AIMS Mathematics

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

AIMS Mathematics, 11(3): 7115–7142.
DOI: 10.3934/math.2026293
Received: 03 February 2026
Revised: 03 March 2026
Accepted: 10 March 2026
Published: 18 March 2026

Research article

A novel modified Liu estimator for the inverse Gaussian regression model to effectively handle multicollinear positive data

Ali T. Hammad[1], **Ehab M. Almetwally**[2], **Hisham Mohamed Almongy**[2,*], **Ahmed M. Gemeay**[1], **Yousef Alharbi**[3], **Ramlah H. Albayyat**[4], **Wafa Ali J. Almohri**[5] **and Manahil SidAhmed Mustafa**[6]

[1] Department of Mathematics, Faculty of Science, Tanta University, Tanta 31527, Egypt
[2] Department of Mathematics and Statistics, College of Science, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh 11432, Saudi Arabia
[3] Department of Mathematics, College of Science, Qassim University, Buraydah 51452, Saudi Arabia
[4] Department of Mathematics, College of Science, Northern Border University, Arar, Saudi Arabia
[5] Department of Mathematics, College of Science, Taibah University, Madinah, Saudi Arabia
[6] Department of Statistics, Faculty of Science, University of Tabuk, Tabuk, Saudi Arabia

* **Correspondence:** Email: hElmongy@imamu.edu.sa.

Abstract: The inverse Gaussian regression model (IGRM) is a commonly used method for modeling multivariate data where the response variable is positively skewed. Parameter estimation in the IGRM is estimated via the maximum likelihood estimator (MLE). While the MLE demonstrates optimal performance under conditions of independent explanatory variables, its efficacy is substantially compromised in the presence of high correlation between explanatory variables, which is known as multicollinearity. This phenomenon leads to inflated variances and standard errors in the coefficient estimates, thereby undermining their statistical efficiency and reliability. To address this inferential challenge, this study introduced a novel modified Liu estimator specifically designed for the IGRM. The proposed estimator aims to mitigate the adverse effects of multicollinearity and enhance the precision of the regression coefficients. The performance of the proposed estimator was rigorously evaluated against the MLE and other established biased estimators. A comprehensive Monte Carlo simulation was utilized for this assessment, whose results indicate the outperformance of the proposed estimator. To further substantiate the scientific utility of this estimator, two applications utilizing real-world multivariate medical data were conducted; the results of these applications were consistent with and reinforced the findings of the simulation study. The synthesized evidence from the simulation study and real-world data analysis suggests that the proposed estimator consistently outperforms competing estimators for the IGRM, achieving greater stability and reliability in the results.

Keywords: inverse Gaussian model; multicollinearity; biased estimator; modified Liu estimators; multivariate medical data

Mathematics Subject Classification: 62J12, 62J07, 62J10, 62P10

1. Introduction

The inverse Gaussian regression model (IGRM) is a helpful alternative to the linear regression model (LRM) when dealing with continuous response variables that are positively skewed and fit an inverse Gaussian distribution. The IGRM is useful in modeling non-normal data in a wide variety of fields like industrial engineering, medical sciences, reliability engineering, chemometrics, social science, and economics [1]. The IGRM can provide the desired flexibility in instances when the traditional Gaussian assumptions are insufficient. The IGRM is also being developed to meet the needs of a wider range of applications. For instance, Fang et al. [2] formulated multivariate degradation models based on IG processes with correlated random effects, and model validation was addressed through novel diagnostics by Amin et al. [3]. The IGRM is also a useful instrument in chemometrics. For example, Akram et al. [4] used it for the quantitative structure–activity relationship (QSAR) modeling. Within the IG framework, Nisa et al. [5] proposed novel multivariate adaptive regression splines. Nisa et al. [6] and Lee et al. [7] contrasted worsening IGRM to gamma regression, and in doing so recognized IGRM as unequivocally advantageous with public health analytics, urban studies, and even quality control of manufacturing processes, as Amin et al. [8] illustrated. Nevertheless, the application of this model is contingent upon its assumptions and certainly involves the complexity associated with its computational demands.

Multicollinearity among explanatory variables poses unstable estimable parameters, and regression coefficients become less interpretable [9, 10]. Most real-life scenarios illustrate this case and provide solution methods with added complications to facilitate estimation and overcome the challenges of real-life scenarios. In cases of extreme skewness, the IGRM performs better than some others, such as gamma regression. With these wide applicability horizons, as with the incomplete estimation of the IGRM for methodologies to combat the multicollinearity problem, biased estimation techniques have been the focus of attention [11, 12].

In the presence of multicollinearity in regression analysis, some biased estimation techniques offer better options than the conventional maximum likelihood estimator (MLE). Hoerl and Kennard [13] pioneered the introduction of the ridge regression estimator, while Kejian [14], later in an example of the Liu estimator, showed increased efficiency in linear regression model applications. Segerstedt [15] extended the applicability of ridge regression to generalized linear models (GLMs), demonstrating that when the shrinkage parameter is optimally set, the mean squared error (MSE) of coefficient estimates can be reduced. The ridge regression methods have, however, been hampered by the complexity in their parameter estimation process. Specifically, the latter become non-linear functions of the shrinkage parameter, k. Conversely, Kejian [14] developed the Liu estimator, which relies linearly on its shrinkage parameter, d, that is, $0 < d < 1$. The last few years have witnessed quite an impressive increase in the family of biased estimators that have been developed to counter multicollinearity. Refining the pioneering ridge and Liu estimators, there is the adjusted Liu [16], modified Liu [17], classical Liu-type [18], and modified ridge-type [19], as well as the restricted estimator of [20], among others.

Innovative approaches for biased estimation have recently been applied to enhance the Liu estimator. This collection includes the modified one-parameter Liu estimator by Lukman et al. [16] and then two parameters by Abonazel [21]. The mentioned studies have proven to have a notable positive effect on practical applications, as shown by Abonazel [22], Abdelwahab et al. [23], Fayomi et al. [24], and Hammad et al. [25]. While these estimators have proven their worth in regression, their applicability in the IGRM has not been explored. This study presents a novel modified two-parameter Liu estimator for the IGRM, which adequately tackles multicollinearity in the framework. Theoretical comparisons, Monte Carlo simulations, and real experiments from medical and biological research have shown the estimator to have remarkable performance and utility in the medical and biological fields. The proposed method demonstrates practical and scientific benefits by providing improved stability and reliability concerning the estimation of parameters.

The structure of this paper is as follows: In Section 2, we describe the IGRM with its MLE approach and elaborate on given biased estimators used to solve issues related to multicollinearity. In Section 3, we discuss our new proposed estimator, and we detail the statistics, the theoretical comparison over the existing estimators, and the methods used for the selection of the biasing parameters. Section 4 describes a Monte Carlo simulation analysis of how the existing estimators and proposed estimators compare with respect to the estimators' performance. In Section 5, we provide a real-life demonstration of the scientific relevance of the proposed estimator. Section 6 contains the conclusion, the summary of the results from the study, and information about future works.

2. Overview of the IGRM and estimation methods

The IGRM constitutes a specialized GLM framework designed for analyzing non-negative, positively skewed response variables that follow the IG distribution. The IG distribution, denoted as $IG(\mu, \tau)$, is characterized by its probability density function:

$$f(y; \mu, \tau) = \left(\frac{\tau}{2\pi y^3}\right)^{1/2} \exp\left[-\frac{\tau(y-\mu)^2}{2y\mu^2}\right], \quad y > 0, \tau > 0, \mu > 0, \tag{2.1}$$

where μ represents the mean parameter with $\mathrm{E}(Y) = \mu$, and τ denotes the precision parameter governing the dispersion, such that $\mathrm{Var}(Y) = \mu^3/\tau$.

Within the GLM framework, the IGRM employs a strictly monotonic and continuously differentiable link function $g(\cdot)$ to relate the mean response $\mu_i = \mathrm{E}(y_i)$ to the linear predictor $\eta_i = x_i'\beta$. Here, x_i denotes the i-th row of the design matrix X containing p covariates, and $\beta \in \mathrm{R}^{p+1}$ represents the vector of regression parameters, including the intercept. The mean response is obtained through the inverse link function:

$$\mu_i = g^{-1}(\eta_i) = g^{-1}(x_i'\beta). \tag{2.2}$$

For the IGRM, the canonical link function corresponds to the reciprocal square transformation $g(\mu_i) = 1/\mu_i^2$, yielding $\mu_i = 1/\sqrt{x_i'\beta}$. Alternative link functions include the logarithmic link $g(\mu_i) = \log(\mu_i)$, which ensures positive mean estimates through $\mu_i = \exp(x_i'\beta)$.

Parameter estimation proceeds via MLE, with the log-likelihood function for n independent observations given by:

$$\ell(\beta,\tau)=\sum_{i=1}^{n}\left\{\frac{1}{\tau}\left(\frac{y_i x_i'\beta}{2}-\sqrt{x_i'\beta}\right)-\frac{1}{2\tau y_i}-\frac{3}{2}\log(2\pi y_i^3)-\frac{1}{2}\log\tau\right\}. \tag{2.3}$$

The score equations are obtained by differentiating (2.3) with respect to β:

$$S(\beta)=\frac{\partial\ell}{\partial\beta}=\frac{1}{2\tau}\sum_{i=1}^{n}\left(y_i-\frac{1}{\sqrt{x_i'\beta}}\right)x_i=0. \tag{2.4}$$

Solving Eq (2.4) requires iterative numerical solutions due to its nonlinearity. The Fisher scoring algorithm provides an efficient solution through the updating scheme:

$$\beta^{(r+1)}=\beta^{(r)}+\mathcal{I}^{-1}(\beta^{(r)})S(\beta^{(r)}), \tag{2.5}$$

where $\mathcal{I}(\beta)=-\mathrm{E}\left[\frac{\partial^2\ell}{\partial\beta\partial\beta'}\right]$ denotes the Fisher information matrix. This procedure is algebraically equivalent to iteratively reweighted least squares, yielding the final estimator:

$$\hat{\beta}_{\text{IGMLE}}=V^{-1}X'\hat{W}\hat{e}, \tag{2.6}$$

where $V=X'\hat{W}X$ with weight matrix $\hat{W}=\mathrm{diag}(\hat{\mu}_i^3)$ and adjusted responses $\hat{e}_i=\hat{\eta}_i+(y_i-\hat{\mu}_i)/\hat{\mu}_i^3$. The asymptotic covariance matrix is given by:

$$\mathrm{Cov}(\hat{\beta}_{\text{IGMLE}})=\tau\,V^{-1}. \tag{2.7}$$

The decomposition is facilitated by the spectral decomposition of the matrix $X'\hat{W}X$, which is central to the covariance structure of the estimator. Specifically, let Υ denote the matrix of eigenvectors such that $X'\hat{W}X=\Upsilon\Theta\Upsilon'$ with $\Theta=\mathrm{diag}(\theta_1,\ldots,\theta_{p+1})$ containing the corresponding eigenvalues. Then, the matrix MSE (MMSE) and MSE decompose into variance components as follows:

$$\mathrm{MMSE}(\hat{\beta}_{\text{IGMLE}})=\tau\Upsilon\Theta^{-1}\Upsilon^{\top}, \tag{2.8}$$

$$\mathrm{MSE}(\hat{\beta}_{\text{IGMLE}})=\tau\,\mathrm{Tr}\left[\mathrm{MMSE}(\hat{\beta}_{\text{IGMLE}})\right]=\tau\sum_{j=1}^{p+1}\frac{1}{\theta_j}. \tag{2.9}$$

2.1. Inverse Gaussian ridge regression estimator

Multicollinearity present among the explanatory variables contributes to the ill-conditioning of the V matrix. The effect is shown by small eigenvalues that cause the variances of the MLEs to be inflated. This leads to unstable estimates of the parameters having large magnitudes, which makes them harder to interpret. To resolve this, Amin et al. [26] proposed the inverse Gaussian ridge regression estimator (IGRRE), which is a generalization of the seminal work of Hoerl and Kennard [13] to the IGRM case, and is given as follows:

$$\hat{\beta}_{\text{IGRRE}}=(V+kI)^{-1}V\hat{\beta}_{\text{IGMLE}}, \tag{2.10}$$

where $k>0$ represents the shrinkage parameter and I denotes the $(p+1)\times(p+1)$ identity matrix. Note that as $k\to 0$, the IGRRE converges to the IGMLE.

The MMSE decomposes into covariance and squared bias components:

$$\text{MMSE}(\hat{\beta}_{\text{IGRRE}}) = \tau\Upsilon\Theta_k^{-1}\Theta\Theta_k^{-1}\Upsilon' + k^2\Upsilon\Theta_k^{-1}\alpha\alpha'\Theta_k^{-1}\Upsilon'. \tag{2.11}$$

The MSE is obtained by taking the trace of the MMSE for the IGRRE, which is given:

$$\text{MSE}(\hat{\beta}_{\text{IGRRE}}) = \tau\sum_{j=1}^{p+1}\frac{\theta_j}{(\theta_j+k)^2} + k^2\sum_{j=1}^{p+1}\frac{\alpha_j^2}{(\theta_j+k)^2}, \tag{2.12}$$

where $\Theta_k = \text{diag}(\theta_1 + k, \ldots, \theta_{p+1} + k)$ and $\alpha = \Upsilon'\beta$.

2.2. Inverse Gaussian Liu estimator

Building on the foundational work of Kejian [14], Hammood et al. [27] proposed the inverse Gaussian Liu estimator (IGLE) as an alternative biased estimation approach to address multicollinearity in the IGRM. The IGLE is defined as:

$$\hat{\beta}_{\text{IGLE}} = (V + I)^{-1}(V + dI)\hat{\beta}_{\text{IGMLE}}, \tag{2.13}$$

where d is the Liu parameter with $d \in (0, 1)$. The estimator exhibits two important limiting cases: as $d \to 1$, it converges to the IGMLE, while for $d < 1$, it produces coefficient estimates with a reduced Euclidean norm compared to the IGMLE.

The MMSE decomposes into covariance and bias components:

$$\text{MMSE}(\hat{\beta}_{\text{IGLE}}) = \tau\Upsilon\Theta_I^{-1}\Theta_d\Theta^{-1}\Theta_d\Theta_I^{-1}\Upsilon' + (d-1)^2\Upsilon\Theta_I^{-1}\beta\beta'\Theta_I^{-1}\Upsilon'. \tag{2.14}$$

Taking the trace yields the MSE as:

$$\text{MSE}(\hat{\beta}_{\text{IGLE}}) = \tau\sum_{j=1}^{p+1}\frac{(\theta_j+d)^2}{\theta_j(\theta_j+1)^2} + (d-1)^2\sum_{j=1}^{p+1}\frac{\alpha_j^2}{(\theta_j+1)^2}, \tag{2.15}$$

where $\Theta_I = \text{diag}(\theta_1 + 1, \ldots, \theta_{p+1} + 1)$ and $\Theta_d = \text{diag}(\theta_1 + d, \ldots, \theta_{p+1} + d)$.

2.3. Inverse Gaussian modified one-parameter Liu estimator

Building upon the modified one-parameter Liu estimator introduced by Lukman et al. [16] for linear regression models, Akram et al. [4] proposed the inverse Gaussian modified one-parameter Liu estimator (IGMOPLE) for addressing multicollinearity in the IGRM. The IGMOPLE is defined as:

$$\hat{\beta}_{\text{IGMOPLE}} = (V + I)^{-1}(V - d_0 I)\hat{\beta}_{\text{IGMLE}}, \tag{2.16}$$

where d_0 is the modified one-parameter Liu shrinkage parameter with $d_0 \in (0, 1)$. The estimator has two key behaviors: when $d \to 1$, it converges to the IGMLE. But when d is less than 1, it gives estimates with a smaller overall size compared to the IGMLE.

The MMSE decomposes into covariance and bias components:

$$\text{MMSE}(\hat{\beta}_{\text{IGMOPLE}}) = \tau\Upsilon\Theta_I^{-1}\Theta_{d_0}\Theta^{-1}\Theta_{d_0}\Theta_I^{-1}\Upsilon' + (d_0+1)^2\Upsilon\Theta_I^{-1}\beta\beta'\Theta_I^{-1}\Upsilon'. \tag{2.17}$$

Taking the trace yields the MSE as:

$$\mathrm{MSE}(\hat{\beta}_{\mathrm{IGMOPLE}}) = \tau \sum_{j=1}^{p+1} \frac{(\theta_j - d_0)^2}{\theta_j(\theta_j+1)^2} + (d_0+1)^2 \sum_{j=1}^{p+1} \frac{\alpha_j^2}{(\theta_j+1)^2}, \tag{2.18}$$

where $\Theta_{d_0} = \mathrm{diag}(\theta_1 - d_0, \ldots, \theta_{p+1} - d_0)$.

3. Proposed estimator

Building upon recent advancements in biased estimation methods, Abonazel [21] developed a modified two-parameter Liu estimator for the LRM. Subsequent research by Abonazel [22] and Hammad et al. [25] has extended this methodology to various regression frameworks. In this study, we introduce a novel modified two-parameter Liu estimator specifically designed for the IGRM, which effectively addresses multicollinearity issues while demonstrating superior performance in minimizing MSE compared to existing estimators. The proposed inverse Gaussian modified two-parameter Liu estimator (IGMTPLE) is derived through the augmentation approach by incorporating the constraint $-(k+d_0)\hat{\beta} = \beta + \varepsilon$ into the MLE framework, yielding:

$$\hat{\beta}_{\mathrm{IGMTPLE}} = (V+I)^{-1}(V-(k+d_0)I)\hat{\beta}_{\mathrm{IGMLE}}, \tag{3.1}$$

where k and d_0 denote the modified two-parameter Liu shrinkage, with $k > 0$ and $0 < d_0 < 1$. When k gets close to 0, it becomes the same as the IGMOPLE.

The estimator's statistical properties are characterized by:

$$\mathrm{Bias}(\hat{\beta}_{\mathrm{IGMTPLE}}) = \mathrm{E}\left(\hat{\beta}_{\mathrm{IGMTPLE}}\right) - \beta = -\Upsilon\Theta_I^{-1}\alpha(k+d_0+1), \tag{3.2}$$

$$\begin{aligned} \mathrm{Cov}(\hat{\beta}_{\mathrm{IGMTPLE}}) &= \mathrm{E}\left[\left(\hat{\beta}_{\mathrm{IGMTPLE}} - \mathrm{E}(\hat{\beta}_{\mathrm{IGMTPLE}})\right)\left(\hat{\beta}_{\mathrm{IGMTPLE}} - \mathrm{E}(\hat{\beta}_{\mathrm{IGMTPLE}})\right)'\right] \\ &= \tau\Upsilon\Theta_I^{-1}\Theta_{k,d_0}\Theta_{k,d_0}\Theta_I^{-1}\Upsilon'. \end{aligned} \tag{3.3}$$

Accordingly, the MMSE and the scalar MSE of the IGMTPLE are given, respectively, by:

$$\begin{aligned} \mathrm{MMSE}\left(\hat{\beta}_{\mathrm{IGMTPLE}}\right) &= \mathrm{E}\left((\hat{\beta}_{\mathrm{IGMTPLE}} - \beta)(\hat{\beta}_{\mathrm{IGMTPLE}} - \beta)'\right) \\ &= \mathrm{Cov}\left(\hat{\beta}_{\mathrm{IGMTPLE}}\right) + \mathrm{Bias}\left(\hat{\beta}_{\mathrm{IGMTPLE}}\right)\mathrm{Bias}'\left(\hat{\beta}_{\mathrm{IGMTPLE}}\right) \\ &= \tau\Upsilon\Theta_I^{-1}\Theta_{k,d_0}\Theta_{k,d_0}\Theta_I^{-1}\Upsilon' + (k+d_0+1)^2\Upsilon\Theta_I^{-1}\alpha\alpha'\Theta_I^{-1}\Upsilon', \end{aligned} \tag{3.4}$$

$$\begin{aligned} \mathrm{MSE}\left(\hat{\beta}_{\mathrm{IGMTPLE}}\right) &= \mathrm{Tr}\left[\mathrm{MMSE}\left(\hat{\beta}_{\mathrm{IGMTPLE}}\right)\right] \\ &= \tau \sum_{j=1}^{p+1} \frac{(\theta_j - (k+d_0))^2}{\theta_j(\theta_j+1)^2} + (k+d_0+1)^2 \sum_{j=1}^{p+1} \frac{\alpha_j^2}{(\theta_j+1)^2}, \end{aligned} \tag{3.5}$$

where $\Theta_{k,d_0} = \mathrm{diag}(\theta_1 - (k+d_0), \ldots, \theta_{p+1} - (k+d_0))$. The optimal shrinkage parameters are derived through MSE minimization:

$$k_{\text{opt}} = \frac{\theta_j(\hat{\tau} - \alpha_j^2)}{\hat{\tau} + \theta_j\alpha_j^2} - d_0, \tag{3.6}$$

$$d_{0(\text{opt})} = \frac{\theta_j(\hat{\tau} - \alpha_j^2) - k(\hat{\tau} + \theta_j\alpha_j^2)}{\hat{\tau} + \theta_j\alpha_j^2}. \tag{3.7}$$

3.1. Theoretical comparison based on MMSE and scalar MSE

Lemma 1. *Let $M > 0$ be a positive definite matrix and α a vector. Then the matrix $M - \alpha\alpha'$ is positive semidefinite if and only if (iff) $\alpha' M^{-1}\alpha \leq 1$ [28].*

Lemma 2. *Consider two linear estimators $\hat{\gamma}_1 = B_1 y$ and $\hat{\gamma}_2 = B_2 y$ of parameter γ. If $D = Cov(\hat{\gamma}_1) - Cov(\hat{\gamma}_2) > 0$, then the difference in their MMSE satisfies $MMSE(\hat{\gamma}_1) - MMSE(\hat{\gamma}_2) > 0$ iff $c_1'(D + c_2c_2')^{-1}c_2 < 1$, where $MMSE(\hat{\gamma}_j) = Cov(\hat{\gamma}_j) + c_jc_j'$ and c_j denotes the bias vector of $\hat{\gamma}_j$ [29].*

Theorem 1. *Under the IGRM, for any shrinkage parameters $k > 0$ and $d_0 > 0$, the IGMTPLE ($\hat{\beta}_{IGMTPLE}$) dominates the IGMLE ($\hat{\beta}_{IGMLE}$) in terms of MMSE, that is, $MMSE\left(\hat{\beta}_{IGMLE}\right) - MMSE\left(\hat{\beta}_{IGMTPLE}\right) > 0$ iff:*

$$\epsilon_1' \left[\hat{\tau}\left(\Upsilon\Theta^{-1}\Upsilon' - \Upsilon\Theta_I^{-1}\Theta_{k,d_0}\Theta_{k,d_0}\Theta_I^{-1}\Upsilon'\right)\right]^{-1} \epsilon_1 < 1,$$

where ϵ_1 is the bias of the IGMTPLE.

Proof. The difference in MMSE between the estimators is:

$$\text{MMSE}\left(\hat{\beta}_{\text{IGMLE}}\right) - \text{MMSE}\left(\hat{\beta}_{\text{IGMTPLE}}\right) = \hat{\tau}\left(\Upsilon\Theta^{-1}\Upsilon' - \Upsilon\Theta_I^{-1}\Theta_{k,d_0}\Theta_{k,d_0}\Theta_I^{-1}\Upsilon'\right) - \epsilon_1'\epsilon_1. \tag{3.8}$$

Expressed in terms of MMSE, the scalar MSE difference becomes:

$$\begin{aligned}\text{MSE}\left(\hat{\beta}_{\text{IGMLE}}\right) - \text{MSE}\left(\hat{\beta}_{\text{IGMTPLE}}\right) &= \hat{\tau}\Upsilon\sum_{j=1}^{p+1}\left[\frac{1}{\theta_j} - \frac{(\theta_j - (k+d_0))^2}{\theta_j(\theta_j+1)^2}\right]\Upsilon' - \epsilon_1'\epsilon_1 \\ &= \hat{\tau}\Upsilon\sum_{j=1}^{p+1}\left[\frac{(\theta_j+1)^2 - (\theta_j - (k+d_0))^2}{\theta_j(\theta_j+1)^2}\right]\Upsilon' - \epsilon_1'\epsilon_1.\end{aligned} \tag{3.9}$$

The matrix $\text{MSE}\left(\hat{\beta}_{\text{IGMLE}}\right) - \text{MSE}\left(\hat{\beta}_{\text{IGMTPLE}}\right)$ is positive definite iff $(\theta_j - 1)^2 > (\theta_j - (k+d_0))^2$, which holds for all j when $k > 0$ and $d_0 > 0$, since $(\theta_j + 1)^2 > (\theta_j - (k+d_0))^2$ for $\theta_j > 0$, $k > 0$, and $d_0 > 0$. The dominance condition then follows from Lemma 2. □

Theorem 2. *Under the IGRM, for any shrinkage parameters $k > 0$ and $d_0 > 0$, the IGMTPLE ($\hat{\beta}_{IGMTPLE}$) dominates the IGRRE ($\hat{\beta}_{IGRRE}$) in terms of MMSE, that is, $MMSE\left(\hat{\beta}_{IGRRE}\right) - MMSE\left(\hat{\beta}_{IGMTPLE}\right) > 0$ iff:*

$$\epsilon_1' \left[\hat{\tau}\left(\Upsilon\Theta_k^{-1}\Theta\Theta_k^{-1}\Upsilon' - \Upsilon\Theta_I^{-1}\Theta_{k,d_0}\Theta_{k,d_0}\Theta_I^{-1}\Upsilon'\right) + \epsilon_2'\epsilon_2\right]^{-1} \epsilon_1 < 1,$$

where ϵ_1 is the bias of the IGMTPLE and ϵ_2 is the bias of the IGRRE.

Proof. The difference in MMSE between the estimators is:

$$\mathrm{MMSE}\left(\hat{\beta}_{\mathrm{IGRRE}}\right) - \mathrm{MMSE}\left(\hat{\beta}_{\mathrm{IGMTPLE}}\right) = \hat{\tau}\left(\Upsilon\Theta_k^{-1}\Theta\Theta_k^{-1}\Upsilon' - \Upsilon\Theta_I^{-1}\Theta_{k,d_0}\Theta_{k,d_0}\Theta_I^{-1}\Upsilon'\right) + \epsilon_2'\epsilon_2 - \epsilon_1'\epsilon_1. \tag{3.10}$$

Expressed in terms of MMSE, the scalar MSE difference becomes:

$$\begin{aligned}\mathrm{MSE}\left(\hat{\beta}_{\mathrm{IGRRE}}\right) - \mathrm{MSE}\left(\hat{\beta}_{\mathrm{IGMTPLE}}\right) &= \hat{\tau}\Upsilon\sum_{j=1}^{p+1}\left[\frac{\theta_j}{(\theta_j+k)^2} - \frac{(\theta_j-(k+d_0))^2}{\theta_j(\theta_j+1)^2}\right]\Upsilon' + \epsilon_2'\epsilon_2 - \epsilon_1'\epsilon_1 \\ &= \hat{\tau}\Upsilon\sum_{j=1}^{p+1}\left[\frac{\theta_j^2(\theta_j+1)^2 - (\theta_j+k)^2(\theta_j-(k+d_0))^2}{\theta_j(\theta_j+k)^2(\theta_j+1)^2}\right]\Upsilon' + \epsilon_2'\epsilon_2 - \epsilon_1'\epsilon_1.\end{aligned} \tag{3.11}$$

The matrix $\mathrm{MSE}\left(\hat{\beta}_{\mathrm{IGRRE}}\right)-\mathrm{MSE}\left(\hat{\beta}_{\mathrm{IGMTPLE}}\right)$ is positive definite iff $\theta_j^2(\theta_j+1)^2 > (\theta_j+k)^2(\theta_j-(k+d_0))^2$, which holds for all j when $k > 0$ and $d_0 > 0$, since $\theta_j^2(\theta_j+1)^2 > (\theta_j+k)^2(\theta_j-(k+d_0))^2$ for $\theta_j > 0$, $k > 0$, and $d_0 > 0$. The dominance condition then follows from Lemma 2. □

Theorem 3. *Under the IGRM, for any shrinkage parameters $k > 0$ and $d_0 > 0$, the IGMTPLE ($\hat{\beta}_{IGMTPLE}$) dominates the IGLE ($\hat{\beta}_{IGLE}$) in terms of MMSE, that is, $MMSE\left(\hat{\beta}_{IGLE}\right) - MMSE\left(\hat{\beta}_{IGMTPLE}\right) > 0$ iff:*

$$\epsilon_1'\left[\hat{\tau}\left(\Upsilon\Theta_I^{-1}\Theta_d\Theta^{-1}\Theta_d\Theta_I^{-1}\Upsilon' - \Upsilon\Theta_I^{-1}\Theta_{k,d_0}\Theta_{k,d_0}\Theta_I^{-1}\Upsilon'\right) + \epsilon_3'\epsilon_3\right]^{-1}\epsilon_1 < 1,$$

where ϵ_1 is the bias of the IGMTPLE and ϵ_3 is the bias of the IGLE.

Proof. The difference in MMSE between the estimators is:

$$\begin{aligned}\mathrm{MMSE}\left(\hat{\beta}_{\mathrm{IGLE}}\right) - \mathrm{MMSE}\left(\hat{\beta}_{\mathrm{IGMTPLE}}\right) &= \hat{\tau}\left(\Upsilon\Theta_I^{-1}\Theta_d\Theta^{-1}\Theta_d\Theta_I^{-1}\Upsilon' - \Upsilon\Theta_I^{-1}\Theta_{k,d_0}\Theta_{k,d_0}\Theta_I^{-1}\Upsilon'\right) \\ &\quad + \epsilon_3'\epsilon_3 - \epsilon_1'\epsilon_1.\end{aligned} \tag{3.12}$$

Expressed in terms of MMSE, the scalar MSE difference becomes:

$$\begin{aligned}\mathrm{MSE}\left(\hat{\beta}_{\mathrm{IGLE}}\right) - \mathrm{MSE}\left(\hat{\beta}_{\mathrm{IGMTPLE}}\right) &= \hat{\tau}\Upsilon\sum_{j=1}^{p+1}\left[\frac{(\theta_j+d)^2}{\theta_j(\theta_j+1)^2} - \frac{(\theta_j-(k+d_0))^2}{\theta_j(\theta_j+1)^2}\right]\Upsilon' + \epsilon_3'\epsilon_3 - \epsilon_1'\epsilon_1 \\ &= \hat{\tau}\Upsilon\sum_{j=1}^{p+1}\left[\frac{(\theta_j+d)^2 - (\theta_j-(k+d_0))^2}{\theta_j(\theta_j+1)^2}\right]\Upsilon' + \epsilon_3'\epsilon_3 - \epsilon_1'\epsilon_1.\end{aligned} \tag{3.13}$$

The matrix $\mathrm{MSE}\left(\hat{\beta}_{\mathrm{IGLE}}\right) - \mathrm{MSE}\left(\hat{\beta}_{\mathrm{IGMTPLE}}\right)$ is positive definite iff $(\theta_j+d)^2 > (\theta_j-(k+d_0))^2$, which holds for all j when $k > 0$ and $d_0 > 0$, since $(\theta_j+d)^2 > (\theta_j-(k+d_0))^2$ for $\theta_j > 0$, $k > 0$, and $d_0 > 0$. The dominance condition then follows from Lemma 2. □

Theorem 4. *Under the IGRM, for any shrinkage parameters $k > 0$ and $d_0 > 0$, the IGMTPLE ($\hat{\beta}_{IGMTPLE}$) dominates the IGMOPLE ($\hat{\beta}_{IGMOPLE}$) in terms of MMSE, that is, $MMSE\left(\hat{\beta}_{IGMOPLE}\right) - MMSE\left(\hat{\beta}_{IGMTPLE}\right) > 0$ iff:*

$$\epsilon_1'\left[\hat{\tau}\left(\Upsilon\Theta_I^{-1}\Theta_{d_0}\Theta^{-1}\Theta_{d_0}\Theta_I^{-1}\Upsilon' - \Upsilon\Theta_I^{-1}\Theta_{k,d_0}\Theta_{k,d_0}\Theta_I^{-1}\Upsilon'\right) + \epsilon_4'\epsilon_4\right]^{-1}\epsilon_1 < 1,$$

where ϵ_1 is the bias of the IGMTPLE and ϵ_4 is the bias of the IGMOPLE.

Proof. The difference in MMSE between the estimators is:

$$\begin{aligned}\text{MMSE}\left(\hat{\beta}_{\text{IGMOPLE}}\right) - \text{MMSE}\left(\hat{\beta}_{\text{IGMTPLE}}\right) = \hat{\tau}\left(\Upsilon\Theta_I^{-1}\Theta_{d_0}\Theta^{-1}\Theta_{d_0}\Theta_I^{-1}\Upsilon' - \Upsilon\Theta_I^{-1}\Theta_{k,d_0}\Theta_{k,d_0}\Theta_I^{-1}\Upsilon'\right)\\ + \epsilon_4'\epsilon_4 - \epsilon_1'\epsilon_1.\end{aligned} \tag{3.14}$$

Expressed in terms of MMSE, the scalar MSE difference becomes:

$$\begin{aligned}\text{MSE}\left(\hat{\beta}_{\text{IGMOPLE}}\right) - \text{MSE}\left(\hat{\beta}_{\text{IGMTPLE}}\right) &= \hat{\tau}\Upsilon\sum_{j=1}^{p+1}\left[\frac{(\theta_j - d_0)^2}{\theta_j(\theta_j+1)^2} - \frac{(\theta_j - (k+d_0))^2}{\theta_j(\theta_j+1)^2}\right]\Upsilon' + \epsilon_4'\epsilon_4 - \epsilon_1'\epsilon_1\\ &= \hat{\tau}\Upsilon\sum_{j=1}^{p+1}\left[\frac{(\theta_j - d_0)^2 - (\theta_j - (k+d_0))^2}{\theta_j(\theta_j+1)^2}\right]\Upsilon' + \epsilon_4'\epsilon_4 - \epsilon_1'\epsilon_1.\end{aligned} \tag{3.15}$$

The matrix $\text{MSE}\left(\hat{\beta}_{\text{IGMOPLE}}\right) - \text{MSE}\left(\hat{\beta}_{\text{IGMTPLE}}\right)$ is positive definite iff $(\theta_j - d_0)^2 > (\theta_j - (k + d_0))^2$, which holds for all j when $k > 0$ and $d_0 > 0$, since $(\theta_j - d_0)^2 > (\theta_j - (k + d_0))^2$ for $\theta_j > 0$, $k > 0$, and $d_0 > 0$. The dominance condition then follows from Lemma 2. □

3.2. Selection of the shrinkage parameter

The performance of biased estimators depends heavily on choosing the right biasing parameters. In the IGRM, we follow approaches from previous research to determine the optimal values for methods like the IGRRE, IGLE, and IGMOPLE. These values are usually selected by minimizing the MSEs, which are often expressed as complex, quadratic, or nonlinear functions involving both the biasing parameters and the unknown parameter α_j.

Following Algamal [11], Lukman et al. [30], and Akram et al. [4], we consider two fundamental ridge parameters for the IGRRE as follows:

$$\hat{k}_1 = \frac{\hat{\tau}}{\sum_{j=1}^{p+1}\hat{\alpha}_j^2}, \quad \hat{k}_2 = \prod_{j=1}^{p+1}\left(\frac{1}{\sqrt{\frac{\hat{\tau}}{\hat{\alpha}_j^2}}}\right)^{1/(p+1)}. \tag{3.16}$$

For the Liu parameter in IGLE, we examine parameters proposed by Hammood et al. [27] and Akram et al. [4] as follows:

$$\hat{d}_1 = \max\left(0, \min\left(\frac{\hat{\tau}(\hat{\alpha}_j^2 - 1)}{1/\theta_j + \hat{\alpha}_j^2}\right)\right), \quad \hat{d}_2 = \max\left(0, \frac{\sum_{j=1}^{p+1}\frac{\hat{\tau}-\hat{\alpha}_j^2}{(1+\theta_j)^2}}{\sum_{j=1}^{p+1}\frac{\hat{\tau}\theta_j+\hat{\alpha}_j^2}{\theta_j(1+\theta_j)^2}}\right). \tag{3.17}$$

Extending the work of Lukman et al. [16], we consider the modified Liu parameter as:

$$\hat{d}_0 = \max\left[0, \min\left(\frac{\theta_j(\hat{\tau} - \hat{\alpha}_j^2)}{\hat{\tau} + \theta_j\hat{\alpha}_j^2}\right)\right]. \tag{3.18}$$

Building on the work of Hammad et al. [25], we propose several novel biasing parameters for the modified two-parameter Liu estimator within the IGMTPLE framework, as detailed below.

First, we set the initial value of $\hat{k}$ to $\hat{k}_1$ as defined in Eq (3.16), and use it to compute $\hat{d}_0$ as follows:

$$\hat{d}_0 = \max\left(0, \min\left(\frac{\theta_j(\hat{\tau} - \hat{\alpha}_j^2) - \hat{k}(1 + \theta_j\hat{\alpha}_j^2)}{\hat{\tau} + \text{median}(\theta_j)\hat{\alpha}_j^2}\right)\right). \tag{3.19}$$

Next, using the obtained value of $\hat{d}_0$, we compute the following biasing parameters $\hat{k}_1$, $\hat{k}_2$, and $\hat{k}_3$ as:

$$\hat{k}_1 = \left(\frac{\min_j(\theta_j)(\hat{\tau} - \min_j(\hat{\alpha}_j^2))}{\hat{\tau} + \min_j(\theta_j)\hat{\alpha}_j^2}\right) - \hat{d}_0, \tag{3.20}$$

$$\hat{k}_2 = \min\left(\frac{\theta_j(\hat{\tau} - \hat{\alpha}_j^2)}{\hat{\tau} + \theta_j\hat{\alpha}_j^2}\right) - \hat{d}_0, \tag{3.21}$$

$$\hat{k}_3 = \frac{1}{2} \cdot \frac{\min\left(\theta_j(\hat{\tau} - \hat{\alpha}_j^2)\right)}{\min\left(\hat{\tau} + \theta_j\hat{\alpha}_j^2\right)} - \hat{d}_0. \tag{3.22}$$

4. Monte Carlo simulation

The simulation study examines the performance of various estimators under different conditions. The response variable y_i is generated from an IG distribution $IG(\mu_i, \tau)$, where the mean structure follows:

$$\mu_i = E(y_i) = \left(\sum_{j=1}^{p+1} \beta_j x_{ij}\right)^{-1/2}, \quad i = 1, \ldots, n; \quad j = 1, \ldots, p+1. \tag{4.1}$$

The explanatory variables are constructed to induce different multicollinearity through the data generation process [31]:

$$x_{ij} = (1 - \rho^2)^{1/2} z_{ij} + \rho z_{i(j+1)}, \tag{4.2}$$

where z_{ij} are independent standard normal variates and ρ determines the intercorrelation strength. Following established simulation practices [25, 32], we consider:

- Multicollinearity level: $\rho \in \{0.80, 0.85, 0.90, 0.95, 0.99\}$.
- Sample size: $n \in \{30, 50, 100, 200, 300, 500\}$.
- Predictor dimensions: $p \in \{3, 7, 10\}$.
- Dispersion parameter: $\tau \in \{0.1, 0.5\}$.

The regression coefficients β are constrained such that $\sum_{j=1}^{p+1} \beta_j^2 = 1$, consistent with common simulation conventions. For each parameter combination, we conduct $R = 5000$ Monte Carlo replications to ensure stable estimates of the MSE:

$$\text{MSE}(\hat{\beta}) = \frac{1}{R} \sum_{i=1}^{R} (\hat{\beta}_i - \beta)'(\hat{\beta}_i - \beta), \tag{4.3}$$

where $(\hat{\beta}_i - \beta)$ denotes the deviation of the estimated parameter vector from the true parameter vector, and R is the total number of simulation replications.

The findings of the simulations shown in Tables 1–6 show how the IGMLE, IGRRE, IGLE, and the new IGMTPLE estimators performed under each of the multiple experimental conditions. These conditions include differences in sample size (n), degree of multicollinearity (ρ), number of predictors (p), and the dispersion parameter (τ).

Table 1. The simulated MSE values of different estimators when $p = 3$ and $\tau = 0.1$.

ρ	n	IGMLE	IGRRE		IGLE		IGMOPLE	IGMTPLE		
		-	$\hat{k}_1$	$\hat{k}_2$	$\hat{d}_1$	$\hat{d}_2$	$\hat{d}_0$	$\hat{k}_1, \hat{d}_0$	$\hat{k}_2, \hat{d}_0$	$\hat{k}_3, \hat{d}_0$
0.80	30	0.05151	0.05021	0.05071	0.04970	0.04970	0.04969	**0.04740**	0.04806	0.05042
	50	0.02616	0.02583	0.02596	0.02572	0.02572	0.02572	**0.02499**	0.02512	**0.02499**
	100	0.01577	0.01554	0.01564	0.01548	0.01548	0.01548	**0.01497**	0.01510	0.01505
	200	0.00731	0.00728	0.00729	0.00726	0.00726	0.00726	0.00721	0.00721	**0.00718**
	300	0.00427	0.00427	0.00427	0.00427	0.00427	0.00427	0.00428	0.00428	0.00429
	500	0.00254	0.00254	0.00254	0.00254	0.00254	0.00254	0.00255	0.00255	0.00257
0.85	30	0.10734	0.10025	0.10263	0.09718	0.09718	0.09718	0.07964	0.08178	**0.07420**
	50	0.03006	0.02954	0.02971	0.02926	0.02926	0.02926	0.02813	0.02827	**0.02791**
	100	0.01875	0.01857	0.01862	0.01844	0.01844	0.01844	0.01803	0.01802	**0.01768**
	200	0.01233	0.01228	0.01230	0.01225	0.01225	0.01225	**0.01214**	0.01216	0.01216
	300	0.00668	0.00666	0.00666	0.00665	0.00665	0.00665	0.00662	0.00662	**0.00659**
	500	0.00318	0.00318	0.00318	0.00318	0.00318	0.00318	0.00318	0.00318	0.00318
0.90	30	0.07185	0.06937	0.06995	0.06769	0.06769	0.06769	0.06167	0.06167	**0.05689**
	50	0.05477	0.05259	0.05325	0.05164	0.05164	0.05164	0.04575	0.04624	**0.04290**
	100	0.03509	0.03415	0.03443	0.03372	0.03372	0.03372	0.03117	0.03136	**0.02985**
	200	0.01782	0.01761	0.01767	0.01750	0.01750	0.01750	0.01693	0.01694	**0.01646**
	300	0.00824	0.00819	0.00820	0.00816	0.00816	0.00816	0.00803	0.00803	**0.00790**
	500	0.00669	0.00666	0.00666	0.00664	0.00664	0.00664	0.00655	0.00655	**0.00647**
0.95	30	0.63733	0.48018	0.54244	0.40193	0.40193	0.40193	0.15127	0.19923	**0.16249**
	50	0.48370	0.37679	0.41654	0.33014	0.33014	0.33014	0.13329	0.16980	**0.12985**
	100	0.21608	0.19087	0.19782	0.17716	0.17716	0.17716	0.11816	0.12165	**0.08162**
	200	0.08239	0.07787	0.07924	0.07577	0.07577	0.07577	0.06344	0.06441	**0.05701**
	300	0.03703	0.03609	0.03636	0.03568	0.03568	0.03568	0.03301	0.03316	**0.03129**
	500	0.02352	0.02310	0.02320	0.02287	0.02287	0.02287	0.02165	0.02163	**0.02021**
0.99	30	4.38805	2.40595	2.90972	0.59857	0.59857	0.59857	0.65823	**0.26810**	0.59643
	50	3.31077	1.74839	2.20699	0.50673	0.50673	0.50673	0.38660	**0.16975**	0.54784
	100	1.14834	0.64965	0.83538	0.43227	0.43227	0.43227	0.12550	**0.11110**	0.13171
	200	0.47893	0.33696	0.39560	0.28804	0.28804	0.28804	**0.06944**	0.11649	0.10189
	300	0.36813	0.28413	0.31478	0.24888	0.24888	0.24888	0.08809	0.11621	**0.07642**
	500	0.21119	0.18426	0.19425	0.17392	0.17392	0.17392	0.10938	0.12077	**0.10214**

Note: Bolded values indicate the minimum MSE in each row.

Table 2. The simulated MSE values of different estimators when $p = 3$ and $\tau = 0.5$.

ρ	n	IGMLE	IGRRE		IGLE		IGMOPLE	IGMTPLE		
		-	$\hat{k}_1$	$\hat{k}_2$	$\hat{d}_1$	$\hat{d}_2$	$\hat{d}_0$	$\hat{k}_1, \hat{d}_0$	$\hat{k}_2, \hat{d}_0$	$\hat{k}_3, \hat{d}_0$
0.80	30	0.29786	0.26837	0.27971	0.25551	0.25551	0.25514	**0.21460**	0.22015	0.22942
	50	0.11545	0.10800	0.11144	0.10642	0.10642	0.10642	**0.09443**	0.09694	0.09606
	100	0.10005	0.09380	0.09706	0.09343	0.09343	0.09327	**0.07977**	0.08472	0.08430
	200	0.03717	0.03674	0.03690	0.03648	0.03648	0.03648	**0.03578**	0.03595	0.03605
	300	0.01916	0.01891	0.01901	0.01882	0.01882	0.01882	**0.01837**	0.01845	0.01839
	500	0.01410	0.01410	0.01410	0.01407	0.01407	0.01407	0.01416	0.01416	0.01433
0.85	30	0.66987	0.52981	0.58939	0.46825	0.46825	0.46461	0.27622	0.30157	**0.27600**
	50	0.17566	0.16377	0.16893	0.15931	0.15931	0.15932	**0.13448**	0.14107	0.14003
	100	0.10430	0.10050	0.10169	0.09817	0.09817	0.09825	0.09003	0.09078	**0.08713**
	200	0.06246	0.06086	0.06147	0.06032	0.06032	0.06032	0.05658	0.05729	**0.05653**
	300	0.02466	0.02438	0.02445	0.02419	0.02419	0.02419	0.02357	0.02357	**0.02308**
	500	0.01525	0.01512	0.01516	0.01506	0.01506	0.01506	0.01476	0.01479	**0.01468**
0.90	30	0.37565	0.32766	0.34695	0.30290	0.30290	0.30105	**0.21803**	0.23473	0.22427
	50	0.29382	0.24822	0.26754	0.23220	0.23220	0.23191	**0.13963**	0.16189	0.14906
	100	0.18314	0.16734	0.17378	0.16254	0.16254	0.16242	0.12448	0.13281	**0.12396**
	200	0.10710	0.10256	0.10416	0.10071	0.10071	0.10071	0.08900	0.09075	**0.08591**
	300	0.04358	0.04263	0.04291	0.04219	0.04219	0.04219	0.03973	0.03989	**0.03841**
	500	0.03799	0.03737	0.03753	0.03704	0.03704	0.03704	0.03531	0.03535	**0.03387**
0.95	30	3.93207	2.74389	3.12859	1.19968	1.19968	1.19812	0.68355	0.58551	**0.50911**
	50	2.59021	1.65651	1.97503	0.82326	0.82326	0.82257	0.25108	0.25511	**0.22184**
	100	0.99637	0.69278	0.80556	0.50195	0.50195	0.50145	**0.13791**	0.19078	0.17925
	200	0.40548	0.32376	0.35860	0.30001	0.30001	0.30001	**0.13236**	0.17360	0.14854
	300	0.20036	0.17974	0.18752	0.17170	0.17170	0.17170	0.12028	0.12966	**0.11281**
	500	0.12458	0.11548	0.11815	0.11104	0.11104	0.11104	0.08716	0.08885	**0.07390**
0.99	30	24.04291	15.18796	16.37264	0.48813	0.48813	0.48815	**0.42871**	0.47653	0.43716
	50	17.90204	10.24984	11.52935	0.46479	0.46479	0.46479	0.37944	0.40706	**0.32141**
	100	6.07598	3.47626	4.04544	0.61969	0.61969	0.61969	0.33293	0.26860	**0.25992**
	200	2.03007	1.00733	1.33118	0.44341	0.44341	0.44329	0.28363	**0.14259**	0.21850
	300	2.27544	1.42291	1.71090	0.66642	0.66642	0.66642	0.24916	0.19205	**0.15041**
	500	0.91576	0.58377	0.71823	0.43882	0.43882	0.43882	**0.08702**	0.13658	0.12089

Note: Bolded values indicate the minimum MSE in each row.

Table 3. The simulated MSE values of different estimators when $p = 7$ and $\tau = 0.1$.

ρ	n	IGMLE	IGRRE		IGLE		IGMOPLE	IGMTPLE		
		-	$\hat{k}_1$	$\hat{k}_2$	$\hat{d}_1$	$\hat{d}_2$	$\hat{d}_0$	$\hat{k}_1, \hat{d}_0$	$\hat{k}_2, \hat{d}_0$	$\hat{k}_3, \hat{d}_0$
0.80	30	0.10846	0.10137	0.10610	0.10314	0.10314	0.10123	0.13400	**0.09460**	0.10845
	50	0.02324	0.02295	0.02314	0.02302	0.02302	0.02295	0.02477	0.02284	**0.02283**
	100	0.01570	0.01552	0.01565	0.01557	0.01557	0.01552	0.01589	**0.01543**	0.01573
	200	0.00716	0.00714	0.00715	0.00715	0.00715	0.00714	0.00721	**0.00714**	0.00717
	300	0.00486	0.00484	0.00485	0.00485	0.00485	0.00484	**0.00476**	0.00480	0.00478
	500	0.00231	0.00231	0.00231	0.00231	0.00231	0.00231	0.00232	**0.00231**	0.00234
0.85	30	0.12668	0.11817	0.12340	0.11975	0.11975	0.11808	**0.08061**	0.09795	0.09230
	50	0.10714	0.10293	0.10514	0.10272	0.10272	0.10265	**0.08487**	0.09119	0.09603
	100	0.04520	0.04408	0.04478	0.04432	0.04432	0.04407	**0.03837**	0.04126	0.03992
	200	0.01141	0.01134	0.01138	0.01135	0.01135	0.01134	0.01127	**0.01122**	0.01190
	300	0.00668	0.00665	0.00667	0.00666	0.00666	0.00665	**0.00654**	0.00659	**0.00654**
	500	0.00280	0.00280	0.00280	0.00280	0.00280	0.00280	0.00281	**0.00280**	0.00283
0.90	30	0.19222	0.17003	0.18457	0.17444	0.17444	0.16726	0.14565	**0.13398**	0.13422
	50	0.14829	0.13862	0.14287	0.13672	0.13672	0.13672	**0.09011**	0.10405	0.10484
	100	0.05982	0.05773	0.05892	0.05793	0.05793	0.05773	0.04549	0.05135	**0.04538**
	200	0.01360	0.01346	0.01355	0.01350	0.01350	0.01346	**0.01268**	0.01313	0.01299
	300	0.00767	0.00764	0.00766	0.00765	0.00765	0.00764	**0.00749**	0.00756	0.00752
	500	0.00548	0.00546	0.00547	0.00546	0.00546	0.00546	0.00540	0.00542	**0.00538**
0.95	30	1.88052	1.17197	1.42453	0.77241	0.77241	0.77224	**0.52032**	0.60750	0.53181
	50	0.35432	0.30676	0.32870	0.29759	0.29759	0.29759	**0.13116**	0.17195	0.22632
	100	0.50132	0.44885	0.47099	0.42679	0.42679	0.42679	**0.18000**	0.25172	0.21784
	200	0.08539	0.08262	0.08383	0.08214	0.08214	0.08214	0.06511	0.07092	**0.05730**
	300	0.06100	0.05962	0.06015	0.05927	0.05927	0.05927	0.05039	0.05300	**0.04733**
	500	0.03065	0.03037	0.03047	0.03027	0.03027	0.03027	**0.02836**	0.02883	0.03705
0.99	30	12.58828	7.82297	9.35000	1.64915	1.64915	1.64915	1.29621	**1.24746**	1.31941
	50	2.82485	1.82918	2.29779	1.23910	1.23910	1.23910	0.90357	**0.44020**	0.90035
	100	1.82663	1.21437	1.50585	0.92552	0.92552	0.92552	0.70898	**0.31772**	0.80846
	200	1.13093	0.81564	0.97986	0.72360	0.72360	0.72360	0.41738	**0.23138**	0.52282
	300	0.38241	0.33210	0.35558	0.31788	0.31788	0.31788	**0.11442**	0.17326	2.40102
	500	0.20950	0.19750	0.20270	0.19438	0.19438	0.19438	**0.12330**	0.14697	0.21046

Note: Bolded values indicate the minimum MSE in each row.

Table 4. The simulated MSE values of different estimators when $p = 7$ and $\tau = 0.5$.

ρ	n	IGMLE	IGRRE		IGLE		IGMOPLE	IGMTPLE		
		-	$\hat{k}_1$	$\hat{k}_2$	$\hat{d}_1$	$\hat{d}_2$	$\hat{d}_0$	$\hat{k}_1, \hat{d}_0$	$\hat{k}_2, \hat{d}_0$	$\hat{k}_3, \hat{d}_0$
0.80	30	0.70057	0.55537	0.63885	0.55081	0.55081	0.53937	0.54129	0.46187	**0.45908**
	50	0.13244	0.12840	0.13089	0.12873	0.12873	0.12828	0.16184	**0.12809**	0.14086
	100	0.08830	0.08526	0.08734	0.08593	0.08593	0.08522	0.10024	**0.08458**	0.09118
	200	0.03572	0.03498	0.03546	0.03515	0.03515	0.03497	0.03424	**0.03388**	0.03399
	300	0.02381	0.02349	0.02369	0.02356	0.02356	0.02349	0.02296	**0.02295**	0.02303
	500	0.01085	0.01081	0.01083	0.01079	0.01079	0.01080	0.01080	**0.01071**	0.01109
0.85	30	0.70472	0.62338	0.67115	0.61944	0.61944	0.60765	0.61361	**0.35034**	0.51190
	50	0.57100	0.51127	0.54590	0.51352	0.51352	0.50586	**0.34945**	0.38090	0.47041
	100	0.18884	0.16801	0.18147	0.17301	0.17301	0.16734	**0.13031**	0.13154	0.13049
	200	0.05194	0.05058	0.05141	0.05080	0.05080	0.05057	0.04834	**0.04816**	0.05973
	300	0.03512	0.03456	0.03493	0.03471	0.03471	0.03455	**0.03319**	0.03360	0.03353
	500	0.01504	0.01494	0.01500	0.01495	0.01495	0.01494	**0.01469**	0.01471	0.01513
0.90	30	1.04385	0.83264	0.95770	0.80676	0.80676	0.77327	0.68761	**0.49088**	0.50035
	50	0.83302	0.69019	0.76666	0.65609	0.65609	0.65197	**0.32796**	0.36227	1.32350
	100	0.32057	0.27114	0.30054	0.27663	0.27663	0.26935	0.23506	**0.18649**	0.20383
	200	0.06674	0.06383	0.06573	0.06465	0.06465	0.06380	**0.05367**	0.05796	0.05660
	300	0.03306	0.03224	0.03277	0.03245	0.03245	0.03223	**0.02897**	0.03049	0.03034
	500	0.02795	0.02772	0.02783	0.02769	0.02769	0.02769	**0.02630**	0.02682	0.02653
0.95	30	12.53613	8.85302	9.78873	2.23171	2.23171	2.23065	**1.73832**	1.78725	1.84573
	50	1.94867	1.44554	1.72683	1.25062	1.25062	1.24429	1.05679	**0.53543**	1.20334
	100	2.81584	2.19513	2.52014	1.72915	1.72915	1.72561	0.96053	0.73761	**0.71686**
	200	0.45418	0.39751	0.42665	0.38972	0.38972	0.38957	**0.15221**	0.23469	0.36108
	300	0.33902	0.30499	0.32130	0.29799	0.29799	0.29799	**0.13266**	0.18997	0.28835
	500	0.17717	0.16949	0.17253	0.16696	0.16696	0.16696	**0.12085**	0.13437	0.22218
0.99	30	68.45847	49.84375	55.02701	2.11871	2.11871	2.11871	1.87716	**1.27748**	1.70110
	50	15.52621	10.53027	12.55010	2.24160	2.24160	2.24160	1.31341	**1.13216**	1.58958
	100	9.15929	5.83413	7.24095	1.91958	1.91958	1.91958	1.23436	**1.00661**	1.16567
	200	6.05710	3.94248	4.89643	1.70439	1.70439	1.70439	0.83528	0.74147	**0.69279**
	300	1.77268	1.21725	1.48636	0.96566	0.96566	0.96566	0.41012	**0.38868**	0.43874
	500	0.98283	0.76319	0.87961	0.70948	0.70948	0.70948	**0.19649**	0.28495	0.23591

Note: Bolded values indicate the minimum MSE in each row.

Table 5. The simulated MSE values of different estimators when $p = 10$ and $\tau = 0.1$.

ρ	n	IGMLE	IGRRE		IGLE		IGMOPLE	IGMTPLE		
		-	$\hat{k}_1$	$\hat{k}_2$	$\hat{d}_1$	$\hat{d}_2$	$\hat{d}_0$	$\hat{k}_1, \hat{d}_0$	$\hat{k}_2, \hat{d}_0$	$\hat{k}_3, \hat{d}_0$
0.80	30	0.19623	0.17989	0.19079	0.18423	0.18423	0.17962	0.14982	**0.14333**	0.17361
	50	0.03186	0.03151	0.03174	0.03161	0.03161	0.03151	0.03268	**0.03110**	0.03331
	100	0.03139	0.03100	0.03125	0.03109	0.03109	0.03100	0.03020	0.03004	**0.02986**
	200	0.01432	0.01425	0.01429	0.01425	0.01425	0.01425	**0.01387**	0.01400	0.01450
	300	0.00313	0.00312	0.00312	0.00312	0.00312	0.00312	**0.00308**	**0.00308**	**0.00308**
	500	0.00207	0.00207	0.00207	0.00207	0.00207	0.00207	0.00211	0.00208	0.00222
0.85	30	0.43445	0.36476	0.40910	0.37562	0.37562	0.36179	0.28874	**0.20699**	0.33908
	50	0.05873	0.05702	0.05814	0.05745	0.05745	0.05701	**0.04809**	0.05206	0.05160
	100	0.02569	0.02518	0.02552	0.02534	0.02534	0.02518	0.02582	**0.02429**	0.02811
	200	0.00511	0.00510	0.00511	0.00510	0.00510	0.00510	0.00511	**0.00508**	0.00538
	300	0.00503	0.00502	0.00503	0.00503	0.00503	0.00502	0.00506	**0.00501**	0.00506
	500	0.00246	0.00246	0.00246	0.00246	0.00246	0.00246	**0.00244**	0.00245	0.00245
0.90	30	0.56236	0.43117	0.51267	0.44875	0.44875	0.42141	0.37759	**0.25215**	0.27355
	50	0.05926	0.05759	0.05862	0.05797	0.05797	0.05758	**0.05241**	0.05307	0.05990
	100	0.04921	0.04805	0.04876	0.04827	0.04827	0.04804	**0.03948**	0.04382	0.03989
	200	0.02739	0.02705	0.02726	0.02713	0.02713	0.02705	0.02428	0.02579	**0.02360**
	300	0.01061	0.01055	0.01059	0.01057	0.01057	0.01055	0.01010	0.01036	**0.01001**
	500	0.00401	0.00401	0.00401	0.00401	0.00401	0.00401	**0.00398**	0.00399	0.00400
0.95	30	1.14966	0.81103	1.00273	0.73395	0.73395	0.69182	0.37728	0.39031	**0.35457**
	50	0.43175	0.37417	0.40524	0.36801	0.36801	0.36800	**0.14667**	0.18837	2.71774
	100	0.22154	0.20206	0.21177	0.20059	0.20059	0.20057	**0.09197**	0.12567	0.82373
	200	0.06098	0.05932	0.06031	0.05960	0.05960	0.05932	**0.04511**	0.05261	0.10886
	300	0.04863	0.04810	0.04830	0.04796	0.04796	0.04796	**0.04299**	0.04451	0.07147
	500	0.00944	0.00942	0.00943	0.00941	0.00941	0.00941	0.00915	0.00925	**0.00809**
0.99	30	20.67271	12.86659	15.37041	2.34151	2.34151	2.34151	**2.27655**	2.28378	2.30604
	50	3.38224	2.26368	2.84606	1.54221	1.54221	1.54221	**0.73226**	0.76893	0.74192
	100	1.71919	1.21032	1.45925	1.01523	1.01523	1.01523	0.68822	**0.51731**	0.65446
	200	0.38960	0.35212	0.36910	0.34311	0.34311	0.34311	**0.15261**	0.20007	0.22573
	300	0.38361	0.35084	0.36615	0.34259	0.34259	0.34259	0.12845	0.20017	**0.12598**
	500	0.11246	0.10904	0.11063	0.10858	0.10858	0.10858	**0.07815**	0.09012	0.13652

Note: Bolded values indicate the minimum MSE in each row.

Table 6. The simulated MSE values of different estimators when $p = 10$ and $\tau = 0.5$.

ρ	n	IGMLE	IGRRE		IGLE		IGMOPLE	IGMTPLE		
		-	$\hat{k}_1$	$\hat{k}_2$	$\hat{d}_1$	$\hat{d}_2$	$\hat{d}_0$	$\hat{k}_1, \hat{d}_0$	$\hat{k}_2, \hat{d}_0$	$\hat{k}_3, \hat{d}_0$
0.80	30	1.06667	0.92720	1.01266	0.90702	0.90702	0.88873	0.70052	**0.54343**	0.80866
	50	0.15657	0.14850	0.15391	0.15069	0.15069	0.14834	0.14290	**0.13219**	0.17935
	100	0.14295	0.13679	0.14077	0.13839	0.13839	0.13668	0.13899	**0.12430**	0.12985
	200	0.06514	0.06403	0.06467	0.06414	0.06414	0.06399	**0.06039**	0.06060	0.14678
	300	0.01651	0.01637	0.01645	0.01639	0.01639	0.01637	0.01767	**0.01639**	0.01795
	500	0.01156	0.01152	0.01154	0.01152	0.01152	0.01152	0.01175	**0.01145**	0.01377
0.85	30	2.92138	2.46099	2.68648	1.94051	1.94051	1.92037	1.24006	**1.14169**	1.53643
	50	0.34327	0.30877	0.33070	0.31611	0.31611	0.30816	0.25191	0.22746	**0.22314**
	100	0.12230	0.11621	0.12010	0.11752	0.11752	0.11614	0.13507	**0.10679**	0.25638
	200	0.02484	0.02468	0.02479	0.02472	0.02472	0.02468	0.02591	**0.02467**	0.03122
	300	0.02570	0.02550	0.02564	0.02556	0.02556	0.02550	0.02762	0.02556	0.02686
	500	0.01164	0.01157	0.01161	0.01159	0.01159	0.01157	0.01156	**0.01146**	0.01171
0.90	30	2.89715	2.34155	2.58895	1.77880	1.77880	1.75643	1.64911	**1.08062**	1.61964
	50	0.28364	0.25134	0.27046	0.25602	0.25602	0.25023	0.22047	**0.21435**	0.23304
	100	0.22740	0.20649	0.21880	0.20985	0.20985	0.20623	0.19063	**0.17401**	0.20081
	200	0.14291	0.13682	0.14073	0.13820	0.13820	0.13678	**0.09313**	0.11655	0.10321
	300	0.04645	0.04550	0.04609	0.04573	0.04573	0.04550	**0.04110**	0.04272	0.04356
	500	0.01858	0.01844	0.01853	0.01847	0.01847	0.01844	**0.01802**	0.01809	0.01837
0.95	30	6.09632	4.21980	5.02622	2.18989	2.18989	2.17898	1.71668	**1.69779**	1.80402
	50	2.22897	1.58418	1.94197	1.33029	1.33029	1.32361	1.15384	**0.98051**	1.26091
	100	0.98305	0.73189	0.86408	0.69930	0.69930	0.69584	0.62395	**0.44115**	0.78767
	200	0.33148	0.29080	0.31612	0.29705	0.29705	0.29015	**0.15009**	0.15182	0.19506
	300	0.21997	0.20931	0.21390	0.20672	0.20672	0.20672	0.12672	**0.10256**	0.14689
	500	0.04971	0.04910	0.04938	0.04903	0.04903	0.04903	**0.04339**	0.04558	0.05307
0.99	30	123.69731	90.02415	101.31767	3.12744	3.12744	3.12744	3.04704	**2.23733**	3.10874
	50	18.86697	12.91225	15.48196	2.89201	2.89201	2.89201	2.21442	**1.74474**	2.45786
	100	7.91361	4.97790	6.25845	2.14369	2.14369	2.14369	**1.14148**	1.67914	1.30512
	200	2.14886	1.59295	1.89774	1.32778	1.32778	1.32778	0.86329	**0.51990**	0.98525
	300	2.12162	1.54795	1.86776	1.31046	1.31046	1.31046	0.66049	**0.45261**	0.89495
	500	0.53696	0.46508	0.50415	0.45712	0.45712	0.45712	**0.12194**	0.22029	0.23135

Note: Bolded values indicate the minimum MSE in each row.

A key finding is the consistently superior performance of the proposed IGMTPLE. Across nearly all scenarios, it achieves the lowest MSE compared to the existing estimators. The results clearly demonstrate that multicollinearity adversely affects all estimators, with MSE values increasing substantially as the correlation level (ρ) rises. However, the IGMTPLE exhibits remarkable robustness to this effect. Among its different parameter combinations, the variants utilizing the shrinkage parameters $(\hat{k}_1, \hat{d}_0)$ and $(\hat{k}_2, \hat{d}_0)$ show particularly strong performance.

As expected, sample size has a positive and substantial impact. For any fixed combination of other factors, the MSE of all estimators decreases as the sample size (n) increases, confirming the large-sample properties of the proposed IGMTPLE. The dispersion parameter (τ) also plays a significant role; an increase in τ leads to a gradual increase in MSE across the board. Despite this, the IGMTPLE, particularly the variant with parameters $(\hat{k}_2, \hat{d}_0)$, maintains a distinct efficiency advantage over the IGMLE, IGRRE, and IGLE.

The IGMLE demonstrates the poorest performance in almost all cases, with its MSE being most severely inflated by changes in the simulation factors, especially under high multicollinearity. In contrast, the proposed IGMTPLE, across all its shrinkage parameter combinations, consistently outperforms the benchmark estimators. The version with parameter combination $(\hat{k}_2, \hat{d}_0)$ emerges as the most effective, reliably delivering the smallest MSE.

The results obtained are displayed in Figures 1–4. Multicollinearity negatively impacts estimation in Figure 2. As the correlation level (ρ) increases, particularly at $\rho = 0.99$, there is a sharp rise in the MSE of the IGMLE. The IGMTPLE is biased, but more so, increases in the correlation level have less effect on the IGMTPLE.

Figure 1 shows the positive effect a larger sample size has on estimation. An increase in sample size leads to a noticeable reduction in MSE among all estimators. The IGMTPLE has a smaller MSE at lower sample sizes and reaches a lower MSE at a faster rate as the sample size increases compared to the other estimators.

Additionally, given a fixed sample size, with a simultaneous increase in collinearity, the number of predictors as in Figure 3, or the dispersion parameter as in Figure 4, the simulated MSE values exhibit a steep increase. This illustrates the problem situations where the IGMTPLE's advantage is most needed.

Overall, the simulation study aligns with the theoretically proposed IGMTPLE for the inverse Gaussian regression model with multicollinearity. The IGMTPLEs proposed offer the most substantial reduction of MSE for the given data conditions. While the reduction in MSE is noted for all versions of the IGMTPLE, the best option is the version with the shrinkage parameter combination $(\hat{k}_2, \hat{d}_0)$, which stands out for its effectiveness in addressing multicollinearity. The simulation results and the theoretical properties of the estimators are consistent with each other.

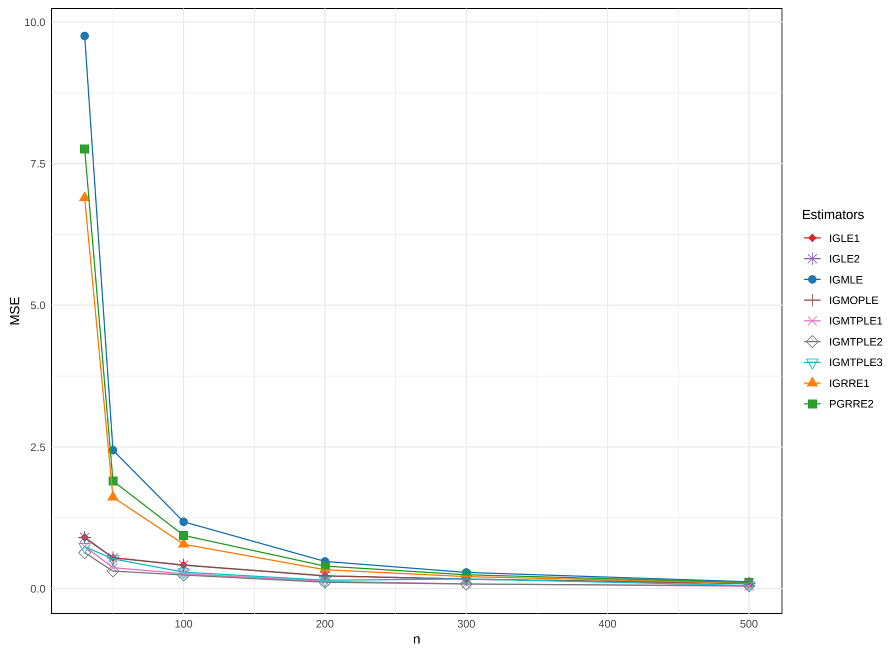

Figure 1. Simulated MSE values of biasing estimators with sample size.

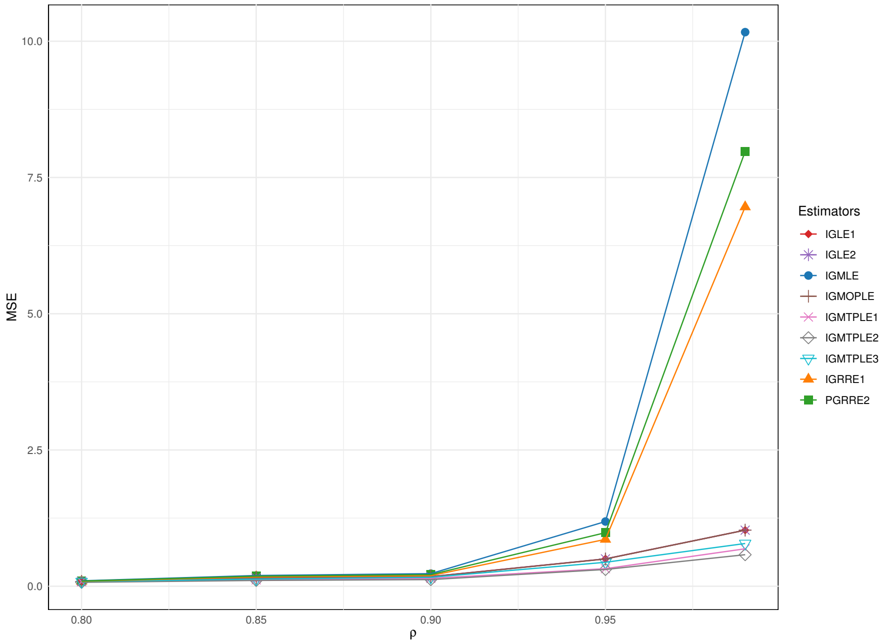

Figure 2. Simulated MSE values of biasing estimators with multicollinearity level.

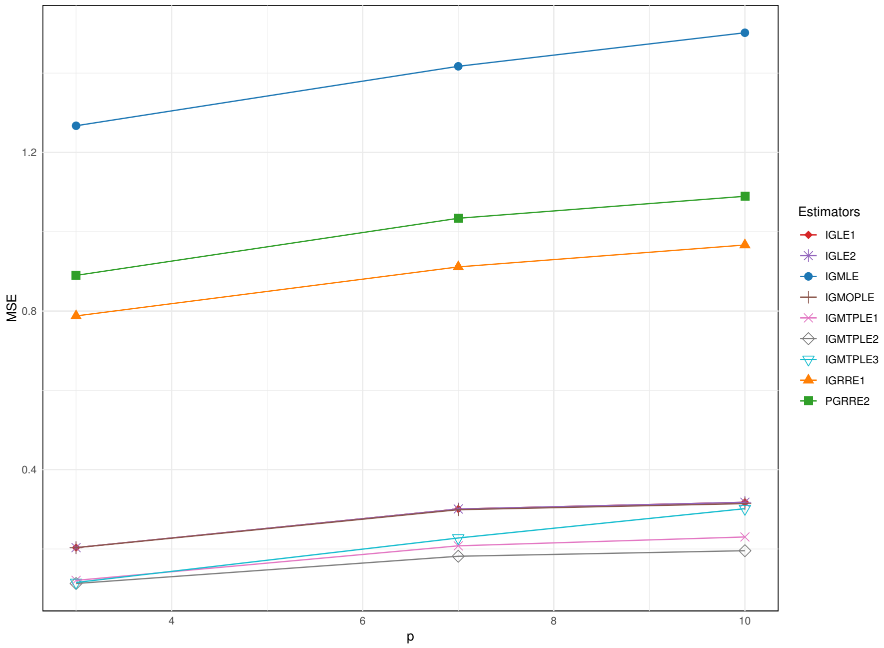

Figure 3. Simulated MSE values of biasing estimators with explanatory variables.

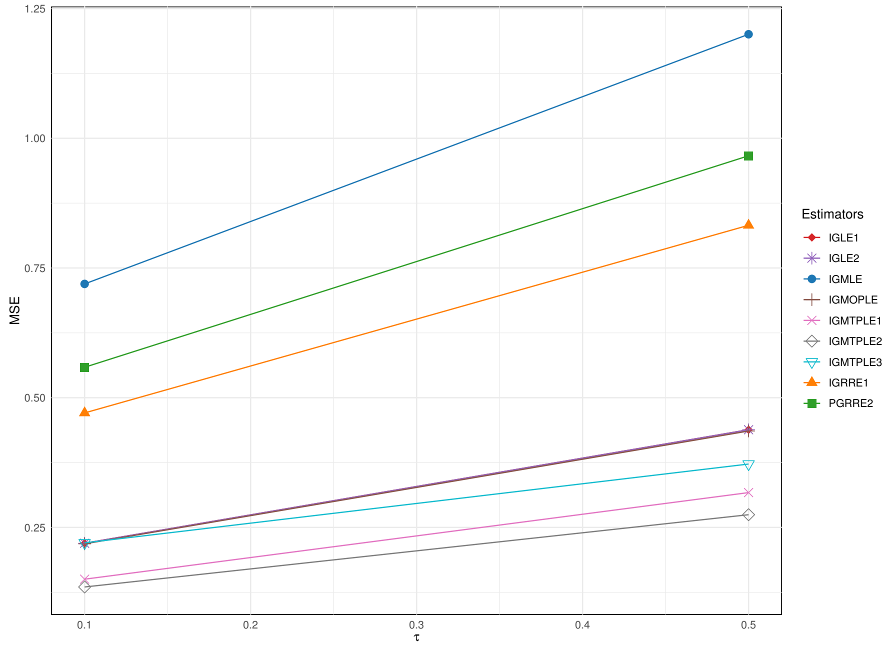

Figure 4. Simulated MSE values of biasing estimators with dispersion parameter.

5. Applications

In evaluating the performance of our proposed estimator for the IGRM in the presence of multicollinearity, we use two different biomedical datasets (see Figures 5 and 6). First, in the Breast Cancer Wisconsin dataset, which includes clinical characteristics of tumors, biological relationships predict multicollinearity, and second, in Anticancer QSAR data with molecular descriptors of azidothymidine derivatives, multicollinearity is caused by structural similarities.

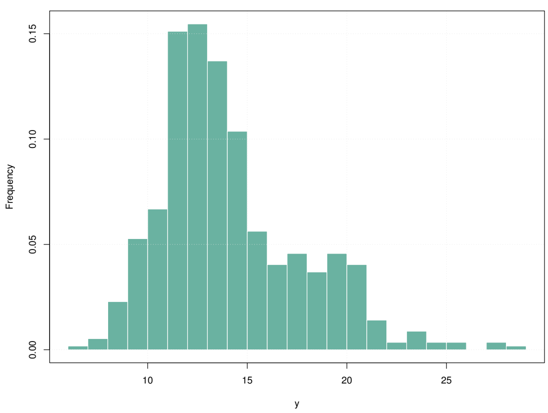

Figure 5. Histogram of the radius mean y in breast cancer data.

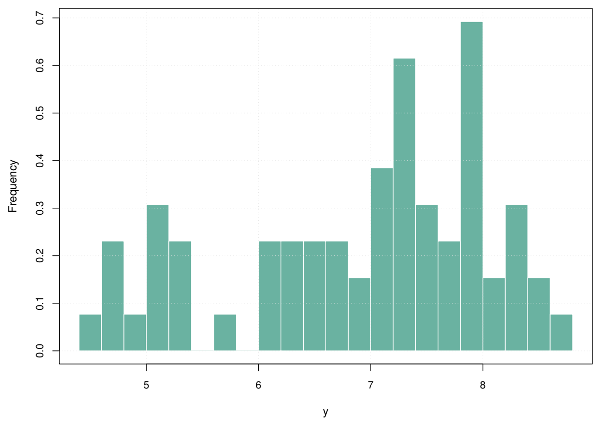

Figure 6. Histogram of the molecules y in anticancer compound data.

5.1. Breast cancer Wisconsin data

We initially analyzed the data for breast cancer to assess the efficacy of the proposed estimator. Breast cancer is still one of the leading health issues that the world is facing. It is the most frequently diagnosed type of cancer and is the second most deadly type of cancer among women. We utilized the Wisconsin Breast Cancer Diagnostic dataset, which contains data relating to the morphological features of tumors [33, 34]. The continuous response variable is the mean radius, which is the average distance of the tumor center to the perimeter, measured in pixels. This variable helps in the measurement and assessment of tumor progression and complements other methods of classification.

The dataset has 569 samples, with the mean tumor radius (y) as the average radius of the tumor from the center of the tumor to the perimeter. Ten predictors were considered: Diagnosis (x_1) is the identifying factor of a tumor as malignant (invasive and potentially metastatic) or benign (non-invasive, requiring less aggressive treatment). Image-based measurements capture tumor geometry and structure: Texture (x_2) captures the variation of gray-scale values; perimeter (x_3) captures the length of the tumor boundary; and area (x_4) is a measurement of the tumor region. Regarding shape, there is smoothness (x_5), which captures the variation of the radius locally, and (x_6), compactness, which is defined as Perimeter^2/Area − 1 (the closer to zero, the more circular the tumor is), and contoured indentations of concavity (x_7). Concave points (x_8) capture the number of indentations and symmetry (x_9), which captures the bilateral equalness, and fractal dimension (x_{10}), which is the measurement of the boundary's complexity.

We assessed multicollinearity using three diagnostic measures: a correlation matrix revealing strong pairwise associations among predictors as in Figure 7, a condition number (CN = 312,853.5), far exceeding the critical threshold (CN > 30), and variance inflation factors (VIFs) ranging from 1.31 to 80.08, with five predictors exhibiting VIF > 10. These results unanimously confirm severe multicollinearity, warranting remedial estimation approaches.

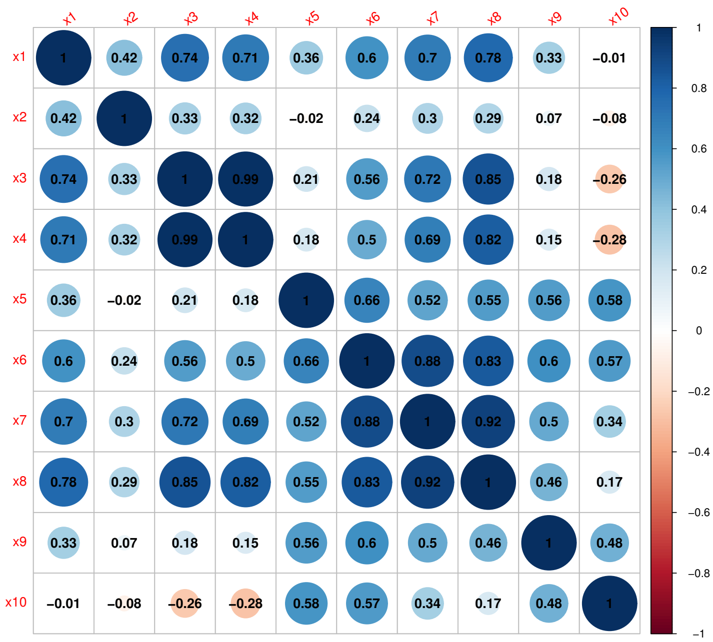

Figure 7. Correlation matrix between explanatory variables in the breast cancer data.

Table 7. Coefficients and MSE values for each estimator of the IGRM using the breast cancer data.

Estimator		β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	β_{10}	MSE
IGMLE	-	0.187851 (0.003439)	-0.000352 (0.000328)	0.000018 (0.000025)	-0.001895 (0.000034)	0.000078 (0.000002)	-0.063377 (0.011603)	0.054402 (0.006581)	-0.001932 (0.003852)	0.052550 (0.010926)	0.008572 (0.004596)	0.103466 (0.033584)	0.003012
IGRRE	$\hat{k}_1$	0.188027 (0.003351)	-0.000354 (0.000328)	0.000018 (0.000025)	-0.001896 (0.000033)	0.000078 (0.000002)	-0.062597 (0.011517)	0.054765 (0.006451)	-0.001818 (0.003842)	0.052099 (0.010864)	0.008572 (0.004591)	0.100053 (0.032215)	0.002829
	$\hat{k}_2$	0.187851 (0.003439)	-0.000352 (0.000328)	0.000018 (0.000025)	-0.001895 (0.000034)	0.000078 (0.000002)	-0.063377 (0.011603)	0.054402 (0.006581)	-0.001932 (0.003852)	0.052550 (0.010926)	0.008572 (0.004596)	0.103465 (0.033584)	0.003012
IGLE	$\hat{d}_1$	0.187856 (0.003436)	-0.000352 (0.000328)	0.000018 (0.000025)	-0.001895 (0.000034)	0.000078 (0.000002)	-0.063355 (0.011601)	0.054413 (0.006578)	-0.001929 (0.003851)	0.052538 (0.010924)	0.008572 (0.004596)	0.103369 (0.033546)	0.003006
	$\hat{d}_2$	0.187851 (0.003439)	-0.000352 (0.000328)	0.000018 (0.000025)	-0.001895 (0.000034)	0.000078 (0.000002)	-0.063377 (0.011603)	0.054402 (0.006581)	-0.001932 (0.003852)	0.052550 (0.010926)	0.008572 (0.004596)	0.103465 (0.033584)	0.003012
IGMOPLE	$\hat{d}_0$	0.188293 (0.003226)	-0.000355 (0.000328)	0.000018 (0.000025)	-0.001898 (0.000033)	0.000078 (0.000002)	-0.061494 (0.011397)	0.055307 (0.006265)	-0.001659 (0.003829)	0.051469 (0.010778)	0.008569 (0.004584)	0.095034 (0.030203)	0.002620
IGMTPLE	$\hat{k}_1, \hat{d}_0$	0.189252 (0.002797)	-0.000362 (0.000327)	0.000018 (0.000025)	-0.001903 (0.000031)	0.000078 (0.000002)	-0.057405 (0.010964)	0.057272 (0.005645)	-0.001068 (0.003781)	0.049124 (0.010457)	0.008561 (0.004560)	0.076730 (0.022863)	0.002465
	$\hat{k}_2, \hat{d}_0$	0.188293 (0.003226)	-0.000355 (0.000328)	0.000018 (0.000025)	-0.001898 (0.000033)	0.000078 (0.000002)	-0.061494 (0.011397)	0.055307 (0.006265)	-0.001659 (0.003829)	0.051469 (0.010778)	0.008569 (0.004584)	0.095034 (0.030203)	0.002620
	$\hat{k}_3, \hat{d}_0$	0.188391 (0.003180)	-0.000356 (0.000328)	0.000018 (0.000025)	-0.001898 (0.000033)	0.000078 (0.000002)	-0.061075 (0.011352)	0.055509 (0.006197)	-0.001599 (0.003824)	0.051229 (0.010745)	0.008568 (0.004582)	0.093160 (0.029451)	0.002560

Note: Values in parentheses are standard errors (SEs).

The application of estimators in the IGRM to breast cancer datasets demonstrates significant differences in predictive performance and the stability of coefficients, as seen in Table 7. The IGMTPLE, the proposed method, has the best predictive performance. It has the best predict MSE of 0.002465. This is a significant improvement to the IGMLE and IGRRE, which have MSEs of 0.003012 and 0.002829, respectively.

Coefficient estimates also show patterns across estimators. The estimates for the first five predictors (β_1 through β_5) are consistent across the models, meaning that these variables are less subject to multicollinearity. However, large differences show up in the other predictors. In particular, β_{10} (fractal dimension) differs from 0.0767 under the IGMTPLE and 0.1035 under the IGMLE. The differences show that multicollinearity has a comparatively large effect on certain predictors, and the shrinkage methods reduce the large estimates.

The standard errors also support the shrinkage effect. The IGMTPLE with tuning parameters $(\hat{k}_1, \hat{d}_0)$ has smaller standard error estimates for most coefficients than the IGMLE, especially for β_5 (0.01096 vs. 0.01160) and β_{10} (0.02286 vs. 0.03358). The reduction in variation shows that the biases in shrinkage are small and thus explains the improved MSE performance.

The presented results are in line with the previously discussed simulation results, reiterating the idea that the IGMTPLE demonstrates consistent performance in real-world applications where some predictors are more important than others and where predictors are correlated. The IGMTPLE demonstrates a practical means of improving the bias-variance tradeoff and, as such, is a viable option for use in complex biomedical problems where the number of dimensions remains high.

5.2. Anticancer compound data: QSAR model

The second dataset was used to evaluate the performance of our proposed method as a QSAR chemometrics application with 65 imidazo [35] pyridine derivatives, where the transformed biological activity (pIC50 $= -\log(\text{IC50})$) serves as the response variable and the 15 molecular descriptors as predictors.

In our QSAR analysis, we employed a dataset comprising 65 observations, with pIC50 (log IC50) serving as the target variable and 15 molecular descriptors as predictors for the anticancer compound dataset. The descriptors include molecular weight (x_1), which measures compound size; IC3 (x_2), reflecting neighborhood symmetry; and SpMaxA_D (x_3), representing the leading eigenvalue of the topological distance matrix. Several autocorrelation descriptors quantify the variability of specific properties along the molecular structure, weighted by volume or polarizability: ATS8v (x_4), MATS7v (x_5), MATS2s (x_6), and GATS4p (x_7). Burden matrix eigenvalue descriptors capture structural features, including SpMax8_Bh(p) (x_8) and SpMax3_Bh(s) (x_9). Electronic properties are represented by P_VSA_e_3 (x_{10}). Three-dimensional structural information is provided by TD808m (x_{11}), RDF100m (x_{12}), and the 3D-MoRSE signals Mor21v (x_{13}) and Mor21e (x_{14}), weighted by van der Waals volume and Sanderson electronegativity, respectively. Finally, HATS6v (x_{15}) describes leverage-weighted autocorrelation.

Multicollinearity diagnostics revealed severe issues: the correlation matrix showed strong pairwise associations as in Figure 8, the CN (43,983.53) far exceeded the critical threshold of 30, and five predictors exhibited VIF values above 10 (range: 1.97–75.57). These consistent findings demonstrate substantial multicollinearity that could compromise standard regression results, indicating the need for specialized estimation techniques to ensure reliable modeling.

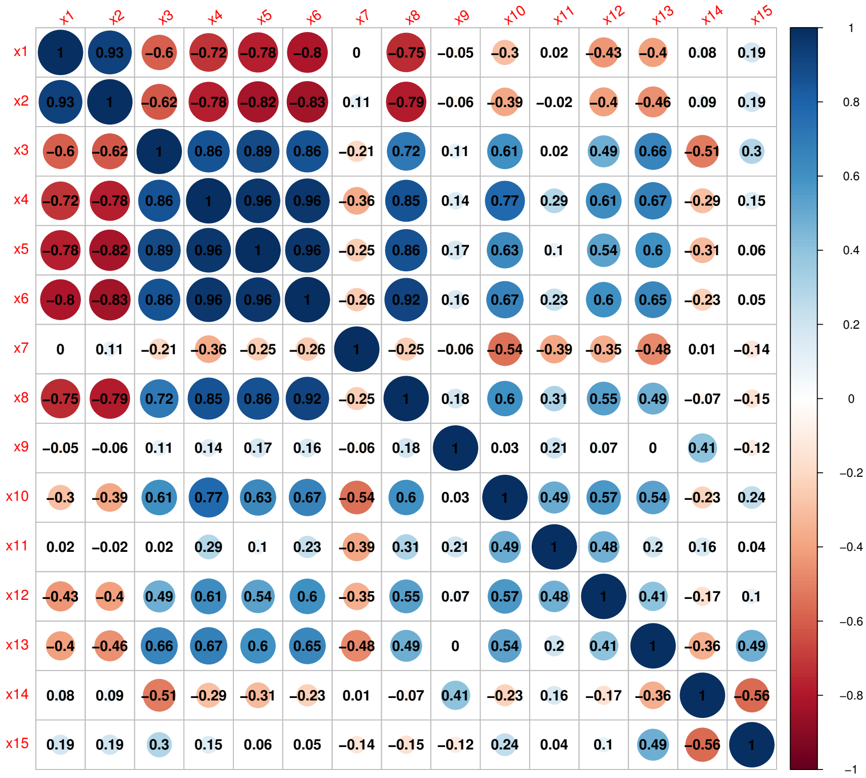

Figure 8. Correlation matrix between explanatory variables in the anticancer compound data.

Table 8 presents the coefficient estimates and prediction errors for all methods applied to this dataset. The proposed IGMTPLE estimator achieves the lowest prediction error (MSE = 0.0196) among all methods examined. This represents a clear improvement over the standard IGMLE (MSE = 0.0381) and the IGMOPLE (MSE = 0.0257).

The coefficient estimates reveal interesting patterns across molecular descriptors. For predictors with small effects, such as β_4 and β_{10}, all methods produce nearly identical estimates. However, several descriptors show large variation depending on which estimator is used. Most notably, β_{15} ranges from -0.0575 under the IGMLE to 0.0070 under the IGMTPLE, a complete change in direction. Similarly, β_8 varies across methods, while β_3 shows differences in magnitude. These patterns indicate that conventional estimators may be unstable for certain molecular features, potentially leading to unreliable conclusions in drug discovery research.

The strong performance of the IGMTPLE has several practical implications. First, more stable coefficient estimates reduce the chance of incorrectly identifying unimportant descriptors as significant. Second, the consistency between the IGMTPLE and the IGMOPLE increases confidence in the estimated effects for key molecular properties. Third, lower prediction error means more accurate activity forecasts for new compounds, which could help prioritize promising drug candidates earlier in development.

The IGMTPLE with tuning parameters $(\hat{k}_1, \hat{d}_0)$ performs particularly well, suggesting this configuration is well matched to the correlation structure of molecular descriptor data. The advantage is most visible for polarizability-related descriptors (β_5 through β_9) and steric parameters (β_{13} through β_{15}), where conventional methods show high variability while the IGMTPLE remains stable.

Table 8. Coefficients and MSE values for each estimator of the IGRM using the anticancer compound data.

Estimator		β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	β_{10}	β_{11}	β_{12}	β_{13}	β_{14}	β_{15}	MSE
IGMLE	-	0.326418	0.019514	-0.000313	-0.008390	-0.000332	0.011580	0.037631	0.065483	-0.072287	-0.010183	-0.000009	0.018313	-0.000025	0.029656	0.049327	-0.057558	0.038070
		(0.060862)	(0.021052)	(0.009500)	(0.014840)	(0.000128)	(0.010736)	(0.021462)	(0.031848)	(0.022344)	(0.004928)	(0.000111)	(0.048644)	(0.000349)	(0.023566)	(0.030550)	(0.133059)	
IGRRE	$\hat{k}_1$	0.313410	0.019698	-0.000791	-0.006670	-0.000333	0.011289	0.036523	0.066617	-0.069796	-0.010081	-0.000008	0.018805	-0.000026	0.028515	0.050914	-0.043819	0.030345
		(0.057995)	(0.020917)	(0.009393)	(0.014490)	(0.000125)	(0.010471)	(0.021210)	(0.031136)	(0.021898)	(0.004871)	(0.000111)	(0.047242)	(0.000348)	(0.022852)	(0.029618)	(0.111284)	
	$\hat{k}_2$	0.326398	0.019515	-0.000314	-0.008388	-0.000332	0.011579	0.037630	0.065484	-0.072284	-0.010183	-0.000009	0.018314	-0.000025	0.029654	0.049329	-0.057534	0.038055
		(0.060858)	(0.021052)	(0.009499)	(0.014839)	(0.000128)	(0.010736)	(0.021461)	(0.031847)	(0.022343)	(0.004928)	(0.000111)	(0.048642)	(0.000349)	(0.023565)	(0.030549)	(0.133022)	
IGLE	$\hat{d}_1$	0.325150	0.019534	-0.000362	-0.008227	-0.000332	0.011557	0.037524	0.065588	-0.072044	-0.010175	-0.000009	0.018372	-0.000025	0.029535	0.049497	-0.056050	0.037138
		(0.060578)	(0.021040)	(0.009488)	(0.014806)	(0.000128)	(0.010708)	(0.021438)	(0.031777)	(0.022299)	(0.004922)	(0.000111)	(0.048508)	(0.000349)	(0.023488)	(0.030450)	(0.130694)	
	$\hat{d}_2$	0.326407	0.019515	-0.000313	-0.008389	-0.000332	0.011579	0.037630	0.065484	-0.072285	-0.010183	-0.000009	0.018314	-0.000025	0.029655	0.049328	-0.057545	0.038061
		(0.060860)	(0.021052)	(0.009500)	(0.014840)	(0.000128)	(0.010736)	(0.021462)	(0.031847)	(0.022344)	(0.004928)	(0.000111)	(0.048642)	(0.000349)	(0.023565)	(0.030549)	(0.133039)	
IGMOPLE	$\hat{d}_0$	0.305411	0.019839	-0.001122	-0.005688	-0.000335	0.011200	0.035850	0.067218	-0.068250	-0.010048	-0.000008	0.019289	-0.000026	0.027661	0.052154	-0.032588	0.025722
		(0.056239)	(0.020843)	(0.009327)	(0.014285)	(0.000124)	(0.010306)	(0.021069)	(0.030715)	(0.021633)	(0.004836)	(0.000111)	(0.046427)	(0.000347)	(0.022386)	(0.029017)	(0.093896)	
IGMTPLE	$\hat{k}_1, \hat{d}_0$	0.272135	0.020352	-0.002403	-0.001408	-0.000338	0.010600	0.033029	0.069966	-0.061854	-0.009834	-0.000007	0.020835	-0.000027	0.024502	0.056632	0.006964	0.019588
		(0.049375)	(0.020519)	(0.009124)	(0.013454)	(0.000118)	(0.009763)	(0.020464)	(0.029101)	(0.020647)	(0.004719)	(0.000111)	(0.043051)	(0.000344)	(0.021078)	(0.027240)	(0.032145)	
	$\hat{k}_2, \hat{d}_0$	0.305411	0.019839	-0.001122	-0.005688	-0.000335	0.011200	0.035850	0.067218	-0.068250	-0.010048	-0.000008	0.019289	-0.000026	0.027661	0.052154	-0.032588	0.025722
		(0.056239)	(0.020843)	(0.009327)	(0.014285)	(0.000124)	(0.010306)	(0.021069)	(0.030715)	(0.021633)	(0.004836)	(0.000111)	(0.046427)	(0.000347)	(0.022386)	(0.029017)	(0.093896)	
	$\hat{k}_3, \hat{d}_0$	0.291315	0.020056	-0.001664	-0.003875	-0.000336	0.010946	0.034655	0.068382	-0.065540	-0.009957	-0.000007	0.019944	-0.000026	0.026323	0.054051	-0.015833	0.021115
		(0.053253)	(0.020705)	(0.009230)	(0.013926)	(0.000121)	(0.010054)	(0.020810)	(0.030002)	(0.021193)	(0.004782)	(0.000111)	(0.044975)	(0.000346)	(0.021742)	(0.028160)	(0.067654)	

Note: Values in parentheses are standard errors (SEs).

6. Conclusions

This paper presents a new modified Liu estimator for the IGRM to overcome inferential problems caused by multicollinearity in multivariate positive-skewed data. Multicollinearity causes the MLEs to have inflated variances and inaccurate coefficient estimates, thereby reducing the reliability of statistical inferences. The new estimator addresses these problems and provides more accurate and consistent estimates of the model parameters. The proposed estimator was the subject of a comparative analysis, in terms of MSE, against the MLE and other biased estimators. Using a Monte Carlo simulation, it was shown that the proposed estimator was superior in all scenarios considered. The simulation results were supported by two case studies based on real multivariate medical data, which demonstrated the usefulness of the estimator. Overall, the modified Liu estimator is reliable and provides a significant improvement over the other estimators for IGRM problems involving multicollinearity. It is important to note that the biased estimator has certain weaknesses, such as its sensitivity to the shrinkage parameters and its focus on multicollinearity without addressing other issues like outliers and high dimensionality. Future lines of inquiry may include refining the estimator to simultaneously address both multicollinearity and outliers, as well as developing cross-validation methodologies to facilitate optimal parameter selection, as highlighted in the works of Hammad et al. [36], Alshangiti et al. [37], and Kamel et al. [38].

Author contributions

All authors contributed equally to the writing and review of this manuscript. All authors have read and approved the final version of the manuscript for publication.

Use of Generative-AI tools declaration

The authors confirm that no artificial intelligence (AI) tools were used in the creation of this article.

Funding

This work was supported and funded by the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University (IMSIU) (grant number IMSIU-DDRSP2602).

Conflict of interest

The authors declare no conflicts of interest.

Data availability

The data supporting the findings of this study are included within the article.

References

1. B. Y. Lemeshko, S. B. Lemeshko, K. A. Akushkina, M. S. Nikulin, N. Saaidia, Inverse Gaussian model and its applications in reliability and survival analysis, In: V. Rykov, N. Balakrishnan, M. Nikulin, *Mathematical and statistical models and methods in reliability*, Statistics for Industry and Technology, Boston, MA: Birkhäuser, 2010, 433–453. https://doi.org/10.1007/978-0-8176-4971-5_33
2. G. Fang, R. Pan, Y. Wang, Inverse Gaussian processes with correlated random effects for multivariate degradation modeling, *Eur. J. Oper. Res.*, **300** (2022), 1177–1193. https://doi.org/10.1016/j.ejor.2021.10.049
3. M. Amin, M. A. Ullah, M. Qasim, Diagnostic techniques for the inverse Gaussian regression model, *Commun. Stat.-Theory Methods*, **51** (2022), 2552–2564. https://doi.org/10.1080/03610926.2020.1777308
4. M. N. Akram, M. Amin, B. M. G. Kibria, M. Arashi, A. F. Lukman, N. Afzal, A new improved Liu estimator for the QSAR model with inverse Gaussian response, *Commun. Stat.-Simul. Comput.*, **53** (2024), 1873–1888. https://doi.org/10.1080/03610918.2022.2059088
5. I. D. Nisa, B. W. Otok, Sutikno, Multivariate adaptive inverse Gaussian regression spline: parameter estimation and statistical hypothesis testing, *AIP Conf. Proc.*, **3201** (2024), 060011. https://doi.org/10.1063/5.0231185
6. E. K. Nisa, R. Maulina, The comparison of inverse Gaussian and Gamma regression: application on stunting data in Jepara, *J. Mat., Stat. Komputasi*, **21** (2024), 334–344. https://doi.org/10.20956/j.v21i1.36351
7. G. Lee, S. Lee, I. S. Na, Correlation between public housing location and city local characteristics: a comparison of gamma and inverse Gaussian models, *J. Asian Archit. Build. Eng.*, **25** (2026), 804–823. https://doi.org/10.1080/13467581.2025.2457388
8. M. Amin, T. Mahmood, S. Kinat, Memory type control charts with inverse-Gaussian response: an application to yarn manufacturing industry, *Trans. Inst. Meas. Control*, **43** (2021), 656–678. https://doi.org/10.1177/0142331220952965
9. T. Kyriazos, M. Poga, Dealing with multicollinearity in factor analysis: the problem, detections, and solutions, *Open J. Stat.*, **13** (2023), 404–424. https://doi.org/10.4236/ojs.2023.133020
10. T. Cheng, X. Wang, X. Liu, J. Wang, Uncertainty-based reliability analysis of cutters via improved Bayesian prior distribution, *Trans. Can. Soc. Mech. Eng.*, **49** (2025), 220–228. https://doi.org/10.1139/tcsme-2024-0102
11. Z. Y. Algamal, Performance of ridge estimator in inverse Gaussian regression model, *Commun. Stat.-Theory Methods*, **48** (2019), 3836–3849. https://doi.org/10.1080/03610926.2018.1481977
12. H. Sun, D. Fan, X. Liu, J. Liu, H. Ge, Composite probability distribution for fatigue life prediction of API X65 steel via Vickers hardness, *Int. J. Damage Mech.*, **35** (2026), 22–45. https://doi.org/10.1177/10567895251331310
13. A. E. Hoerl, R. W. Kennard, Ridge regression: biased estimation for nonorthogonal problems, *Technometrics*, **12** (1970), 55–67. https://doi.org/10.1080/00401706.1970.10488634
14. L. Kejian, A new class of blased estimate in linear regression, *Commun. Stat.-Theory Methods*, **22** (1993), 393–402. https://doi.org/10.1080/03610929308831027

15. B. Segerstedt, On ordinary ridge regression in generalized linear models, *Commun. Stat.-Theory Methods*, **21** (1992), 2227–2246. https://doi.org/10.1080/03610929208830909

16. A. F. Lukman, B. M. G. Kibria, K. Ayinde, S. L. Jegede, Modified one-parameter Liu estimator for the linear regression model, *Model. Simul. Eng.*, **2020** (2020), 1–17. https://doi.org/10.1155/2020/9574304

17. A. Araveeporn, Modified Liu parameters for scaling options of the multiple regression model with multicollinearity problem, *Mathematics*, **12** (2024), 3139. https://doi.org/10.3390/math12193139

18. K. Liu, Using Liu-type estimator to combat collinearity, *Commun. Stat.-Theory Methods*, **32** (2003), 1009–1020. https://doi.org/10.1081/sta-120019959

19. A. F. Lukman, K. Ayinde, S. Binuomote, O. A. Clement, Modified ridge-type estimator to combat multicollinearity: application to chemical data, *J. Chemom.*, **33** (2019), e3125. https://doi.org/10.1002/cem.3125

20. M. R. Özkale, S. Kaçiranlar, The restricted and unrestricted two-parameter estimators, *Commun. Stat.-Theory Methods*, **36** (2007), 2707–2725. https://doi.org/10.1080/03610920701386877

21. M. R. Abonazel, A new biased estimation class to combat the multicollinearity in regression models: modified two–parameter Liu estimator, *Comput. J. Math. Stat. Sci.*, **4** (2025), 316–347.

22. M. R. Abonazel, New modified two-parameter Liu estimator for the Conway–Maxwell Poisson regression model, *J. Stat. Comput. Simul.*, **93** (2023), 1976–1996. https://doi.org/10.1080/00949655.2023.2166046

23. M. M. Abdelwahab, M. R. Abonazel, A. T. Hammad, A. M. El-Masry, Modified two-parameter Liu estimator for addressing multicollinearity in the Poisson regression model, *Axioms,* **13** (2024), 46. https://doi.org/10.3390/axioms13010046

24. A. Fayomi, A. S. Hassan, E. M. Almetwally, Inference and quantile regression for the unit-exponentiated Lomax distribution, *Plos One*, **18** (2023), e0288635. https://doi.org/10.1371/journal.pone.0288635

25. A. T. Hammad, E. H. Hafez, U. Shahzad, E. Yıldırım, E. M. Almetwally, B. M. G. Kibria, New modified Liu estimators to handle the multicollinearity in the beta regression model: simulation and applications, *Mod. J. Stat.*, **1** (2025), 58–79. https://doi.org/10.64389/mjs.2025.01111

26. M. Amin, M. Qasim, S. Afzal, K. Naveed, New ridge estimators in the inverse Gaussian regression: Monte Carlo simulation and application to chemical data, *Commun. Stat.-Simul. Comput.*, **51** (2022), 6170–6187. https://doi.org/10.1080/03610918.2020.1797794

27. N. M. Hammood, D. M. Jabur, Z. Y. Algamal, A Liu estimator in inverse Gaussian regression model with application in chemometrics, *Math. Stat. Eng. Appl.*, **71** (2022), 248–266.

28. R. W. Farebrother, Further results on the mean square error of ridge regression, *J. R. Stat. Soc. Ser. B (Methodol.)*, **38** (1976), 248–250. https://doi.org/10.1111/j.2517-6161.1976.tb01588.x

29. G. Trenkler, H. Toutenburg, Mean squared error matrix comparisons between biased estimators—An overview of recent results, *Stat. Papers*, **31** (1990), 165–179. https://doi.org/10.1007/bf02924687

30. A. F. Lukman, Z. Y. Algamal, B. M. G. Kibria, K. Ayinde, The KL estimator for the inverse Gaussian regression model, *Concurr. Comput.: Pract. Exper.*, **33** (2021), e6222. https://doi.org/10.1002/cpe.6222

31. R. Alharbi, A. S. Alghamdi, A new almost unbiased estimator for beta regression model under multicollinearity, *AIMS Math.*, **11** (2026), 85–126. https://doi.org/10.3934/math.2026005

32. A. A. Alzahrani, Modified two-parameter ridge estimator for the Beta logistic model to mitigate multicollinearity, *AIMS Math.*, **11** (2026), 543–557. https://doi.org/10.3934/math.2026023

33. M. M. Abo El Nasr, A. A. Abdelmegaly, D. A. Abdo, Performance evaluation of different regression models: application in a breast cancer patient data, *Sci. Rep.*, **14** (2024), 12986. https://doi.org/10.1038/s41598-024-62627-6

34. V. Jaiswal, P. Suman, D. Bisen, An improved ensembling techniques for prediction of breast cancer tissues, *Multimed. Tools Appl.*, **83** (2024), 31975–32000. https://doi.org/10.1007/s11042-023-16949-8

35. Z. Y. Algamal, M. H. Lee, A novel molecular descriptor selection method in QSAR classification model based on weighted penalized logistic regression, *J. Chemom.*, **31** (2017), e2915. https://doi.org/10.1002/cem.2915

36. A. T. Hammad, I. Elbatal, E. M. Almetwally, M. M. Abd El-Raouf, M. A. El-Qurashi, A. M. Gemeay, A novel robust estimator for addressing multicollinearity and outliers in Beta regression: simulation and application, *AIMS Math.*, **10** (2025), 21549–21580. https://doi.org/10.3934/math.2025958

37. A. M. Alshangiti, A. M. Gemeay, M. E. Bakr, A. M. Yousuf, M. A. El-Qurashi, O. S. Balogun, et al., New robust estimator for handling outliers and multicollinearity in gamma regression model with application to breast cancer data, *Sci. Rep.*, **15** (2025), 38436. https://doi.org/10.1038/s41598-025-25231-w

38. A. Kamel, M. Abonazel, A. Youssef, A novel robust M-estimator for the random-coefficients regression model: simulation and its application in energy management systems, *Comput. J. Math. Stat. Sci.*, **4** (2025), 738–767.

AIMS Press