}\right) = \frac{1}{\tau}\mathbf{B}_2\mathbf{Q}^{-1}\mathbf{B}_2^T + (d+\lambda)^4\mathbf{A}_{\lambda}^{-2}\boldsymbol{\gamma}\boldsymbol{\gamma}^T\mathbf{A}_{\lambda}^{-2},$$
$$\mathrm{SMSE}\left(\hat{\boldsymbol{\gamma}}_{\mathrm{BJLTE}}\right) = \frac{1}{\tau}\sum_{i=1}^{p+1}\frac{\left(q_i - d\right)^2\left(q_i + d + 2\lambda\right)^2}{q_i\left(q_i+\lambda\right)^4} + (d+\lambda)^4\sum_{i=1}^{p+1}\frac{\gamma_i^2}{\left(q_i+\lambda\right)^4}. \tag{26}$$

4.3. Properties of BMJLTE. The canonical form of BMJLTE of $\boldsymbol{\gamma}$ is given by

$$\hat{\boldsymbol{\gamma}}_{\mathrm{BMJLTE}} = \mathbf{B}_2\mathbf{B}_1\hat{\boldsymbol{\gamma}}_{\mathrm{BMLE}}; \quad \lambda > 0, d \in \mathbb{R}, \tag{27}$$

where $\mathbf{B}_1 = I_p - (\lambda+d)\mathbf{A}_{\lambda}^{-1}$. The bias and covariance of $\hat{\boldsymbol{\gamma}}_{\mathrm{BMJLTE}}$ are given by

$$\mathbb{B}\left(\hat{\boldsymbol{\gamma}}_{\mathrm{BMJLTE}}\right) = -(d+\lambda)\mathbf{H}\mathbf{A}_{\lambda}^{-1}\boldsymbol{\gamma}, \tag{28}$$

$$\mathrm{Cov}\left(\hat{\boldsymbol{\gamma}}_{\mathrm{BMJLTE}}\right) = \frac{1}{\tau}\mathbf{B}_2\mathbf{B}_1\mathbf{Q}^{-1}\mathbf{B}_1^T\mathbf{B}_2^T, \tag{29}$$

where $\mathbf{H} = I_p + (d+\lambda)\mathbf{A}_{\lambda}^{-1} - (d+\lambda)^2\mathbf{A}_{\lambda}^{-2}$. The MMSE and SMSE are given, respectively, by

$$\mathrm{MMSE}\left(\hat{\boldsymbol{\gamma}}_{\mathrm{BMJLTE}}\right) = \frac{1}{\tau}\mathbf{B}_2\mathbf{B}_1\mathbf{Q}^{-1}\mathbf{B}_1^T\mathbf{B}_2^T + (d+\lambda)^2\mathbf{H}\mathbf{A}_{\lambda}^{-1}\boldsymbol{\gamma}\boldsymbol{\gamma}^T\mathbf{A}_{\lambda}^{-1}\mathbf{H}^T,$$
$$\mathrm{SSMSE}\left(\hat{\boldsymbol{\gamma}}_{\mathrm{BMJLTE}}\right) = \frac{1}{\tau}\sum_{i=1}^{p+1}\frac{\left(q_i - d\right)^4\left(q_i + d + 2\lambda\right)^2}{q_i\left(q_i+\lambda\right)^6}$$
$$+ (d+\lambda)^2\sum_{i=1}^{p+1}\frac{\gamma_i^2\left(\left(q_i+\lambda\right)^2 + (\lambda+d)\left(q_i-d\right)\right)^2}{\left(q_i+\lambda\right)^6}. \tag{30}$$

5. Theoretical Comparison of Estimators

In this section, we conduct a theoretical comparison between the proposed estimators and the BLTE, focusing on the squared bias (SB) and MMSE. We begin by evaluating the SB, which is defined as follows:

$$\mathrm{SB}(\widehat{\beta}) = [\mathbb{B}(\widehat{\beta})]^T[\mathbb{B}(\widehat{\beta})]. \tag{31}$$

For comparison among the MMSE estimators, we need the following lemma.

Lemma 1. *Let $\widehat{\beta}_i$ for $i = 1,2$ be two estimators of $\boldsymbol{\beta}$ with the covariance matrix $Cov(\widehat{\beta}_i)$ and the bias vector b_i, then*

$$\mathrm{MMSE}\left(\widehat{\beta}_1\right) - \mathrm{MMSE}\left(\widehat{\beta}_2\right) = \mathbf{D} + b_1b_1^T - b_2b_2^T > 0, \tag{32}$$

if and only if

$$b_2^T\left[\mathbf{D} + b_1b_1^T\right]b_2 < 1, \tag{33}$$

where $\mathbf{D} = \mathbf{Cov}(\widehat{\beta}_1) - \mathbf{Cov}(\widehat{\beta}_2)$ is a positive defined matrix [33].

5.1. Comparison between BLTE and BJLTE

Theorem 2. *The SB of BJLTE is smaller than the SB of BLTE if*

$$\min_{1\le i\le p+1}\{(q_i-d)(q_i+d+2\lambda)\}>0. \tag{34}$$

Proof. By using (20) and (24), we have

$$\begin{aligned}
\mathrm{SB}\left(\hat{\boldsymbol{\gamma}}_{\mathrm{BLTE}}\right)-\mathrm{SB}\left(\hat{\boldsymbol{\gamma}}_{\mathrm{BJLTE}}\right) &= (d+\lambda)^2\boldsymbol{\gamma}^T\mathbf{A}_\lambda^{-2}\boldsymbol{\gamma}-(d+\lambda)^4\boldsymbol{\gamma}^T\mathbf{A}_\lambda^{-2}\boldsymbol{\gamma}\boldsymbol{\gamma}\\
&= (d+\lambda)^2\sum_{i=1}^{p+1}\frac{\gamma_i^2}{(q_i+\lambda)^2}-(d+\lambda)^4\sum_{i=1}^{p+1}\frac{\gamma_i^2}{(q_i+\lambda)^4}\\
&= (d+\lambda)^2\sum_{i=1}^{p+1}\frac{\gamma_i^2}{(q_i+\lambda)^4}\left((q_i+\lambda)^2-(d+\lambda)^2\right)\\
&= (d+\lambda)^2\sum_{i=1}^{p+1}\frac{\gamma_i^2}{(q_i+\lambda)^4}(q_i-d)(q_i+d+2\lambda).
\end{aligned} \tag{35}$$

This equation is positive if $(q_i-d)(q_i+d+2\lambda)>0$ for $i=1,2,\ldots,p+1$ then the proof is completed. □

Theorem 3. *When* $\max_{1\le i\le p+1}\{(\lambda+d)(2q_i+d+3\lambda)\}<0$, *the BMJLTE is superior to the BLTE in terms of MMSE if the following inequality holds.*

$$\mathbb{B}\left(\hat{\boldsymbol{\gamma}}_{\mathrm{BJLTE}}\right)^T\left[\mathbf{D}_1+\left[\mathbb{B}\left(\hat{\boldsymbol{\gamma}}_{\mathrm{BLTE}}\right)\right]^T\left[\mathbb{B}\left(\hat{\boldsymbol{\gamma}}_{\mathrm{BLTE}}\right)\right]\right]^{-1}\mathbb{B}\left(\hat{\boldsymbol{\gamma}}_{\mathrm{BJLTE}}\right)<1, \tag{36}$$

where $\mathbf{D}_1=Cov(\hat{\boldsymbol{\gamma}}_{BLTE})-Cov(\hat{\boldsymbol{\gamma}}_{BJLTE})$.

Proof. By following Lemma 1, it is required to only show that $\mathbf{D}_1$ is a defined positive matrix.

$$\begin{aligned}
\mathbf{D}_1 &= \mathrm{Cov}\left(\hat{\boldsymbol{\gamma}}_{\mathrm{BLTE}}\right)-\mathrm{Cov}\left(\hat{\boldsymbol{\gamma}}_{\mathrm{BJLTE}}\right)\\
&= \frac{1}{\tau}\left[\mathbf{A}_\lambda^{-1}\left(\mathbf{Q}-dI_p\right)\mathbf{Q}^{-1}\left(\mathbf{Q}-dI_p\right)^T\mathbf{A}_\lambda^{-1}-\mathbf{B}_2\mathbf{Q}^{-1}\mathbf{B}_2^T\right]\\
&= \frac{1}{\tau}\mathrm{diag}\left(\frac{(q_i-d)^2}{q_i(q_i+\lambda)^2}-\frac{(q_i-d)^2(q_i+d+2\lambda)^2}{q_i(q_i+\lambda)^4}\right)\\
&= \frac{1}{\tau}\mathrm{diag}\left(\frac{(q_i-d)^2}{q_i(q_i+\lambda)^4}\{-(d+\lambda)(2q_i+d+3\lambda)\}\right).
\end{aligned} \tag{37}$$

Thus, $\mathbf{D}_1$ will be positive if $(\lambda+d)(2q_i+d+3\lambda)<0$, for $i=1,2,\ldots,p+1$. Hence, the proof is completed. □

5.2. Comparison between BLTE and BMJLTE

Theorem 4. *The SB of BMJLTE is smaller than the SB of BLTE if for $i = 1,2,\ldots,p+1$, we have $d < -q_i - 2\lambda$ or $-\lambda < d < q_i$ or $d > 2q_i + \lambda$.*

Proof. By using (20) and (28), we have

$$\begin{aligned}\mathrm{SB}\left(\hat{\boldsymbol{\gamma}}_{\mathrm{BLTE}}\right) - \mathrm{SB}\left(\hat{\boldsymbol{\gamma}}_{\mathrm{BMJLTE}}\right) &= (d+\lambda)^2 \boldsymbol{\gamma}^T \mathbf{A}_\lambda^{-2} \boldsymbol{\gamma} - (d+\lambda)^2 \boldsymbol{\gamma}^T \mathbf{A}_\lambda^{-1} \mathbf{D}^T \mathbf{D} \mathbf{A}_\lambda^{-1} \boldsymbol{\gamma} \\ &= (d+\lambda)^2 \sum_{i=1}^{p+1} \frac{\gamma_i^2}{(q_i+\lambda)^2} - (d+\lambda)^2 \sum_{i=1}^{p+1} \frac{\gamma_i^2}{(q_i+\lambda)^6}\left((q_i+\lambda)^2 + (\lambda+d)(q_i-d)\right)^2 \\ &= (d+\lambda)^2 \sum_{i=1}^{p+1} \frac{\gamma_i^2}{(q_i+\lambda)^6}\left((q_i+\lambda)^4 - \left((q_i+\lambda)^2 + (\lambda+d)(q_i-d)\right)^2\right) \\ &= (d+\lambda)^2 \sum_{i=1}^{p+1} \frac{\gamma_i^2}{(q_i+\lambda)^6}\left(2(q_i+\lambda)^2(\lambda+d)(d-q_i) - (\lambda+d)^2(q_i+d)^2\right).\end{aligned} \tag{38}$$

This equation is positive if for $i = 1,2,\ldots,p+1$, $h_i(d) = 2(q_i+\lambda)^2(\lambda+d)(d-q_i) - (\lambda+d)^2(q_i+d)^2$ be positive. This function has four roots $d_1 = q_i$, $d_2 = -\lambda$, $d_3 = 2q_i + \lambda$, and $d_4 = -2\lambda - q_i$. Since $-2\lambda - q_i < -\lambda < q_i < 2q_i + \lambda$, the function $h_i(d)$ is positive if $d < -q_i - 2\lambda$ or $-\lambda < d < q_i$ or $d > 2q_i + \lambda$ for $i = 1,2,\ldots,p+1$. □

Theorem 5. *When $d \in (\lambda - \sqrt{2}(q_i+\lambda), \lambda + \sqrt{2}(q_i+\lambda))$ for $i = 1,2,\ldots,p+1$, the BMJLTE superior the BLTE in terms of MMSE if the following inequality holds*

$$\mathbb{B}\left(\hat{\boldsymbol{\gamma}}_{\mathrm{BMJLTE}}\right)^T\left[\mathbf{D}_2 + \left[\mathbb{B}\left(\hat{\boldsymbol{\gamma}}_{\mathrm{BLTE}}\right)\right]^T\left[\mathbb{B}\left(\hat{\boldsymbol{\gamma}}_{\mathrm{BLTE}}\right)\right]\right]^{-1}\mathbb{B}\left(\hat{\boldsymbol{\gamma}}_{\mathrm{BMJLTE}}\right) < 1, \tag{39}$$

where $\mathbf{D}_2 = Cov(\hat{\boldsymbol{\gamma}}_{BLTE}) - Cov(\hat{\boldsymbol{\gamma}}_{BMJLTE})$.

Proof. By following Lemma 1, it is enough to show that $\mathbf{D}_2$ is a defined positive matrix.

$$\begin{aligned}\mathbf{D}_2 &= \mathrm{Cov}\left(\hat{\boldsymbol{\gamma}}_{\mathrm{BLTE}}\right) - \mathrm{Cov}\left(\hat{\boldsymbol{\gamma}}_{\mathrm{BMJLTE}}\right) \\ &= \frac{1}{\tau}\left[\mathbf{A}_\lambda^{-1}\left(\mathbf{Q} - dI_p\right)\mathbf{Q}^{-1}\left(\mathbf{Q} - dI_p\right)^T \mathbf{A}_\lambda^{-1} - \mathbf{B}_2\mathbf{B}_1\mathbf{A}_\lambda^{-1}\mathbf{B}_1^T\mathbf{B}_2^T\right] \\ &= \frac{1}{\tau}\mathrm{diag}\left(\frac{(q_i-d)^2}{q_i(q_i+\lambda)^2} - \frac{(q_i-d)^4(q_i+d+2\lambda)^2}{q_i(q_i+\lambda)^6}\right) \\ &= \frac{1}{\tau}\mathrm{diag}\left(\frac{(q_i-d)^2}{q_i(q_i+\lambda)^6}\left\{(q_i+\lambda)^4 - (q_i-d)^2(q_i+d+2\lambda)^2\right\}\right) \\ &= \frac{1}{\tau}\mathrm{diag}\left(\frac{(q_i-d)^2}{q_i(q_i+\lambda)^6}\left\{-d^2 - 2d\lambda + \lambda^2 + 2q_i^2 + 4\lambda q_i\right\}\right).\end{aligned} \tag{40}$$

Thus, $\mathbf{D}_2$ will be positive if for $i = 1,2,\ldots,p+1$, $f_i(d) = -d^2 - 2d\lambda + \lambda^2 + 2q_i^2 + 4\lambda q_i$ be positive. The discrimination of $f_i(d)$ is $\Delta = 8(q_i+\lambda)^2 > 0$; therefore, we will have two following real roots:

$$d = \frac{2\lambda \pm 2\sqrt{2}(q_i+\lambda)}{-2} = \lambda \pm \sqrt{2}(q_i+\lambda). \tag{41}$$

Thus, $f_i(d)$ is positive if $\lambda - \sqrt{2}(q_i+\lambda) < d < \lambda + \sqrt{2}(q_i+\lambda)$ and the proof is finished. □

5.3. Bias Parameter Selection. In the following subsection, we will derive an estimator for the parameter d. In the search for this estimator, we consider the BJLTE. Therefore, the first derivation of the SMSE of BJLTE with respect to d is calculated as follows:

$$\frac{\partial}{\partial d}\mathrm{SMSE}\left(\hat{\boldsymbol{\gamma}}_{\mathrm{BJLTE}}\right) = \frac{1}{\tau}\sum_{i=1}^{p+1}\frac{-2\left(q_i - d\right)\left(q_i + d + 2\lambda\right)^2 + \left(q_i - d\right)^2\left(q_i + d + 2\lambda\right)}{q_i\left(q_i + \lambda\right)^2} + 4\left(d+\lambda\right)^3\sum_{i=1}^{p+1}\frac{\gamma_i^2}{\left(q_i + \lambda\right)^4} = \frac{1}{\tau}\sum_{i=1}^{p+1}\frac{4\left(d+\lambda\right)}{q_i\left(q_i + \lambda\right)^4}\left(\tau\left(d+\lambda\right)^2 q_i\gamma_i^2 - \left(q_i - d\right)\left(q_i + d + 2\lambda\right)\right). \tag{42}$$

This equation equals zero if $g_i(d) = \tau(d+\lambda)^2 q_i\gamma_i^2 - (q_i - d)(q_i + d + 2\lambda) = 0$ which is a quadratic function of d. This function leads to the following real roots:

$$d = -\lambda \pm \frac{\left(\lambda + q_i\right)}{\sqrt{1 + \tau q_i\gamma_i^2}}. \tag{43}$$

Although there are various estimators for d, we only recommend and utilize the following estimator in the simulation study.

$$d_{\mathrm{opt}} = \min_{1\le i\le p+1}\left(-\lambda + \frac{\left(\lambda + q_i\right)}{\sqrt{1 + \tau q_i\gamma_i^2}}\right). \tag{44}$$

6. Simulation Study

In this section, we will evaluate the performance of the proposed estimators in the BRM through a simulation study. By considering various values for parameter $(n = 50,100,200)$, the number of independent variables $(p = 3,6,12)$, and the precision parameter $(\tau = 2,4)$, we will present multiple potential outcomes that demonstrate the efficacy of the proposed estimator.

The n observations of the covariates are generated by

$$x_{ij} = \sqrt{1-\theta^2}u_{ij} + \theta u_{ip+1}, \quad i = 1,2,\cdots,n,\ j = 1,2,\cdots,p, \tag{45}$$

where u_{ij} are generated from standard normal distribution and θ is supplied to control the intensity of correlation among the covariates. The value of θ is considered to be 0.90, 0.095, and 0.99 to determine how the estimators are affected by varying degrees of collinearity. The values of coefficients are considered as $\beta_j = 1/\sqrt{p+1}$ such that $\boldsymbol{\beta}^T\boldsymbol{\beta} = 1$. Finally, the n observations of the response variable in the BRM with logit link function are generated from the beta distribution, $\mathrm{beta}(\mu_i\tau, (1-\mu_i)\tau)$, where

$$\mu_i = \log\left(\frac{e^{\eta_i}}{1+e^{\eta_i}}\right);\quad \eta_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip},\ i = 1,2,\cdots,n. \tag{46}$$

In order to determine the BLTE, we utilize the estimators described in [16] for λ and d as follows:

$$\hat{\lambda} = \frac{q_{\min}}{\hat{\tau}\gamma_{\min}^2}, \quad \hat{d} = \frac{\sum_{j=1}^{p+1}\left(1 - \hat{\tau}\hat{\lambda}\gamma_j^2/\left(q_j + \hat{\lambda}\right)^2\right)}{\sum_{j=1}^{p+1}\left(1 + \hat{\tau}\hat{\lambda}\gamma_j^2/q_j\left(q_j + \hat{\lambda}\right)^2\right)}, \tag{47}$$

where

$$\hat{\tau} = \frac{1}{n-p-1}\sum_{i=1}^{n}\left(\frac{y_i - \hat{\mu}_i}{\hat{\mu}_i}\right)^2. \tag{48}$$

To evaluate the performance of the BJLTE and BMJLTE, we utilize the same λ for both estimators. For the parameter $\hat{d}$, we employ the estimator provided in equation (44). We also conclude the BRE and BLE estimators in our simulation study as well. For the ridge parameter, we use estimator $\hat{\lambda}$ in (47) and for the Liu estimator, we use the following estimator proposed by Karlsson et al. [32]:

$$\hat{d} = \frac{\sum_{j=1}^{p+1}\left(1 - \hat{\tau}\gamma_j^2/\left(q_j + 1\right)^2\right)}{\sum_{j=1}^{p+1}\left(1 + \hat{\tau}\gamma_j^2/q_j\left(q_j + 1\right)^2\right)}. \tag{49}$$

In comparing the performance of the estimators, we specifically focus on the simulated MSE in addition to the squared bias. Therefore, we repeat the experiment 1000 times and calculate the criteria using the following formulas:

Table 1: MSEs of estimators for different values of p, n, τ, and θ.

n	θ	$\tau = 2$						$\tau = 4$					
		BMLE	BRE	BLE	BLTE	BJLTE	BMJLTE	BMLE	BRE	BLE	BLTE	BJLTE	BMJLTE
							$p = 3$						
	0.90	0.3022	0.3384	0.2923	0.2832	0.2147	0.1103	0.2478	0.2142	0.2319	0.1912	0.1543	0.0798
50	0.95	0.5794	0.6134	0.5468	0.5350	0.2909	0.1145	0.4436	0.3636	0.3873	0.2865	0.2282	0.0910
	0.99	2.2172	2.8775	1.7931	1.7655	1.0641	0.2832	2.0901	1.6876	1.7618	1.2805	0.7917	0.2122
	0.90	0.1333	0.1465	0.1129	0.0969	0.0917	0.0664	0.1127	0.0983	0.1096	0.0958	0.0865	0.0517
100	0.95	0.2543	0.2804	0.2127	0.2106	0.1654	0.0745	0.2122	0.1865	0.1975	0.1806	0.1302	0.0587
	0.99	1.1973	1.2859	1.0951	0.9460	0.4728	0.1253	0.9914	0.8548	0.9764	0.7054	0.4091	0.1144
	0.90	0.0893	0.0942	0.0827	0.0782	0.0646	0.0448	0.0525	0.0521	0.0524	0.0520	0.0464	0.0319
200	0.95	0.1306	0.1405	0.1224	0.1053	0.0994	0.0514	0.0965	0.0888	0.0937	0.0831	0.0712	0.0379
	0.99	0.6009	0.6327	0.5507	0.5165	0.2801	0.0838	0.4587	0.3885	0.4271	0.3689	0.2175	0.0672
							$p = 6$						
	0.90	0.8585	0.7101	0.8112	0.6779	0.5602	0.2451	0.6874	0.6137	0.6288	0.5241	0.4639	0.2073
50	0.95	1.7104	1.1514	1.3928	0.9291	0.9904	0.3846	1.2851	1.0236	1.1967	0.9009	0.7720	0.3015
	0.99	8.6889	7.2763	7.4530	6.7188	4.5850	1.6188	5.9755	3.8911	4.5802	3.8891	3.1749	1.1342
	0.90	0.3986	0.3477	0.3920	0.3255	0.3059	0.1569	0.2873	0.2803	0.2815	0.2742	0.2296	0.1201
100	0.95	0.7306	0.6585	0.6968	0.5220	0.4874	0.2020	0.5359	0.4857	0.5184	0.4563	0.3746	0.1594
	0.99	3.5323	2.6366	2.9676	2.0615	1.8698	0.5925	2.6909	2.0213	2.2174	1.8216	1.4400	0.4598
	0.90	0.2890	0.2430	0.2748	0.2355	0.1629	0.1024	0.1605	0.1535	0.1598	0.1351	0.1239	0.0772
200	0.95	0.4578	0.4176	0.4204	0.4034	0.2751	0.1376	0.2846	0.2661	0.2730	0.2548	0.2038	0.1032
	0.99	1.5683	1.1589	1.3526	1.1169	0.8971	0.2909	1.2089	1.0617	1.1856	0.8241	0.7236	0.2431
							$p = 12$						
	0.90	2.5243	1.6977	1.9620	1.1786	1.4256	0.9571	1.9009	1.2977	1.6209	0.8938	0.8331	0.4962
50	0.95	5.0245	4.2807	4.8692	4.0890	3.5342	3.2079	3.6854	2.7807	3.3092	2.4606	2.3068	1.3676
	0.99	24.6431	12.8085	19.7217	9.6211	9.1361	8.7632	18.4751	10.9167	16.2137	6.7541	4.1255	2.2051
	0.90	1.0469	0.8181	0.9835	0.6289	0.4407	0.3488	0.7490	0.6345	0.7184	0.4629	0.7442	0.6787
100	0.95	1.9974	1.3483	1.6154	0.9259	0.9349	0.6284	1.4818	1.0734	1.2275	0.6999	1.4447	1.2420
	0.99	9.9571	5.2884	8.4150	3.7668	3.1380	2.8220	7.2307	4.7531	5.1546	3.5274	2.9461	1.6707
	0.90	0.4948	0.4095	0.4673	0.3818	0.3940	0.2719	0.3489	0.2745	0.3076	0.2799	0.3484	0.2333
200	0.95	0.9220	0.8566	0.9074	0.8478	0.6173	0.4505	0.6735	0.4526	0.5647	0.4105	0.3700	0.3224
	0.99	4.3973	3.9470	4.1994	3.6736	2.8949	1.7720	3.2880	2.5741	2.8634	1.1837	1.9341	0.9379

$$\text{MSE}(\widehat{\beta}) = \frac{1}{1000}\sum_{m=1}^{1000}\left(\widehat{\beta}_m - \boldsymbol{\beta}\right)^T\left(\widehat{\beta}_m - \boldsymbol{\beta}\right),$$
$$\text{SB}(\widehat{\beta}) = (\overline{\overline{\beta}} - \boldsymbol{\beta})^T(\overline{\overline{\beta}} - \boldsymbol{\beta}), \tag{50}$$

where $\widehat{\beta}_m$ is the estimation of $\boldsymbol{\beta}$ at the m th repetition of simulation and $\overline{\overline{\beta}}$ is the mean of estimated values.

The MSE and SB of estimators are presented in Tables 1 and 2, respectively. Based on the information provided in the tables, the following conclusions can be drawn:

(i) Based on Table 1, BMJLTE always outperforms other estimators in terms of MSE.

(ii) In terms of MSE, BLE is worse than other estimators.

(iii) The MSE of BRE is only less than BLE.

(iv) The BJLTE exhibits better performance than the BLTE, BLE, and BRE based on the MSE values.

(v) When the intensity of correlation increases, the MSE for all estimators increases.

(vi) When the dispersion value changes from $\tau = 2$ to $\tau = 4$, the MSE of all estimators reduces.

(vii) The MSE of estimators tends to increase as the number of covariates increases.

(viii) For a fixed value of p, θ, and τ, increasing the sample size results in a decrease in the MSE for all estimators.

(ix) Table 2 shows that when $p = 3$ and 6, the BJLTE has the lowest SB among the estimators.

(x) However, when $p = 12$ and there is a high correlation ($\theta = 0.99$), the BMJLTE has the lowest SB and for $\theta = 0.90$ and 0.95, the BJLTE still performs better.

(xi) The BLE has the largest squared bias among the estimators in all scenarios.

Table 2: SBs of estimators for different values of p, n, τ, and θ.

n	θ	$\tau = 2$					$\tau = 4$				
		BRE	BLE	BLTE	BJLTE	BMJLTE	BRE	BLE	BLTE	BJLTE	BMJLTE
						$p = 3$					
	0.90	0.0068	0.0097	0.0070	0.0027	0.0063	0.0029	0.0077	0.0032	0.0003	0.0019
50	0.95	0.0051	0.0074	0.0052	0.0019	0.0044	0.0035	0.0019	0.0037	0.0025	0.0034
	0.99	0.0046	0.0145	0.0108	0.0011	0.0034	0.0067	0.0030	0.0082	0.0016	0.0026
	0.90	0.0083	0.0210	0.0056	0.0015	0.0036	0.0016	0.0097	0.0023	0.0003	0.0019
100	0.95	0.0059	0.0134	0.0063	0.0031	0.0058	0.0017	0.0079	0.0014	0.0004	0.0012
	0.99	0.0046	0.0270	0.0064	0.0035	0.0044	0.0019	0.0079	0.0046	0.0004	0.0022
	0.90	0.0038	0.0282	0.0036	0.0013	0.0026	0.0004	0.0096	0.0009	0.0001	0.0005
200	0.95	0.0035	0.0184	0.0034	0.0014	0.0040	0.0011	0.0131	0.0011	0.0003	0.0009
	0.99	0.0044	0.0061	0.0061	0.0041	0.0045	0.0015	0.0068	0.0021	0.0012	0.0014
						$p = 6$					
	0.90	0.0141	0.0521	0.0148	0.0088	0.0102	0.0023	0.0085	0.0035	0.0012	0.0013
50	0.95	0.0150	0.3049	0.0168	0.0101	0.0128	0.0024	0.0087	0.0031	0.0017	0.0026
	0.99	0.0179	2.8894	0.0176	0.0156	0.0162	0.0039	0.0916	0.0065	0.0035	0.0045
	0.90	0.0161	0.0134	0.0166	0.0109	0.0130	0.0040	0.0081	0.0043	0.0023	0.0034
100	0.95	0.0169	0.0316	0.0192	0.0151	0.0186	0.0033	0.0112	0.0045	0.0028	0.0041
	0.99	0.0228	1.2182	0.0203	0.0172	0.0186	0.0071	0.0142	0.0040	0.0032	0.0035
	0.90	0.0188	0.0231	0.0164	0.0119	0.0134	0.0040	0.0094	0.0037	0.0022	0.0032
200	0.95	0.0191	0.0251	0.0190	0.0153	0.0120	0.0040	0.0135	0.0040	0.0028	0.0039
	0.99	0.0189	0.1154	0.0206	0.0196	0.0201	0.0053	0.0097	0.0048	0.0044	0.0046
						$p = 12$					
	0.90	0.0831	0.2812	0.0829	0.0772	0.0803	0.0681	0.2812	0.0668	0.0648	0.0665
50	0.95	0.0978	0.7763	0.0955	0.0938	0.0948	0.0878	0.7763	0.0865	0.0850	0.0858
	0.99	0.1104	7.6194	0.1124	0.1103	0.1023	0.1104	7.6194	0.1104	0.1016	0.1007
	0.90	0.0712	0.2431	0.1619	0.0581	0.0645	0.0803	0.0295	0.0741	0.0703	0.0726
100	0.95	0.0873	0.7990	0.1542	0.0766	0.0786	0.0864	0.0608	0.0914	0.0898	0.0903
	0.99	0.1090	6.8720	0.1029	0.1007	0.0890	0.1109	6.8720	0.1122	0.1107	0.1005
	0.90	0.0767	0.2249	0.1213	0.0515	0.0555	0.0767	0.2249	0.0740	0.0698	0.0717
200	0.95	0.0907	0.6941	0.1492	0.0666	0.0693	0.0917	0.0941	0.0918	0.0889	0.0910
	0.99	0.1109	0.1663	0.1007	0.0803	0.0751	0.1109	0.1663	0.1056	0.1025	0.1001

(xii) In general, as the sample size increases, the SB values decrease. However, in most scenarios, the SB values increase whenever the number of covariates or the intensity of correlation increases.

(xiii) When the dispersion value changes from $\tau = 2$ to 4, the SB values decrease.

7. Application to Real Data

In this section, two chemical datasets are utilized to demonstrate the performance of proposed estimators.

7.1. Gasoline Yield Data. The first dataset used in this study is sourced from Prater [2]. The objective is to investigate the impact of several covariates on a response variable, which represents the percentage of crude oil converted to gasoline through the process of distillation and fractionation. Initially, this dataset was analyzed by [34] using a linear regression model. However, it was discovered that the error term distribution was not symmetrical. As a result, the data were transformed to ensure that the dependent variable took values along the real number line. Lemonte et al. [35] adopted an alternative approach to address this issue. They used the beta distribution for analysis and found that it provided more robust outcomes for influential observations compared to the method employed by Atkinson [34]. The covariates in this dataset are: the crude oil gravity (x_1), the vapor pressure of the crude oil (x_2), the temperature at which 10 percent of crude oil has vaporized (x_3), and finally the temperature at which all the gasoline is vaporized (x_4). Qasim et al. [12] also used this dataset to illustrate the application of ridge beta regression.

The left-side plot in Figure 1 reveals a significant correlation among the variables, particularly between X_2 and X_3. Furthermore, we compute the condition index ($CI = \sqrt{q_{\max}/q_{\min}}$) of matrix $\mathbf{X}^T\widehat{\mathbf{V}}\mathbf{X}$ as 10613.01. Both the correlation plot and CI indicate the presence of multicollinearity in the dataset. Therefore, we apply the proposed estimators described in this paper. To evaluate the performance of the proposed estimators, we employ the bootstrapping method with sample size $n = 15$ and 1000 bootstrap iterations. We consider the BMLE as the true value of the coefficient vector and compute the mean squared error (MSE) and the mean absolute error (MAE) of various estimators by

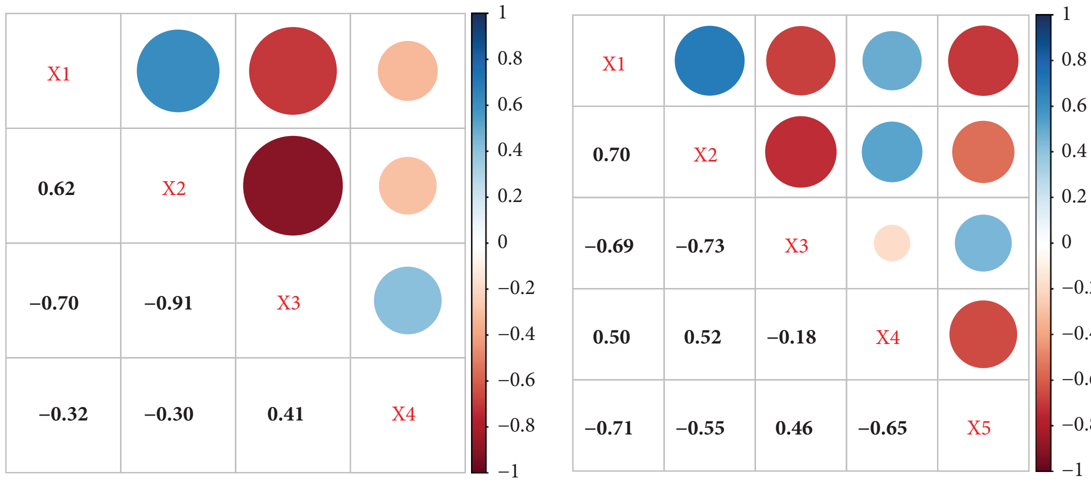

Figure 1: Visualization of correlation among the covariates for datasets.

Table 3: MSE and MAE of estimators for gasoline yield data.

	BRE		BLE		BLTE		BJLTE		BMJLTE	
	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
Intercept	10.077	2.8984	119.65	30.339	8.0875	2.6083	0.9055	0.2408	1.8979	0.8661
x_1	0.0006	0.0223	0.0578	0.2152	0.0005	0.0200	1.5 $E-05$	0.0015	6.0 $E-05$	0.0063
x_2	0.0058	0.0585	1.0483	0.7870	0.0047	0.0526	0.0011	0.0066	0.0019	0.0197
x_3	6.4 $E-05$	0.0070	0.0083	0.0793	5.2 $E-5$	0.0063	8.0 $E-6$	0.0006	1.5 $E-5$	0.0021
x_4	2.0 $E-06$	0.0006	4.40 $E-05$	0.0038	2.0 $E-06$	0.0006	1.7 $e-09$	2.4 $E-05$	4.7 $e-08$	0.0001

$$\mathrm{MSE}\left(\hat{\beta}_j\right) = \frac{1}{1000}\sum_{m=1}^{1000}\left(\hat{\beta}_j^{(m)} - \hat{\beta}_{j(\mathrm{BMLE})}^{(m)}\right)^2, \quad (51)$$

$$\mathrm{MAE}\left(\hat{\beta}_j\right) = \frac{1}{1000}\sum_{m=1}^{1000}\left|\hat{\beta}_j^{(m)} - \hat{\beta}_{j(\mathrm{BMLE})}^{(m)}\right|, \quad (52)$$

where $\hat{\beta}_{j(\mathrm{BMLE})}^{(m)}$ is the maximum likelihood estimation and the $\hat{\beta}_j$ is one of the considered estimators of β_j, e.g., BRE, BLE, BLTE, BJLTE, and BMJLTE, in the m th bootstrap replication. The results in Table 3 show that the Jackknifed estimators have significantly lower MSE and MAE values compared to BLTE, BLE, and BRE. Specifically, the BMJLTE outperforms the other estimators in terms of MSE, and BJLTE outperforms the other estimators in terms of MAE, which is consistent with the simulation study results. Furthermore, Table 4 reports the estimation of coefficients using the mean of the bootstrap estimators.

7.2. Heat Treating Test Data. The second dataset considered in this study is the heat-treating test data obtained from [36]. It comprises five covariates: furnace temperature (X_1), carbon concentration and duration of the carburizing cycle (soakpct and soaktime) denoted as X_2 and X_3, and carbon concentration and duration of the defuse time (Difftime and Diffpct) indicated as X_4 and X_5. The response variable captures the quality of a sound determined by the rate of vibrations or the level of something which is referred to as PITCH in the dataset and denotes the product presentation to the customer's heart. Since the response variable follows a ratio form, we employ the beta regression model to analyze it. However, before applying the beta regression model, we verify whether the values of y follow a beta distribution. We conduct an Anderson–Darling (AD) test using the ad test function from the goftest package in the R programming language. The computed test statistic is 0.85967, with a p value of 0.439. The estimated parameter values are $a = 4.9995$ and $b = 182.1799$. The p value suggests that the beta distribution is suitable for modeling the response variable.

We fit a model by including an intercept term and compute the condition index of the matrix $\mathbf{X}^T\hat{\mathbf{V}}\mathbf{X}$, which results in a value of 303772.7. The correlation matrix of the covariates is displayed in the right-side plot of Figure 1. Both observations indicate the presence of multicollinearity in the dataset. Consequently, we apply the proposed estimators to this dataset as well. To evaluate the efficiency of the proposed estimators, similar to the first dataset, we employ the bootstrapping method with a sample size of $n = 15$ and 1000 bootstrap iterations.

We calculate the MSE and MAE of each estimator by using (51) and (52), respectively. Based on Table 5, the Jackknifed estimators are superior to the BLTE, BLE, and BRE due to having the lowest value of MSE and MAE. The MSE of the BMJLTE is lower than that of BJLTE but for the MAE, it is the opposite. The estimation of the coefficients by using the proposed estimators is presented in Table 6 which is obtained by using the mean of the estimation for all bootstrap estimations.

Table 4: Estimation of coefficients for gasoline yield data.

	MBLE	BRE	BLE	BLTE	BJLTE	BMJLTE
Intercept	−2.8362	−0.0019	−31.9338	−0.2892	−2.6504	−2.0311
x_1	0.0050	−0.0168	0.2060	−0.0145	0.0037	−0.001
x_2	0.0356	−0.0193	0.7818	−0.0136	0.031	0.0188
x_3	−0.0109	−0.0175	0.0649	−0.0168	−0.0113	−0.0128
x_4	0.0107	0.0103	0.0101	0.0103	0.0107	0.0107

Table 5: MSE and MAE of estimators for heat treating test data.

	BRE		BLE		BLTE		BJLTE		BMJLTE	
	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
Intercept	381.678	18.253	397.05	98.722	345.55	14.947	377.64	10.673	341.74	11.330
x_1	$2.4\ E-02$	0.0022	0.6929	0.0593	$1.4\ E-03$	0.0007	$1.2\ E-03$	0.0005	$1.1\ E-03$	0.0009
x_2	0.0518	0.0054	0.6205	0.1997	0.0091	0.0026	0.0025	0.0012	0.0023	0.0015
x_3	4.5216	1.5143	6.5441	1.6992	4.2757	1.1854	2.4942	0.0843	1.4043	0.9981
x_4	0.1018	0.0401	0.4679	0.3817	0.0383	0.0148	0.0101	0.0019	0.0093	0.0054
x_5	0.3748	0.1778	0.5453	0.4451	0.2816	0.1635	0.2335	0.0465	0.1965	0.0811

Table 6: Estimation of coefficients for heat treating test data.

	MBLE	BRE	BLE	BLTE	BJLTE	BMJLTE
Intercept	−5.0144	−2.0309	−62.4927	−0.8049	−1.0367	−0.4186
x_1	0.0000	−0.0014	0.0358	−0.0016	−0.0014	−0.0018
x_2	0.0888	0.0888	−0.0411	0.0842	0.0829	0.0836
x_3	0.6477	0.1659	−0.206	−0.4477	−0.4923	−0.554
x_4	0.3747	0.3814	$-3.0\ E-04$	0.3977	0.4022	0.4012
x_5	−0.0343	−0.1885	0.0069	−0.3696	−0.3664	−0.3893

8. Conclusion

In this paper, we have addressed the bias issue in the Liu-type estimator used in BRMs. By applying the Jackknife methodology, we were able to reduce the bias of the beta Liu-type estimator and introduce a modified estimator. We have analytically established the conditions under which both proposed estimators outperform the beta Liu-type estimator. To evaluate the performance of the proposed estimators and compare them to the BMLE, BRE, BLE, and BLTE, we conducted a comprehensive simulation study. The simulation experiment considered various aspects to observe the behavior of the proposed estimators. The results indicate that the proposed estimators, especially the modified estimator, outperformed the BRE, BLE, and BLTE in terms of MSE and squared bias (SB). Furthermore, we have demonstrated the efficiency of the proposed estimators through two real-life examples in the field of chemometrics. In both cases, the proposed Jackknifed estimators exhibited smaller MSE and MAE compared to the alternative estimators. Based on these findings, we recommend researchers utilize the proposed estimators, especially the modified Jackknifed Liu-type estimator whenever multicollinearity is present in BRMs. The proposed estimators offer improved performance in terms of bias reduction and estimation accuracy.

Data Availability

The data supporting this paper are from previously reported studies and datasets, which have been cited.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] M. Amin, M. Qasim, S. Afzal, and K. Naveed, "New ridge estimators in the inverse Gaussian regression: Monte Carlo simulation and application to chemical data," *Communications in Statistics-Simulation and Computation*, vol. 51, no. 10, pp. 6170–6187, 2020.

[2] N. H. Prater, "Estimate gasoline yields from crudes," *Petrol Refiner*, vol. 35, no. 5, pp. 236–238, 1956.

[3] Z. Y. Algamal, M. H. Lee, A. M. Al-Fakih, and M. Aziz, "High-dimensional QSAR prediction of anticancer potency of imidazo[4,5-b] pyridine derivatives using adjusted adaptive LASSO," *Journal of Chemometrics*, vol. 29, no. 10, pp. 547–556, 2015.

[4] Z. Y. Algamal, M. K. Qasim, M. H. Lee, and H. T. Mohammad Ali, "High-dimensional QSAR/QSPR classification modeling based on improving pigeon optimization algorithm," *Chemometrics and Intelligent Laboratory Systems*, vol. 206, Article ID 104170, 2020.

[5] S. Ferrari and F. Cribari-Neto, "Beta regression for modelling rates and proportions," *Journal of Applied Statistics*, vol. 31, no. 7, pp. 799–815, 2004.

[6] Z. Y. Algamal, M. R. Abonazel, and A. F. Lukman, "Modified Jackknife ridge estimator for Beta regression model with application to Chemical Data," *International Journal of Mathematics, Statistics, and Computer Science*, vol. 1, pp. 15–24, 2023.

[7] Z. Y. Algamal, "A particle swarm optimization method for variable selection in beta regression model," *Electronic Journal of Applied Statistical Analysis*, vol. 12, no. 2, pp. 508–519, 2019.

[8] S. Seifollahi and H. Bevrani, "James-stein type estimators in beta regression model: simulation and application," *Hacettepe Journal of Mathematics and Statistics*, vol. 52, no. 4, pp. 1046–1065, 2023.

[9] M. H. Quenouille, "Notes on bias in estimation," *Biometrics*, vol. 43, no. 3-4, pp. 353–360, 1956.

[10] M. R. Abonazel and I. M. Taha, "Beta ridge regression estimators: simulation and application," *Communications in Statistics-Simulation and Computation*, vol. 52, no. 9, pp. 4280–4292, 2021.

[11] 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 Mathematics*, vol. 7, no. 1, pp. 35–57, 2021.

[12] M. Qasim, K. Mansson, and B. M. Golam Kibria, "On some beta ridge regression estimators: method, simulation and application," *Journal of Statistical Computation and Simulation*, vol. 91, no. 9, pp. 1699–1712, 2021.

[13] R. Arabi Belaghi, Y. Asar, and R. Larsson, "Improved shrinkage estimators in the beta regression model with application in econometric and educational data," *Statistical Papers*, vol. 64, no. 6, pp. 1891–1912, 2022.

[14] 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," *Frontiers in Applied Mathematics and Statistics*, vol. 7, pp. 1–10, 2022.

[15] M. R. Abonazel, I. Dawoud, F. A. Awwad, and A. F. Lukman, "Dawoud-Kibria estimator for beta regression model: simulation and application," *Frontiers in Applied Mathematics and Statistics*, vol. 8, pp. 1–12, 2022.

[16] Z. Y. Algamal and M. R. Abonazel, "Developing a Liu-type estimator in beta regression model," *Concurrency and Computation: Practice and Experience*, vol. 34, no. 5, 2021.

[17] A. Erkoc, E. Ertan, Z. Y. Algamal, and K. U. Akay, "The beta Liu-type estimator: simulation and application," *Hacettepe Journal of Mathematics and Statistics*, vol. 52, no. 3, pp. 828–840, 2023.

[18] K. Ohtani, "On small sample properties of the almost unbiased generalized ridge estimator," *Communications in Statistics-Theory and Methods*, vol. 15, no. 5, pp. 1571–1578, 1986.

[19] M. Amin, M. Qasim, A. Yasin, and M. Amanullah, "Almost unbiased ridge estimator in the gamma regression model," *Communications in Statistics-Simulation and Computation*, vol. 51, no. 7, pp. 3830–3850, 2020.

[20] J. Wu and Y. Asar, "On almost unbiased ridge logistic estimator for the logistic regression model," *Hacettepe Journal of Mathematics and Statistics*, vol. 45, no. 69, pp. 1–98, 2015.

[21] J. Wu and Y. Asar, "Performance of the almost unbiased ridge-type principal component estimator in logistic regression model," *Communications in Statistics-Simulation and Computation*, vol. 47, no. 10, pp. 2925–2937, 2018.

[22] N. Varathan and P. Wijekoon, "Modified almost unbiased Liu estimator in logistic regression," *Communications in Statistics-Simulation and Computation*, vol. 50, no. 11, pp. 3530–3546, 2019.

[23] Y. Asar and M. Korkmaz, "Almost unbiased Liu-type estimators in gamma regression model," *Journal of Computational and Applied Mathematics*, vol. 403, Article ID 113819, 2022.

[24] J. W. Tukey, "Bias and confidence in not quite large samples," *The Annals of Mathematical Statistics*, vol. 29,614 pages, 1958.

[25] B. Singh, Y. P. Chaubey, and T. D. Dwivedi, "An almost unbiased ridge estimator," *Sankhya: The Indian Journal of Statistics, Series B*, vol. 48, no. 3, pp. 23–46, 1986.

[26] F. S. M. Batah, T. V. Ramanathan, and S. D. Gore, "The efficiency of modified Jackknife and ridge type regression estimators: a comparison," *Surveys in Mathematics and its Applications*, vol. 3, pp. 111–122, 2008.

[27] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society: Series B*, vol. 58, no. 1, pp. 267–288, 1996.

[28] Z. Y. Algamal, "Developing a ridge estimator for the gamma regression model," *Journal of Chemometrics*, vol. 32, no. 10, 2018.

[29] S. Mahmood, N. Seyala, and Z. Algamal, "Adjusted R^2-type measures for beta regression model," *Electronic Journal of Applied Statistical Analysis*, vol. 13, no. 02, pp. 350–357, 2020.

[30] N. Yildiz, "On the performance of the Jackknifed Liu-type estimator in linear regression model," *Communications in Statistics-Theory and Methods*, vol. 47, no. 9, pp. 2278–2290, 2018.

[31] Y. P. Chaubey, M. Khurana, and S. Chandra, "An investigation into properties of jackknifed and Bootstrapped Liu-type estimator," *Far East Journal of Mathematical Sciences*, vol. 106, no. 1, pp. 159–170, 2018.

[32] P. Karlsson, K. Mansson, and B. G. Kibria, "A Liu estimator for the beta regression model and its application to chemical data," *Journal of Chemometrics*, vol. 34, no. 10, 2020.

[33] G. Trenkler and H. Toutenburg, "Mean squared error matrix comparisons between biased estimators: an overview of recent results," *Statistical Papers*, vol. 31, no. 1, pp. 165–179, 1990.

[34] A. C. Atkinson, *Plots, Transformations and Regression: An Introduction to Graphical Methods of Diagnostic Regression Analysis*, Oxford University Press, Oxford, UK, 1985.

[35] A. J. Lemonte, S. L. Ferrari, and F. Cribari-Neto, "Improved likelihood inference in Birnbaum-saunders regressions," *Computational Statistics and Data Analysis*, vol. 54, no. 5, pp. 1307–1316, 2010.

[36] D. C. Montgomery and G. C. Runger, *Applied Statistics and Probability for Engineers*, John Wiley and Sons, Hoboken, NJ, USA, 2014.

[37] T. A. Bancroft, "On biases in estimation due to the use of preliminary tests of significance," *The Annals of Mathematical Statistics*, vol. 15, no. 2, pp. 190–204, 1944.

[38] M. Meloun and J. Militký, "Detection of single influential points in OLS regression model building," *Analytical Chemistry Journal*, vol. 439, no. 2, pp. 169–191, 2001.