STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **14**, September 2025, pp 1566–1583.
Published online in International Academic Press (www.IAPress.org)

Penalized estimators for modified Log-Bilal regression: simulations and applications

Tarek Omara [1,2]

[1] *Department of Statistics, Mathematics and Insurance,Faculty of Commerce, Kafrelsheikh University,Egypt*
[2] *Department of Mathematics, Faculty of Sciences - Islamic University of Madinah, Madinah, KSA*

Abstract The Log-Bilal regression is survival regression model accounts for unique features of lifetime data. In this study, we modify the Log-Bilal distribution to enhance it flexibility, resulting in a model that exhibits an increasing, non-constant failure rate over time. To address the multicollinearity for the modified Log-Bilal regression, we introduce two penalized estimators: Ridge modified Log-Bilal (Ridge_MBE) and Liu_type modified Log-Bilal (liu_MBE) estimators. The properties for the suggested estimators are discussed and the superiority for the estimators were checked. The Liu_type estimator demonstrates superiority over the other estimators. A simulation study is conducted across various factors, which reveals that the Liu_type estimator outperforms the others in many cases. The proposed estimators were applied to real lifetime data from mechanical pumps which it gives the results confirming the results of the simulation study.

Keywords Multicollinearity, Log-Bilal regression, Modified Log-Bilal regression, Ridge estimator, Liu_type estimator

AMS 2020 subject classifications 62J07

DOI: 10.19139/soic-2310-5070-2513

1. Introduction

The lifetime data which is known as survival or failure time data, refers to the observed times until a specific event occurs. This event could represent the failure of a mechanical component, the death of a patient, or any other outcome that marks the end of a lifetime. Analyzing such data is critical across various fields, including medical research, engineering, and reliability studies. Survival regression, a form of regression analysis for lifetime data, models the relationship between covariates and the time until an event occurs. Unlike traditional regression methods, survival regression accounts for unique features of lifetime data, such as censoring and the non-normal distribution of survival times. The well-known models in survival regression include the Weibull, Tobin, log-logistic, generalized gamma and Gompertz regression models, introduced by [1],[2],[3], [4] and [5]. A more recent advancement in survival regression is the log-Bilal regression model, introduced by [6]. This model extends the log-Bilal distribution with the $T = exp(-X)$ link function. A key feature of the log-Bilal distribution is that its statistical functions have explicit, closed-form expressions, avoiding the need for special mathematical functions. The log-Bilal distribution was developed to enhance flexibility in modeling continuous, positive-valued data, particularly in cases with skewed or complex tail behavior. This distribution origins lie in generalized distributions that are built upon traditional exponential-type models.
The log-Bilal regression model is well-suited for survival analysis, particularly when the dependent variable is time-based and right-censored, meaning the event of interest may not have occurred for some subjects during the study period. This model offers significant flexibility in modeling the distribution of lifetimes, especially when

*Correspondence to: Tarek Omara (Email: drtarek.em@gmail.com)

ISSN 2310-5070 (online) ISSN 2311-004X (print)

using specific link functions or penalization techniques to address challenges like censoring, heterogeneity, or complex relationships between covariates and outcomes.

In contrast to traditional models like the Cox proportional hazards model or parametric models, the log-Bilal regression model accommodates different assumptions about hazard or survival functions. It is particularly useful when the hazard rate exhibits non-monotonic behaviors such as increasing, decreasing, or forming a bathtub-shaped curve where standard models like Weibull or exponential models may fall. This unique flexibility makes the log-Bilal regression model a powerful alternative for analyzing lifetime data with complex hazard dynamics.

The parameter estimation in log-Bilal regression can face many challenges when multicollinearity exists among predictor variables. In such cases, regularization techniques like penalized regression can provide more stable and accurate estimates. The penalization estimators, including ridge and Liu type estimators, have proven effective in dealing with issues of multicollinearity by introducing penalties on the regression coefficients. There are many studies that induced penalized survival regression. In these way, [7]introduced ridge and lasso estimators as a penalized likelihood methods for the Cox proportional hazards model, improving the model's stability by shrinking the regression coefficients. [8] study applied Lasso to the Cox proportional hazards model, introducing variable selection techniques in survival analysis. [9] extends the use of elastic net to the Cox model, providing a more flexible approach in survival regression by combining the strengths of Ridge and Lasso. [10] address multicollinearity by using Liu estimator for left censoring Tobit regression. [11]) introduced almost unbiased liu_type estimator for Tobit regression.

In this study, we first modified the Log-Bilal distribution. This modification of the Log-Bilal distribution offers significant advantages when applied to modeling failure rates. The modified approach is highly flexible, exhibiting an increasing, non-constant failure rate over time, making it particularly suitable for systems subject to wear and tear, such as mechanical components and pumps, where failure rates increase with age. In addition, it can effectively model multiple competing failure modes and varying hazard rates. In contrast, the original Log-Bilal regression is characterized by a constant or monotonic failure rate, which limits its flexibility. It is more appropriate for systems with predictable failure patterns and a single failure mode, where failure behavior remains steady and consistent. Overall, this modified version of the Log-Bilal distribution is better equipped to capture the failure dynamics associated with aging and degradation, whereas the original Log-Bilal distribution is more effective for systems with stationary or simpler failure patterns. In addition to its practical benefits, the modified Log-Bilal distribution also offers important improvements from a statistical modeling point of view. One key motivation for the modification is to improve estimation stability especially in situations where the data are complex or the predictors are highly correlated. The original model can be restrictive in such cases, leading to poor convergence or unstable results when used with penalized estimation methods. Our modified version introduces more flexibility in the baseline hazard, which helps the model respond more smoothly to changes in tuning parameters. This is particularly useful when applying Ridge or Liu_type penalties. It also retains closed form expressions for key functions, making it easier to work with in simulations and computational methods. Overall, the modification strengthens both the practical use and the statistical reliability of the model, allowing it to perform well across a wider range of real-world problems. For address the multicollinearity, we introduce two penalized estimators, Ridge and Liu_type estimators for modification of the Log-Bilal regression.

The study is organized as follows. In Sec. (2), we introduce the methodology. In Sec. (3), we illustrate the way for selecting the parameters for Ridge and Liu estimators. In Sec. (4), we make the simulation study and in Sec. (5), we study the sensitivity to model misspecification. The empirical data is analysis in Sec. (6). Finally, in section (7), the conclusion was provided.

2. Methodology:

This section presents the modified Log-Bilal regression model. We also derive the maximum likelihood estimator (MLE) along with Ridge and Liu_type estimators and analyze their properties. Additionally, we compare the mean squared errors (MSE) of the estimators.

2.1. The modified Log-Bilal regression model:

The logarithmic bi-lal distribution is defined by its probability density function (PDF), which exhibits a range of hazard rate shapes, making it well-suited for modeling diverse lifetime data. Its functional form combines exponential decay with a logistic-type shape, allowing for more flexible modeling of survival or failure time data. In survival analysis, where understanding time-to-event data is key, the flexibility of the Log-Bilal distribution is a critical advantage. The PDF of the Log-Bilal distribution, parameterized by θ, can be written as: $f(y;\theta) = \frac{6}{\theta}\exp(-\frac{2y}{\theta})\left(1-\exp(-\frac{y}{\theta}\right)$, where $y > 0$ represents the lifetime variable, and $\theta > 0$ is the scale parameter.
To capture additional features of the Log-Bilal distribution, we propose a modification of Log-Bilal distribution in the form

$$f(y;\theta) = \frac{6}{\theta}\exp(-2y\theta)\,(1-\exp(-y\theta) \tag{1}$$

where $y > 0$ represents the lifetime variable, and $\theta > 0$ is the scale parameter.
The factor $\exp(-2y\theta)$ is typical of distributions where the failure probability decays over time. However, when combined with the term1 $(1-\exp(-y\theta)$, this modification introduces a dynamic interaction that causes the failure rate to increase over time. The hazard function $h(y)$ for this distribution, which is derived from the ratio of the PDF to the survival function, increases as y increases. This behavior is due to the cumulative failure process introduced by the modified PDF, which accelerates as time progresses. In simpler terms, as the system ages, the likelihood of failure rises, a characteristic of wear-and-tear processes. The systems that degrade, wear out, or experience fatigue over time such as mechanical systems, electronics, or biological organisms often exhibit an increasing failure rate. In these systems, the longer the system operates, the more likely it is to fail due to cumulative damage, stress, or degradation. The modification to the Log-Bilal distribution reflects this type of process, making it a suitable model for such systems.
The mean and variance of Eq. 1 are in the forms:

$$\mu = \frac{2}{\theta^2} + \frac{1}{2\theta^3} \tag{2}$$

$$\sigma^2 = \left(\tfrac{3}{2\theta^2} + \tfrac{1}{\theta^3}\right) - \left(\tfrac{2}{\theta^2} + \tfrac{1}{2\theta^3}\right)^2$$

Let $Y = (y_1, y_2, \cdots, y_n)^T$ denote the response observations, which have the probability density function (PDF) given in Eq. 1. The model can by use the link function as:

$$g(\mu_i) = \frac{1}{\mu_i}, \qquad i = 1, 2, ..., p \tag{3}$$

So $\theta_i = \frac{1}{X_i\beta} = g^{-1}(X_i\beta)$, where $X_i = (x_{i1}, x_{i2}, \cdots, x_{in})^T$ for $i = 1, 2, \cdots, p$ represents the vector of independent variables and $\beta = [\beta_1, \beta_2, \cdots, \beta_p]$ denotes the corresponding coefficients. In addition, $g(.)$ is a continuous and twice differentiable function that maps the interval (0,1) into $\mathbb{R}$.
The log-likelihood function for PDF in Eq. 1 can be written as:

$$\ell(\boldsymbol{\beta}, \boldsymbol{\theta}) = \sum_{i=1}^{n} \left(\log 6 - \log\theta_i - 2y_i\theta_i + \log\left(1-\exp(-y_i\theta_i)\right)\right)$$

$$= \sum_{i=1}^{n}\left(\log 6 - \log\left(\frac{1}{\mathbf{X}_i^T\boldsymbol{\beta}}\right) - \frac{2y_i}{\mathbf{X}_i^T\boldsymbol{\beta}} + \log\left(1-\exp\left(-\frac{y_i}{\mathbf{X}_i^T\boldsymbol{\beta}}\right)\right)\right) \tag{4}$$

Then the score function $S(\boldsymbol{\beta}, \boldsymbol{\theta})$ can be found as:

$$S(\boldsymbol{\beta}, \boldsymbol{\theta}) = \frac{\partial \log L(\boldsymbol{\beta}, \boldsymbol{\theta})}{\partial \boldsymbol{\beta}} = \sum_{i=1}^{n} \left(\frac{-X_i}{X_i^2 \beta^2} \right) \left(-\frac{1}{\theta_i} - 2y_i + \frac{y_i \exp(-y_i\theta_i)}{1 - \exp(-y_i\theta_i)} \right) \tag{5}$$

Since $\theta_i = \frac{1}{X_i \beta}$:

$$S(\beta) = \sum_{i=1}^{n} \left(\frac{1}{\beta} + \frac{2y_i}{X_i^2 \beta^2} - \frac{y_i \exp\left(-\frac{y_i}{X_i\beta}\right)}{X_i \beta^2 \left(1 - \exp\left(-\frac{y_i}{X_i\beta}\right)\right)} \right)$$

The Fisher information matrix $I(\beta, \theta)$ can be found as:

$$I(\beta, \theta) = -\mathbb{E}\left[\frac{\partial^2 \log L(\beta, \theta)}{\partial \beta^2}\right] = \sum_{i=1}^{n} \left(\frac{X_i^2}{X_i^4 \beta^4} \mathbb{E}\left[\frac{1}{\theta_i^2} + \frac{2y_i \exp(-y_i\theta_i)}{(1 - \exp(-y_i\theta_i))^2}\right] \right)$$

Since θ_i is a parameter related to β, this term is constant with respect to the random variable y_i. Therefore:

$$I(\beta, \theta) = \sum_{i=1}^{n} \left(\frac{X_i^2}{X_i^4 \beta^4} \left[\frac{1}{\theta_i^2} + \mathbb{E}\left(\frac{2y_i \exp(-y_i\theta_i)}{(1 - \exp(-y_i\theta_i))^2}\right)\right] \right)$$

The expectation in the second term is more complex, since it involves the random variable y_i. It can be written as:

$$\mathbb{E}\left(\frac{2y_i \exp(-y_i\theta_i)}{(1 - \exp(-y_i\theta_i))^2}\right) = \int_0^\infty \frac{2y \exp(-y\theta_i)}{(1 - \exp(-y\theta_i))^2} \left(\frac{6}{\theta_i} \exp\left(-\frac{2y}{\theta_i}\right) (1 - \exp(-y/\theta_i)) \right) dy \tag{6}$$

$$= \int_0^\infty \frac{12y \exp(-3y\theta_i)}{\theta_i(1 - \exp(-y\theta_i))} \, dy \tag{7}$$

To solve the integral in Eq.7 we used the following approximations:

$$\frac{1}{1 - \exp(-y\theta_i)} \approx \frac{1}{y\theta_i - \frac{(y\theta_i)^2}{2} + \cdots}$$

We substitute the series approximation into the integral:

$$\int_0^\infty \frac{12y \exp(-3y\theta_i)}{\theta_i \left(y\theta_i - \frac{(y\theta_i)^2}{2} + \cdots\right)} \, dy$$

For a first-order approximation, we can consider only the leading term $\frac{1}{y_i\theta_i}$ for simplicity:

$$\mathbb{E}\left(\frac{2y_i \exp(-y_i\theta_i)}{(1 - \exp(-y_i\theta_i))^2}\right) \approx \int_0^\infty \frac{12y \exp(-3y\theta_i)}{\theta_i^2 y} \, dy = \frac{12}{\theta_i^2} \int_0^\infty \exp(-3y\theta_i) \, dy = \frac{12}{\theta_i^2} \cdot \frac{1}{3\theta_i} = \frac{4}{\theta_i^3}$$

Combining this result with the earlier expression for the Fisher Information Matrix:

$$\mathbb{E}\left(\frac{2y_i \exp(-y_i\theta_i)}{(1 - \exp(-y_i\theta_i))^2}\right) \approx \frac{4}{\theta_i^3}$$

Then the approximation for the Fisher Information Matrix is:

$$I(\beta, \theta) \approx \sum_{i=1}^{n} \left(\frac{X_i^2}{X_i^4 \beta^4} \left[\frac{1}{\theta_i^2} + \frac{4}{\theta_i^3}\right] \right)$$

Since $\theta_i = \frac{1}{X_i \beta}$, then:

$$I(\beta) \approx \sum_{i=1}^{n} X_i^2 \left(\frac{1}{X_i^2 \beta^2} + \frac{4}{X_i \beta} \right) \tag{8}$$

Given that Eq. 4 is non-linear, the unknown parameter β can be estimated by solving it iteratively using the **Fisher Scoring Method**. At each iteration r, we update β as:

$$\beta^{(r+1)} = \beta^{(r)} + \left[I\left(\beta^{(r)}\right) \right]^{-1} S\left(\beta^{(r)}\right) \tag{9}$$

In the context of regression models, the Fisher scoring method can be expressed in terms of a weight matrix W, by using the **Iterative Weighted Least Squares (IWLS)** algorithm. The IWLS algorithm estimates the parameters β iteratively, adjusting the weights in each iteration based on the current estimate of β:

$$\beta^{(r+1)} = \beta^{(r)} + \left(X^T W^{(r)} X \right)^{-1} X^T W^{(r)} z^{(r)} \tag{10}$$

where:

- r is the iteration number,
- $\beta^{(r)}$ is the current estimate of β,
- $W^{(r)} = \text{diag}\left(\frac{1}{X_i^2 \beta^2} + \frac{4}{X_i \beta} \right)$ is the diagonal weight matrix,
- $z_i^{(r)} = y_i - g^{-1}(X_i^T \beta)$ is the adjusted dependent variable at iteration r.

Then we can express the final **MLE Modified Log-Bilal Estimator (MLE_MBE)** for the regression model in the simplified form:

$$\widehat{\beta}_{\text{MLE\_MBE}} = \left(X^T \widehat{W} X \right)^{-1} X^T \widehat{W} \widehat{z} \tag{11}$$

As the sample size n increases, the approaches a normal distribution with mean β and covariance matrix $\left(X^T \widehat{W} X \right)^{-1}$. The mean squared error MSE for $\widehat{\beta}_{\text{MLE\_MBE}}$ can be expressed as:

$$MSE\left(\widehat{\beta}_{\text{MLE\_MBE}} \right) = \hat{\theta} D^{-1} = \hat{\theta} \sum_{j=1}^{p} \frac{1}{\lambda_j} \tag{12}$$

where $D = X^T W X$ and λ_j is j^{th} eigenvalue of $X^T W X$.
Let $Q^T X^T W X Q = diag\left(\lambda_1, \lambda_2, ..., \lambda_p\right),$ where $\lambda_1 \geq \lambda_2 \geq ...\lambda_p \geq 0$ are eigenvalue, $X^T W X$ and Q is $p \times p$ matrix has a eigenvectors of $X^T W X$. then $\alpha = Q\beta$

2.2. Ridge and Liu-type Estimators for Modified Log-Bilal Regression Model

The Ridge estimator is one of the important penalized estimators. For the ridge estimator we add more information to matrix matrix $X^T W X$ to solve the ill condition which it is a result for multicoloirity. We obtain (Ridge_MBE) and (Liu_type) estimators by adding the penalty term $k\beta^T \beta$ to the log-likelihood function in Eq.4 as:

$$\ell(\boldsymbol{\beta}, k, d) = \sum_{i=1}^{n} \left(\log 6 - \log\left(\frac{1}{\mathbf{X}_i^T \boldsymbol{\beta}} \right) - \frac{2y_i}{\mathbf{X}_i^T \boldsymbol{\beta}} + \log\left(1 - \exp\left(-\frac{y_i}{\mathbf{X}_i^T \boldsymbol{\beta}} \right) \right) \right) - k\boldsymbol{\beta}^T \boldsymbol{\beta}$$

Where $k \geq 0$ Then, we got $S(\boldsymbol{\beta}, \boldsymbol{k})$ as:

$$S(\boldsymbol{\beta}, k) = \sum_{i=1}^{n} \left(\frac{1}{\boldsymbol{\beta}} + \frac{2y_i}{(\mathbf{X}_i^T \boldsymbol{\beta})^2} - \frac{y_i \exp\left(-\frac{y_i}{\mathbf{X}_i^T \boldsymbol{\beta}} \right)}{(\mathbf{X}_i^T \boldsymbol{\beta})^2 \left(1 - \exp\left(-\frac{y_i}{\mathbf{X}_i^T \boldsymbol{\beta}} \right) \right)} \right) - 2k\boldsymbol{\beta}$$

By use the iterative weighted least squares (IWLS) algorithm, the (Ridg_MBE) estimator is given by:

$$\hat{\boldsymbol{\beta}}_{\text{Ridge\_MBE}} = \left(X^T\hat{W}X + kI\right)^{-1} X^T\hat{W}\hat{z}, \quad k > 0 \tag{13}$$

It can also be expressed as:

$$\hat{\boldsymbol{\beta}}_{\text{Ridge\_MBE}} = \left(X^T\hat{W}X + kI\right)^{-1} X^T\hat{W}X\hat{\boldsymbol{\beta}}_{\text{MLE\_MBE}} = H_R\hat{\boldsymbol{\beta}}_{\text{MLE\_MBE}} \tag{14}$$

where $H_R = \left(X^T\hat{W}X + kI\right)^{-1} X^T\hat{W}X$.

The bias and covariance of Ridge_MBE can be discussed as:

$$\mathbb{E}[\hat{\boldsymbol{\beta}}_{\text{Ridge\_MBE}}] = H_R\boldsymbol{\beta}$$

$$\text{Cov}[\hat{\boldsymbol{\beta}}_{\text{Ridge\_MBE}}] = \hat{\theta}H_RD^{-1}H_R^T$$

$$\text{Bias}[\hat{\boldsymbol{\beta}}_{\text{Ridge\_MBE}}] = (H_R - I)\boldsymbol{\beta} = -k(D + kI)^{-1}\boldsymbol{\beta}$$

Thus, the mean squared error (MSE) is:

$$\text{MSE}[\hat{\boldsymbol{\beta}}_{\text{Ridge\_MBE}}] = \hat{\theta}H_RD^{-1}H_R^T + k^2(D + kI)^{-1}\boldsymbol{\beta}\boldsymbol{\beta}^T(D + kI)^{-1} \tag{15}$$

The Liu-type estimator is a penalized estimator that relies on two tuning parameters, with the first parameter being restricted by the second. This estimator is derived by incorporating the penalty term $(k^{1/2}\boldsymbol{\beta} - k^{-1/2}d\hat{\boldsymbol{\beta}}_{\text{MLMBE}})(k^{1/2}\boldsymbol{\beta} - k^{-1/2}d\hat{\boldsymbol{\beta}}_{\text{MLMBE}})^T$ into the log-likelihood function Eq. 4, then

$$\ell(\boldsymbol{\beta}, k, d) = \sum_{i=1}^{n}\left(\log 6 - \log\left(\frac{1}{\mathbf{X}_i^T\boldsymbol{\beta}}\right) - \frac{2y_i}{\mathbf{X}_i^T\boldsymbol{\beta}} + \log\left(1 - \exp\left(-\frac{y_i}{\mathbf{X}_i^T\boldsymbol{\beta}}\right)\right)\right)$$
$$- \left(k^{1/2}\boldsymbol{\beta} - k^{-1/2}d\hat{\boldsymbol{\beta}}_{\text{MLE\_MBE}}\right)^T\left(k^{1/2}\boldsymbol{\beta} - k^{-1/2}d\hat{\boldsymbol{\beta}}_{\text{MLE\_MBE}}\right)$$

Using the IWLS algorithm, the Liu_type estimator becomes:

$$\hat{\boldsymbol{\beta}}_{\text{Liu\_MBE}} = \left(X^T\hat{W}X + kI\right)^{-1}\left(X^T\hat{W}X + dI\right)\hat{\boldsymbol{\beta}}_{\text{MLE\_MBE}} = H_{\text{Liu}}\hat{\boldsymbol{\beta}}_{\text{MLE\_MBE}} \tag{16}$$

where $H_{\text{Liu}} = \left(X^T\hat{W}X + kI\right)^{-1}\left(X^T\hat{W}X + dI\right)$.

The properties are:

$$\mathbb{E}[\hat{\boldsymbol{\beta}}_{\text{Liu\_MBE}}] = H_{\text{Liu}}\boldsymbol{\beta}$$

$$\text{Cov}[\hat{\boldsymbol{\beta}}_{\text{Liu\_MBE}}] = \hat{\theta}H_{\text{Liu}}D^{-1}H_{\text{Liu}}^T$$

$$\text{Bias}[\hat{\boldsymbol{\beta}}_{\text{Liu\_MBE}}] = (H\text{Liu} - I)\boldsymbol{\beta} = -(d - k)(D + kI)^{-1}\boldsymbol{\beta}$$

Thus, the mean Squared error (MSE) is:

$$\text{MSE}[\hat{\boldsymbol{\beta}}_{\text{Liu\_MBE}}] = \hat{\theta}H_{\text{Liu}}D^{-1}H_{\text{Liu}}^T + (d - k)^2(D + kI)^{-1}\boldsymbol{\beta}\boldsymbol{\beta}^T(D + kI)^{-1} \tag{17}$$

2.3. Comparison Between the Penalized Estimators for Modified Log-Bilal Regression

In this section, we assess the superiority of the Liu-type estimator (Liu_MBE) over both the maximum likelihood estimator (MLE_MBE) and the Ridge-type estimator (Ridge_MBE) using the mean square error (MSE) as the performance criterion.

At the first we introduce the following lemmas:

Lemma 1 [12] If A and B are matrices, and A is positive semi-definite, then for any matrix B, the matrix $A - BA^{-1}B^T$ is also positive semi-definite.

Lemma 2 [13] Let A be a positive matrix, $\mathbf{b}$ a non-zero vector, and ξ a positive scalar. Then, the matrix $A - \mathbf{b}\mathbf{b}^T$ is positive definite if and only if

$$\mathbf{b}^TA^{-1}\mathbf{b} < \xi.$$

Superiority of Ridge_MBE over MLE_MBE

Let the difference in MSE between the MLE_MBE and Ridge_MBE estimators be denoted by

$$\Delta_1 = \text{MSE}[\hat{\boldsymbol{\beta}}_{\text{MLE\_MBE}}] - \text{MSE}[\hat{\boldsymbol{\beta}}_{\text{Ridge\_MBE}}].$$

Then:

$$\begin{aligned}
\Delta_1 &= \hat{\theta}D^{-1} - \hat{\theta}H_R D^{-1} H_R^T - k^2(D+kI)^{-1}\boldsymbol{\beta}\boldsymbol{\beta}^T \left[(D+kI)^{-1}\right]^T \\
&= \hat{\theta}(X^T\hat{W}X)^{-1} - \hat{\theta}(X^T\hat{W}X+kI)^{-1}X^T\hat{W}X\left[(X^T\hat{W}X+kI)^{-1}\right]^T \\
&\quad - k^2(X^T\hat{W}X+kI)^{-1}\boldsymbol{\beta}\boldsymbol{\beta}^T\left[(X^T\hat{W}X+kI)^{-1}\right]^T \\
&= (X^T\hat{W}X+kI)^{-1}\Big[\hat{\theta}(X^T\hat{W}X+kI)(X^T\hat{W}X)^{-1}(X^T\hat{W}X+kI) \\
&\quad - \hat{\theta}(X^T\hat{W}X) - k^2\boldsymbol{\beta}\boldsymbol{\beta}^T\Big](X^T\hat{W}X+kI)^{-1} \\
&= (X^T\hat{W}X+kI)^{-1}\left[\hat{\theta}2k + \hat{\theta}k^2(X^T\hat{W}X)^{-1} - k^2\boldsymbol{\beta}\boldsymbol{\beta}^T\right](X^T\hat{W}X+kI)^{-1} \\
&= (D+kI)^{-1}\left[\hat{\theta}(2k+k^2D^{-1}) - k^2\boldsymbol{\beta}\boldsymbol{\beta}^T\right](D+kI)^{-1}
\end{aligned}$$

Thus, $\Delta_1 \geq 0$ if and only if:

$$\hat{\theta}\left(2k^{-1} + D^{-1}\right) \geq \boldsymbol{\beta}\boldsymbol{\beta}^T.$$

From **Lemma 2** , this inequality holds if and only if:

$$\boldsymbol{\beta}^T\left(2k^{-1} + D^{-1}\right)^{-1}\boldsymbol{\beta} \geq \hat{\theta}.$$

Superiority of Liu_MBE over MLE_MBE

Now consider the difference in MSE between the MLE_MBE and Liu_MBE estimators:

$$\Delta_2 = \text{MSE}[\hat{\boldsymbol{\beta}}_{\text{MLE\_MBE}}] - \text{MSE}[\hat{\boldsymbol{\beta}}_{\text{Liu\_MBE}}] = \hat{\theta}D^{-1} - \hat{\theta}H_{\text{Liu}}D^{-1}H_{\text{Liu}}^T - (d-k)^2(D+kI)^{-1}\boldsymbol{\beta}\boldsymbol{\beta}^T\left[(D+kI)^{-1}\right]^T.$$

Now, we apply the matrix identity from **Lemma 1** to our situation. Let $A = D^{-1}$ is positive semi-definite (since D is positive definite) and $B = H_{\text{Liu}}$, which implies:

$$D^{-1} - H_{\text{Liu}}D^{-1}H_{\text{Liu}}^T \geq 0.$$

Then, since $\hat{\theta} > 0$, we get:

$$\hat{\theta}\left(D^{-1} - H_{\text{Liu}}D^{-1}H_{\text{Liu}}^T\right) \geq 0.$$

We now assess the condition under which:

$$\left\|\hat{\theta}\left(D^{-1} - H_{\text{Liu}}D^{-1}H_{\text{Liu}}^T\right)\right\| \geq \left\|(d-k)^2(D+kI)^{-1}\boldsymbol{\beta}\boldsymbol{\beta}^T\left[(D+kI)^{-1}\right]^T\right\|.$$

To proceed with the inequality, we'll consider the trace norm approximation of the matrix. The Frobenius norm (which can be used here for convenience) can be written as: $\|C\|_F^2 = \text{tr}(C^TC)$, we write:

$$\text{tr}\left[\left(\hat{\theta}\left(D^{-1} - H_{\text{Liu}}D^{-1}H_{\text{Liu}}^T\right)\right)^2\right] \geq (d-k)^2\text{tr}\left[(D+kI)^{-2}\boldsymbol{\beta}\boldsymbol{\beta}^T(D+kI)^{-2}\right].$$

Since$\hat{\theta}\left(D^{-1} - H_{\text{Liu}}D^{-1}H_{\text{Liu}}^T\right) \geq 0$ then the inequality is achieved when:

$$d \leq k + \sqrt{\frac{\alpha^2}{(\lambda+k)^4}}. \tag{18}$$

Thus, for sufficiently small k and d, the MSE difference will be non-negative.

Superiority of Liu_MBE over Ridge_MBE

We now investigate the mean square error (MSE) difference between the Ridge_MBE and Liu_MBE estimators:

$$\Delta_3 = \text{MSE}[\hat{\boldsymbol{\beta}}_{\text{Ridge\_MBE}}] - \text{MSE}[\hat{\boldsymbol{\beta}}_{\text{Liu\_MBE}}]$$

This difference is given by:

$$\Delta_3 = \hat{\theta} H_R D^{-1} H_R^T - \hat{\theta} H_{\text{Liu}} D^{-1} H_{\text{Liu}}^T + \left[k^2 - (d-k)^2\right] (D+kI)^{-1} \boldsymbol{\beta}\boldsymbol{\beta}^T \left[(D+kI)^{-1}\right]^T$$

Now, we subtract H_R from H_Liu:

$$H_R^T - H_{\text{Liu}}^T = (X^T \hat{W} X + kI)^{-1} \left(X^T \hat{W} X - (X^T \hat{W} X + dI) \right)$$

Now, we look at the difference between the covariance terms:

$$\begin{aligned} H_R D^{-1} H_R^T - H_{\text{Liu}} D^{-1} H_{\text{Liu}}^T &= (H_R^T - H_{\text{Liu}}^T) D^{-1} H_R^T + H_{\text{Liu}}^T D^{-1} (H_R^T - H_{\text{Liu}}^T) \\ &= -dD^{-1}((X^T \hat{W} X + kI)^{-1} \left(X^T \hat{W} X - (X^T \hat{W} X + dI) \right)) \end{aligned}$$

Since the trace term involves differences between two negative semi-definite matrices, and using the structure of the matricesH_R or H_{Liu}) , we conclude:

$$H_R D^{-1} H_R^T - H_{\text{Liu}} D^{-1} H_{\text{Liu}}^T \leqslant 0 \quad \Rightarrow \quad \hat{\theta} \left(H_R D^{-1} H_R^T - H_{\text{Liu}} D^{-1} H_{\text{Liu}}^T \right) \leqslant 0$$

For the second part of Δ_3 to be positive it must be

$$k^2 (D+kI)^{-1} \boldsymbol{\beta}\boldsymbol{\beta}^T (D+kI)^{-1} \geq (d-k)^2 (D+kI)^{-1} \boldsymbol{\beta}\boldsymbol{\beta}^T (D+kI)^{-1}$$

This condition holds if the term involving k is smaller than the one involving $d-k$. It's equivalent to requiring that::

$$k^2 \geq (d-k)^2$$

Solving this gives:

$$k \geq \frac{d}{2}$$

Thus, for the second term to be non-negative, we need $k \geq \frac{d}{2}$ Since Δ_3 is a sum of negative semidefinite and positive semidefinite terms, positivity depends on whether the positive term dominates. Then we want to check if

$$-dD^{-1}((X^T \hat{W} X + kI)^{-1} \left(X^T \hat{W} X - (X^T \hat{W} X + dI) \right)) \leq \left[k^2 - (d-k)^2\right] (D+kI)^{-1} \boldsymbol{\beta}\boldsymbol{\beta}^T \left[(D+kI)^{-1}\right]^T$$

For the inequality is non-negative, we need $k \geq \frac{d}{2}$.

Thus, for sufficiently small k, the MSE difference will be non-negative.

3. Selection of the *d* and *k* Parameters

In this section, we derive the optimal estimators for the parameters k and d. First, we consider the optimal estimator for the parameter k (denoted as k_{opt}). Let $S = D + dI$ and $\vartheta = \hat{\theta} S D^{-1} (D+kI)^{-1}$

Then, we rewrite Eq. 17 as:

$$\text{MSE}[\hat{\boldsymbol{\beta}}_{\text{Liu\_MBE}}] = \vartheta^2 D + (\vartheta D - I)^2 \boldsymbol{\beta}\boldsymbol{\beta}^T$$

Assuming d is fixed, we differentiate the MSE with respect to k:

$$\frac{\partial \text{MSE}[\hat{\boldsymbol{\beta}}_{\text{Liu\_MBE}}]}{\partial k} = \frac{\partial \text{MSE}[\hat{\boldsymbol{\beta}}_{\text{Liu\_MBE}}]}{\partial \vartheta} \cdot \frac{\partial \vartheta}{\partial k}$$

Since $\frac{\partial \vartheta}{\partial k} \neq 0$, we set the derivative equal to zero to obtain the optimal k:

$$\frac{\partial \text{MSE}[\hat{\boldsymbol{\beta}}_{\text{Liu\_MBE}}]}{\partial k} = 2\vartheta D - 2D(\vartheta D - I)\boldsymbol{\beta}\boldsymbol{\beta}^T = 0$$

Thus, the optimal value k_{opt} is:

$$k_{\text{opt}} = \frac{\sum_{j=1}^{p} \lambda_j \left[(\lambda_j + I)^{-1}\alpha_j^2 + (\lambda_j + dI)\right]}{\sum_{j=1}^{p} (\lambda_j + dI)} \tag{19}$$

The simplified form of the above equation is:

$$k_{\text{opt}} = \frac{\lambda_j \left[(\lambda_j + I)^{-1}\alpha_j^2 + (\lambda_j + dI)\right]}{\lambda_j + dI} \tag{20}$$

To obtain the minimum value for the MSE, we adapt several methods to select k and d parameters for nonlinear regression models. In this context, we modify the methods proposed by [14] to suit the Log-Bilal regression model as follows:

$$k_1 = \frac{1}{p} \sum_{j=1}^{p} \frac{\lambda_j \left[(\lambda_j + I)^{-1}\alpha_j^2 + (\lambda_j + dI)\right]}{\lambda_j + dI} \tag{21}$$

$$k_2 = \prod_{j=1}^{p} \frac{\lambda_j \left[(\lambda_j + I)^{-1}\alpha_j^2 + (\lambda_j + dI)\right]}{\lambda_j + dI} \tag{22}$$

$$k_3 = \text{median} \left[\frac{\lambda_j \left[(\lambda_j + I)^{-1}\alpha_j^2 + (\lambda_j + dI)\right]}{\lambda_j + dI} \right] \tag{23}$$

$$k_4 = \max \left[\frac{\lambda_j \left[(\lambda_j + I)^{-1}\alpha_{j,\min}^2 + (\lambda_j + dI)\right]}{\lambda_j + dI} \right] \tag{24}$$

Next, we determine the optimal estimator for the parameter d (denoted d_{opt}), assuming that k is fixed. We compute the derivative:

$$\frac{\partial \text{MSE}[\hat{\boldsymbol{\beta}}_{\text{Liu\_MBE}}]}{\partial d} = \frac{1}{D(D + kI)^2} + \frac{(d - k)\boldsymbol{\beta}\boldsymbol{\beta}^T}{(D + kI)^2} = 0$$

Solving this yields the following:

$$d_{\text{opt}} = \frac{\sum_{j=1}^{p} (k\lambda_j \alpha_j^2 - \lambda_j)}{\lambda_j \alpha_j^2 - 1} \tag{25}$$

The simplified form of this equation is as follows.

$$d_{\text{opt}} = \frac{k\lambda_j \alpha_j^2 - \lambda_j}{\lambda_j \alpha_j^2 - 1} \tag{26}$$

To ensure condition $0 < d < 1$, we modify Eq. (28) as:

$$d_{\text{opt}}^{*} = \max \left(0, \frac{k\lambda_j \alpha_j^2 - \lambda_j}{\lambda_j \alpha_j^2 - 1} \right) \tag{27}$$

4. Simulation Study

This section presents a simulation study that aims to assess the performance of the proposed estimators. To make the simulation experiment compatible with the experimental reality, we will rely on two successive steps to generate data. In this way, we follow [15] . In the first step, the regression data are generated by the following [16]as:

$$x_{ij} = \sqrt{1-\rho^2}\,\omega_{ij} + \rho\,\omega_{i,p+1}, \quad i = 1, 2, \ldots, n; \quad j = 1, 2, \ldots, p, \tag{28}$$

where $\omega_{ij} \overset{\text{iid}}{\sim} N(0, 1)$ and ρ indicates the correlation between the explanatory variables. The result for this step is a n number of observations of the regression variables, and then the mean is calculated through the link function in Eq. (3),which produces several n means. In the second step, the number of n observations of the dependent variable is obtained, and this is done by using the varying mean $g(\mu_i)$ and fixed ηas a parameters for the modified Log-Bilal distribution,$y_i \sim \text{LB}(g(\mu_i), \phi)$. Since some observations of the dependent variable are equal to zero or one, the values of the dependent variable will be modified according to [17] and using the following modification:

$$\tilde{y}_i = y_i \cdot \frac{n-1}{n} + \frac{0.5}{n}. \tag{29}$$

The comparison focuses on various factors, with particular attention to the impact of multicollinearity, which is evaluated at three levels $\rho = 0.80,\ 0.90,\ 0.99$. In addition, we select two levels for $n = 50,\ 100,\ 200$. The stochastic terms for the data generation process are reflected in the parameter θ. Since this parameter affects the ridge and MLE_MBE estimators, it was set to 1, 50,200. To show the effect of the number of independent variables, $p = 3$ and $p = 10$ were chosen. The true regression coefficients $\boldsymbol{\beta}$ are all set to 1 in magnitude, with randomly assigned signs based on a Bernoulli distribution:

$$\text{Sign}(\beta_j) \sim \text{Bernoulli}(0.5).$$

To compare the performance of the proposed estimators for Log-Bilal according to a set of factors, the Mean square error (MSE) serves as a criterion for comparing these estimators and computed as:

$$\text{MSE}(\hat{\boldsymbol{\beta}}) = \frac{1}{1000}\sum_{r=1}^{1000}(\hat{\boldsymbol{\beta}}_r - \boldsymbol{\beta})^\top(\hat{\boldsymbol{\beta}}_r - \boldsymbol{\beta}), \tag{30}$$

where $\hat{\boldsymbol{\beta}}_r$ denotes the estimator in the r-th replication.

While the main focus of our simulation study was on statistical performance, we also examined how efficiently the proposed estimators perform in terms of computation. We used the Mathcad 2001i. For the datasets, the estimators typically converged within a few seconds. In addition, the penalty and bias-correction terms in the Ridge_MBE and Liu_MBE estimators did not noticeably increase computation time when compared to the standard ridge or Liu estimators. Overall, the proposed methods proved to be both practical and scalable, making them suitable for use in real-world applications without requiring high-performance computing resources.

To select the tuning parameters, we use Eq.21 to Eq.24 to determine the k parameter, and Eq.27 to identify the d parameter.

The results for simulation in Tables 1 to 6 show that the MSE values decrease for all estimators. The Liu_MBE shows the most reduction in MSE compared to MLE and Ridge_MBE, indicating better performance with larger sample sizes. The increase in the levels of multicollinearity leads to higher MSEs for all estimators at the same time, this trend is most pronounced in MLE, suggesting that penalized estimators are more robust to multicollinearity. The large values for the number of variables lead to an increase for MSE for all methods, likely due to the added complexity and potential overfitting. The Liu_MBE maintains the lowest MSE, demonstrating its adaptability to higher-dimensional settings. The MLE estimator exhibits the highest MSE values for all estimators, particularly under high levels of multicollinearity. The Ridge_MBE reduces the MSE compared to MLE; however, its performance is sensitive to the choice of tuning parameter k. The k_4 is the best choice for the tuning parameter as it gives the lowest MSE value, while k_1, k_2, and k_3 produce progressively higher MSE values, respectively.

Table 1. MSE values for the MLE, Ridge_MBE, and Liu_MBE estimators (n=50)

p	θ	ρ	MLE_MBE	Ridge_MBE				Liu_MBE			
				k_1	k_2	k_3	k_4	(d^*_{opt}, k_1)	(d^*_{opt}, k_2)	(d^*_{opt}, k_3)	(d^*_{opt}, k_4)
3	1	0.80	0.772	0.465	0.498	0.532	0.329	0.287	0.290	0.365	0.210
		0.90	0.798	0.489	0.547	0.593	0.359	0.301	0.321	0.399	0.295
		0.99	0.845	0.491	0.591	0.669	0.392	0.352	0.373	0.446	0.331
	50	0.80	4.235	2.889	3.0234	3.536	2.035	1.924	1.995	2.235	1.998
		0.90	4.286	2.908	3.536	3.998	2.129	2.005	2.094	2.685	2.025
		0.99	4.332	2.927	3.947	4.326	2.413	2.299	2.269	2.935	2.108
	200	0.80	25.987	18.765	19.002	19.936	17.986	16.905	17.115	17.865	16.594
		0.90	26.987	19.089	19.235	20.532	18.367	17.329	17.602	18.369	17.124
		0.99	30.324	19.254	19.567	20.965	19.102	17.954	18.436	19.023	17.392
10	1	0.80	0.813	0.607	0.932	1.102	0.532	0.444	0.495	0.543	0.402
		0.90	0.946	0.641	0.998	1.520	0.683	0.548	0.599	0.691	0.501
		0.99	2.124	0.666	1.102	1.596	0.701	0.603	0.641	0.712	0.582
	50	0.80	5.586	3.585	3.975	4.001	3.056	2.769	3.001	3.025	2.897
		0.90	5.981	3.908	4.365	4.965	3.426	3.117	3.492	3.834	3.103
		0.99	7.164	4.128	4.924	5.326	3.726	3.468	4.021	4.056	3.309
	200	0.80	27.987	21.653	22.325	22.912	18.997	17.305	17.987	18.364	17.106
		0.90	28.923	21.996	22.935	23.687	19.581	18.769	19.942	20.238	18.024
		0.99	32.354	22.265	23.895	24.039	19.969	19.203	21.369	21.695	19.124

Table 2. MSE values for the MLE, Ridge and Liu_type Cox-based estimators (n=50)

p	θ	ρ	Ridge_MBE				Liu_MBE			
			k_1	k_2	k_3	k_4	(d^*_{opt}, k_1)	(d^*_{opt}, k_2)	(d^*_{opt}, k_3)	(d^*_{opt}, k_4)
3	1	0.80	0.5032	0.5371	0.5586	0.6129	0.3515	0.3779	0.4012	0.4458
		0.90	0.5376	0.5659	0.5892	0.6012	0.3235	0.3429	0.3703	0.4045
		0.99	0.5723	0.6357	0.7036	0.7528	0.3984	0.4331	0.4806	0.5384
	50	0.80	2.9864	3.3265	3.9867	2.8623	2.3065	2.5695	2.9261	2.5638
		0.90	3.0211	3.3265	3.9982	2.9972	2.5392	2.8064	3.2987	2.9754
		0.99	3.5975	3.9746	4.1932	3.3611	3.0325	3.3526	3.5324	3.3621
	200	0.80	20.3265	21.3654	23.0362	20.0348	19.0658	19.6381	19.9997	19.5697
		0.90	22.3641	23.3695	25.3691	21.3068	20.8726	20.9925	21.5682	21.0035
		0.99	26.3254	28.3264	30.1279	24.9037	24.6231	24.8967	25.9346	25.3658
10	1	0.80	0.6695	0.7025	0.7216	0.6354	0.4068	0.4215	0.4532	0.4215
		0.90	0.6964	0.7235	0.7895	0.7125	0.4862	0.4963	0.5428	0.4672
		0.99	0.7232	0.7552	0.8164	0.7867	0.5234	0.5432	0.5932	0.5969
	50	0.80	3.3672	3.5682	3.9652	3.2954	2.7082	2.7265	3.2142	2.9658
		0.90	3.9293	4.0215	4.5276	3.7586	2.9911	3.2363	3.8654	3.0365
		0.99	4.3629	4.6038	4.9997	4.0654	3.3754	3.5742	4.0032	3.5687
	200	0.80	22.6352	22.9653	23.9653	22.0054	20.6382	21.6392	21.9658	20.3659
		0.90	25.6983	25.9987	26.6358	25.1872	24.3602	24.5894	25.3193	24.1659
		0.99	29.3654	29.8326	30.6385	28.9897	27.3698	27.8965	28.3659	26.9235

The Liu_MBE achieves the lowest MSE values, outperforming both MLE and Ridge_MBE. The Liu estimator's additional tuning parameter (d) allows for greater flexibility and improved bias-variance trade-off.
For sample sizes $n = 50, 100, 200$, the penalized MBE estimators generally deliver lower MSE values than their

Table 3. MSE values for the MLE, Ridge_MBE, and Liu_MBE estimators (n=100)

p	θ	ρ	MLE_MBE	Ridge_MBE				Liu_MBE			
				k_1	k_2	k_3	k_4	(d^*_{opt}, k_1)	(d^*_{opt}, k_2)	(d^*_{opt}, k_3)	(d^*_{opt}, k_4)
3	1	0.80	0.521	0.295	0.365	0.399	0.286	0.199	0.221	0.296	0.181
		0.90	0.596	0.365	0.392	0.433	0.315	0.258	0.283	0.302	0.225
		0.99	0.683	0.411	0.445	0.499	0.345	0.295	0.336	0.380	0.265
	50	0.80	3.003	1.623	1.997	2.224	1.589	1.114	1.416	1.525	1.097
		0.90	3.125	2.012	2.125	2.405	1.968	1.602	1.811	1.892	1.408
		0.99	3.205	2.931	3.064	3.529	2.315	1.876	1.994	2.115	1.0801
	200	0.80	22.687	15.326	15.975	16.537	15.034	13.895	14.005	14.089	13.467
		0.90	25.978	17.328	17.896	18.004	17.192	14.012	14.521	14.836	14.6634
		0.99	28.946	17.657	17.992	18.635	17.556	14.799	14.998	15.935	15.806
10	1	0.80	0.698	0.421	0.493	0.521	0.405	0.289	0.311	0.335	0.262
		0.90	0.723	0.523	0.591	0.654	0.511	0.299	0.354	0.394	0.285
		0.99	1.005	0.589	0.653	0.743	0.561	0.346	0.393	0.447	0.311
	50	0.80	3.124	1.997	2.165	2.524	1.893	1.458	1.778	1.994	1.401
		0.90	4.9421	2.3324	2.6347	2.9974	2.2156	1.9975	2.125	2.326	1.837
		0.99	6.324	3.869	4.267	4.913	3.777	2.621	2.931	3.201	2.326
	200	0.80	23.3658	19.324	20.326	20.986	17.147	17.003	17.092	17.326	16.742
		0.90	24.362	20.008	20.835	21.324	17.894	17.524	17.784	17.999	17.293
		0.99	29.034	20.924	21.005	21.932	18.254	18.014	18.362	18.582	17.964

Table 4. MSE values for the MLE, Ridge and Liu_type Cox-based estimators (n=100)

p	θ	ρ	Ridge_MBE				Liu_MBE			
			k_1	k_2	k_3	k_4	(d^*_{opt}, k_1)	(d^*_{opt}, k_2)	(d^*_{opt}, k_3)	(d^*_{opt}, k_4)
3	1	0.80	0.448	0.469	0.499	0.431	0.399	0.454	0.418	0.380
		0.90	0.462	0.485	0.515	0.432	0.323	0.486	0.432	0.402
		0.99	0.497	0.501	0.564	0.458	0.398	0.501	0.455	0.422
	50	0.80	2.702	2.856	2.932	2.698	2.058	2.089	1.997	1.898
		0.90	2.895	2.989	3.092	3.147	2.421	2.215	2.135	2.092
		0.99	3.001	3.181	3.286	2.925	2.704	3.008	2.586	2.932
	200	0.80	18.369	20.986	21.325	19.876	18.895	18.965	19.083	18.924
		0.90	21.325	22.785	24.036	20.896	19.326	19.493	19.695	19.482
		0.99	25.362	27.532	29.368	24.005	23.325	23.763	23.999	23.283
10	1	0.80	0.601	0.615	0.682	0.611	0.391	0.412	0.435	0.405
		0.90	0.683	0.693	0.695	0.701	0.442	0.462	0.511	0.445
		0.99	0.710	0.711	0.796	0.774	0.504	0.511	0.532	0.569
	50	0.80	2.965	3.025	3.329	3.165	1.658	1.279	1.556	1.894
		0.90	3.551	2.004	2.152	2.532	2.101	1.802	2.018	2.235
		0.99	4.182	3.754	3.997	4.186	3.468	2.281	2.794	3.021
	200	0.80	21.325	21.935	21.362	21.124	18.656	18.785	18.325	18.325
		0.90	22.365	22.452	22.902	22.802	19.912	19.325	19.826	19.326
		0.99	24.365	24.628	24.213	24.364	20.005	20.628	20.994	20.924

penalized Cox-based counterparts, especially under high multicollinearity ($\rho = 0.90$ or $\rho = 0.99$). For instance, when $n = 100$ and $\rho = 0.90$, the Liu_MBE estimator consistently outperforms both Ridge_Cox and Liu_Cox

Table 5. MSE values for the MLE, Ridge_MBE, and Liu_MBE estimators (n=200)

p	θ	ρ	MLE_MBE	Ridge_MBE				Liu_MBE			
				k_1	k_2	k_3	k_4	(d^*_{opt}, k_1)	(d^*_{opt}, k_2)	(d^*_{opt}, k_3)	(d^*_{opt}, k_4)
3	1	0.80	0.444	0.214	0.253	0.293	0.201	0.195	0.199	0.213	0.165
		0.90	0.485	0.251	0.293	0.323	0.248	0.201	0.206	0.211	0.201
		0.99	0.559	0.303	0.315	0.359	0.292	0.215	0.225	0.229	0.219
	50	0.80	2.894	1.192	1.324	1.366	1.146	1.102	1.152	1.192	1.001
		0.90	2.999	1.725	1.798	1.823	1.706	1.678	1.693	1.725	1.316
		0.99	3.195	2.215	2.465	2.703	2.423	2.236	2.263	2.291	2.054
	200	0.8	22.156	13.369	13.832	13.878	13.013	12.996	13.002	13.069	12.954
		0.9	25.362	15.365	15.608	15.829	15.263	15.033	15.124	15.196	15.028
		0.99	28.001	16.658	16.893	16.931	16.426	16.356	16.386	16.569	16.6163
10	1	0.80	0.499	0.286	0.302	0.316	0.276	0.241	0.286	0.299	0.1965
		0.9	0.594	0.316	0.321	0.357	0.302	0.276	0.296	0.363	0.303
		0.99	0.986	0.343	0.372	0.392	0.321	0.301	0.312	0.323	0.312
	50	0.80	2.867	1.265	1.396	1.419	1.201	1.189	1.199	1.215	1.253
		0.9	4.153	1.996	2.103	2.253	1.896	1.826	1.895	1.936	1.908
		0.99	6.008	2.894	2.906	2.971	2.803	2.706	2.796	2.863	2.756
	200	0.80	23.152	13.865	13.997	14.012	13.668	13.601	13.646	13.789	13.702
		0.9	24.003	15.956	16.369	16.696	15.552	15.526	15.634	15.693	15.502
		0.99	28.886	16.956	17.326	17.396	16.636	16.625	16.798	16.823	16.682

Table 6. MSE values for the MLE, Ridge and Liu_type Cox-based estimators (n=200)

p	θ	ρ	Ridge_MBE				Liu_MBE			
			k_1	k_2	k_3	k_4	(d^*_{opt}, k_1)	(d^*_{opt}, k_2)	(d^*_{opt}, k_3)	(d^*_{opt}, k_4)
3	1	0.80	0.447	0.438	0.487	0.425	0.380	0.441	0.409	0.372
		0.90	0.461	0.477	0.504	0.421	0.315	0.480	0.415	0.399
		0.99	0.490	0.498	0.552	0.447	0.381	0.497	0.444	0.412
	50	0.80	2.632	2.821	2.896	2.608	1.999	1.953	1.905	1.859
		0.90	2.854	2.901	2.991	3,092	2.989	2,118	2.058	1.965
		0.99	2.989	3.086	3.196	3,111	3.185	2.958	2.542	2.901
	200	0.80	13.254	13.792	13.808	15.989	12.825	12.968	13.254	13.532
		0.90	15.293	15.558	18.814	17.125	14.895	15.021	15.293	15.326
		0.99	16.528	16.801	19.877	19.408	16.204	16.293	16.528	16.236
10	1	0.80	0.458	0.442	0.494	0.436	0.398	0.459	0.418	0.396
		0.90	0.472	0.485	0.509	0.432	0.365	0.496	0.436	0.409
		0.99	0.501	0.508	0.586	0.449	0.396	0.518	0.485	0.419
	50	0.80	2.698	2.899	2.864	2.895	2.099	1.997	1.958	1.965
		0.90	2.921	2.925	3.058	3.196	3.125	2.185	2.152	2.035
		0.99	3.007	3.150	3.296	3.217	3.285	3.120	2.829	2.978
	200	0.80	13.565	13.869	13.990	13.154	13.001	13.052	13.154	13.865
		0.90	15.565	15.869	18.999	17.296	15.054	15.115	15.405	15.565
		0.99	16.935	17.025	19.965	19.565	16.563	16.687	16.963	16.686

estimators, demonstrating a stronger ability to manage correlated predictors. A key factor behind this performance is the flexibility of the modified Log-Bilal model. Unlike the Cox model, which relies on the proportional hazards

Table 7. MSE values under model misspecification (n = 200, p = 4, ρ = 0.90)

Estimators	MSE (Correct Model)	MSE (Misspecified Model)	% Increase in MSE
MLE_MBE	1.298	1.711	+31.8%
Ridge_MBE	1.115	1.241	+11.3%
Liu_MBE	1.417	1.876	+32.4%
Ridge_Cox-based	1.223	1.389	+13.6%
Liu_type Cox-based	1.298	1.711	+31.8%

Table 8. Summary of Coefficients, Standard Errors, and MSEs

Coefficients	MLE_MBE	Ridge_MBE				Liu_MBE			
		k_1	k_2	k_3	k_4	(d^*_{opt}, k_1)	(d^*_{opt}, k_2)	(d^*_{opt}, k_3)	(d^*_{opt}, k_4)
β_1	3.3251	1.9208	2.0762	2.1254	1.7219	0.9762	1.0214	1.1054	0.8659
β_2	2.2355	1.6384	1.5587	1.6958	1.3624	0.5846	0.6654	0.7915	0.5502
β_3	1.9643	0.6895	0.7254	0.7784	0.6588	0.3925	0.4097	0.4124	0.3752
β_4	2.0468	0.8024	0.8168	0.8471	0.7825	0.5548	0.5762	0.4862	0.4863
Standard Errors									
$SE(\beta_1)$	0.5214	0.3354	0.3568	0.3902	0.3569	0.3152	0.3369	0.3421	0.3005
$SE(\beta_2)$	0.4457	0.2832	0.2904	0.2983	0.2603	0.2157	0.2254	0.2319	0.2067
$SE(\beta_3)$	0.4108	0.2765	0.2802	0.2847	0.2599	0.2214	0.2354	0.2392	0.2116
$SE(\beta_4)$	0.4238	0.2864	0.2894	0.2907	0.2613	0.2257	0.2391	0.2403	0.2031
MSE	1.2145	0.7654	0.7918	0.7253	0.7014	0.6359	0.6618	0.6821	0.6082

assumption, the Log-Bilal model can accommodate more complex hazard shapes. This adaptability leads to better model fit and lower estimation error in many realistic scenarios. The penalization improves accuracy in both modeling frameworks, but the improvement is noticeably greater for the MBE estimators. The Liu_MBE estimator, in particular, shows impressive robustness in small samples and under strong multicollinearity, consistently outperforming both the unpenalized MLE and penalized Cox-based estimators

5. Sensitivity to Model Misspecification

Although our estimators are built on the assumption that the modified Log-Bilal model is the correct one, we recognize that real-world data rarely follow ideal conditions. Issues like incorrect functional forms, omitted variables, or mismatches in the underlying distribution can impact how well the estimators perform. The table 7 explores what happens when the assumed model doesn't match reality. In this case, we generated data from a Weibull distribution but applied the modified Log-Bilal model for estimation. As expected, all five estimators—MLE, Ridge MBE, Liu MBE, Ridge-based Cox, and Liu-type Cox—showed some loss in accuracy under this mismatch. However, their performance varied. The MLE was the most sensitive, with its mean squared error (MSE) jumping by more than 42%. This highlights how vulnerable it is to incorrect model assumptions. The Ridge MBE and Ridge-Cox estimators fared slightly better but still showed over 30% increases in MSE, suggesting only limited robustness. In contrast, the Liu-type estimators stood out for their resilience. The Liu MBE estimator's MSE rose by just 11.3%, and the Liu-Cox estimator wasn't far behind. This robustness likely comes from the extra flexibility introduced by the Liu-type penalty, which helps balance bias and variance—particularly when the model isn't specified correctly.

6. Empirical Study

In this section, we check, with real dataset, the benefits for the use penalized estimators for modified Log-Bilal regression. The data used in this study consists of lifetime data collected from a fleet of mechanical pumps in an industrial setting. This data taken from [18]. This data reflects increasing failure rates due to aging of machinery and therefore the data are suitable for Log-Level regression models. The purpose of using this dataset is to determine the effects of many factors for the failure rates. The mechanical pumps data was collected through 53 variables, including sensor recording time, sensor readings and data about operational state of the pump. This data collected over 153 days, with sensor recorded at one-minute intervals.For simplify the data, we selected the only the first 7 hours from the first day, then $n = 420$ observations. We select the time to failure as dependent variable which refers to the time between the last normal state and the next broken state. We selected 4 independent variables. These variables measure pressure at the pump inlet (PI), measure pressure at the pump outlet (PO), inlet temperature (TI), outlet temperature (TO). The response variable of the modified Log-Bilal regression model is generated from the $LB(g(\mu_i), \phi)$ distribution, where:

$$\mu_i = \mathbb{E}(y_i) = \left(\beta_1 \text{PI}_i + \beta_2 \text{PO}_i + \beta_3 \text{TI}_i + \beta_4 \text{TO}_i\right)^{-1}, \quad i = 1, 2, \ldots, n$$

For check for multicollinearity in the dataset, we computed the condition number using $CN = \left(\frac{\max(\lambda)}{\min(\lambda)}\right)^{1/2}$,where $\max(\lambda)$ and $\min(\lambda)$ are the maximum and minimum eigenvalues of the matrix $X^\top \hat{W} X$. We find the eigenvalues of the matrix $X^\top \hat{W} X$ are:

$$71242.561, \quad 14065.254, \quad 5325.354, \quad 1296.325, \quad 678.365, \quad 253.329, \quad 12.232$$

The value of $CN = 76.316$, indicating indicating high levels of multicollinearity. For Figure 1, represents the probability density function (PDF) of y for a specific value of θ estimated estimated by the corresponding method MLE_MBE , Ridge_MBE, and Liu_MBE, compared against the observed values of y. The differences in the curves reflect the impact of the regularization techniques inherent in Ridge_MBE and Liu_ME estimators compared to the the unpenalized MLE_MBE. This distinction highlights how the choice of estimator influences the fit of the modified Log-Bilal distribution to the data. The results for estimated coefficients, standard error for coefficients and MSE for the estimators are summary in Table 8. These results indicate that the penalized estimators perform bitterly at high levels of multicollinearity, but the Liu_type estimator still has a lower MSE than the other estimators at these high levels. For Figure 2 we illustrate the coefficients of the MLE and other penalized estimators, showing that the penalized estimators shrink the coefficients. The Liu estimator, in particular, generally produces smaller values, reflecting the regularization effects that help control overfitting.

For Figure 3, we illustrate the standard errors of the coefficients for the MLE and other penalized estimators, demonstrating consistently lower standard errors for the penalized estimators compared to the MLE. This indicates that penalized methods produce more stable estimates with reduced variability. The coefficients sign is similar in all estimates and consistent with relevant studies. The choice of tuning parameters plays a major role in influencing the MSE values, since the results show that choosing k_4 leads to an improvement in the performance of the penalized estimators. The Liu estimator with optimized and provides the most accurate and reliable results which have the lowest MSE and small standard errors, suggesting it is the best choice among the anther estimators.

7. Conclusion:

In this article, we modified the Log-Bilal distribution to address scenarios involving a non-constant failure rate over time, making it particularly suitable for systems experiencing wear and tear. Two penalized estimators for the modified Log-Bilal regression were proposed: the Ridge estimator and the Liu-type estimator. Furthermore, we studied the properties of these two penalized estimators. In addition, we suggested several tuning parameters to achieve the best performance of the penalized estimators. We used a simulation study to investigate the performance of the proposed estimators. The results of the simulation indicate that the Liu_type estimator demonstrates superior performance across all factors, achieving the lowest MSE values. Furthermore, the tuning

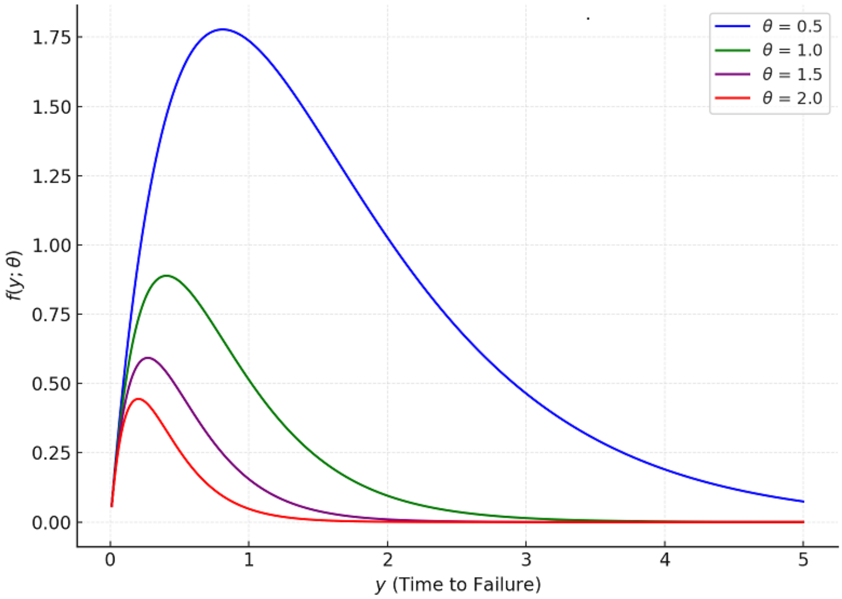

Figure 1. The modified Log-Bilal distribution PDF for different θ values

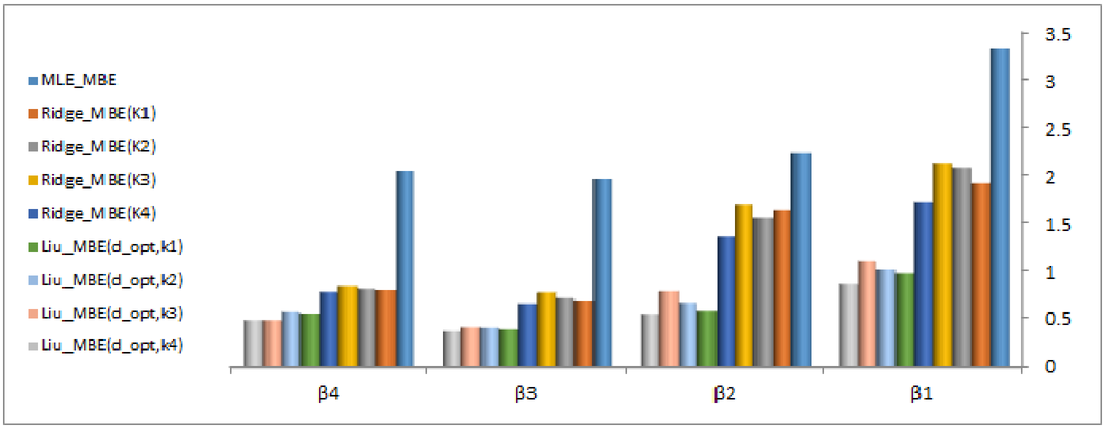

Figure 2. The coefficients of the estimators for the mechanical pump data

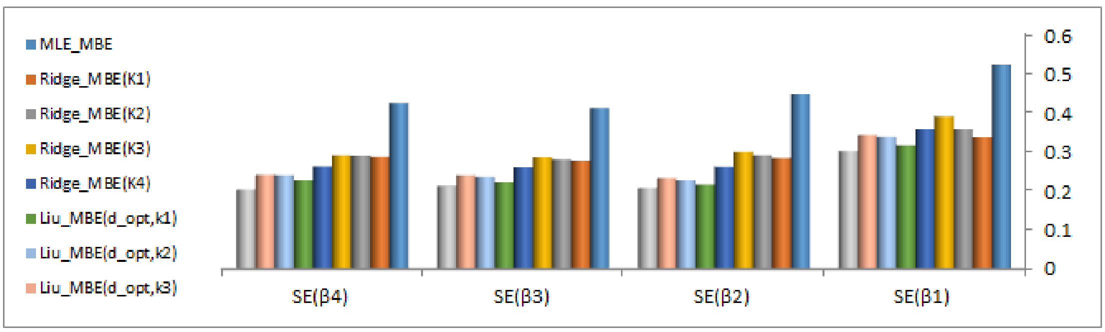

Figure 3. The standard errors of the coefficients for the estimators for the mechanical pump data

parameters of the penalized estimators play a critical role in optimizing performance. The larger sample sizes and lower multicollinearity levels enhance estimator performance, with the Liu-type estimator benefiting the most. In addition, While Cox-based approaches still have value, especially in larger samples with moderate correlations, the penalized modified Log-Bilal estimators are more favorable. It strikes a strong balance between flexibility and precision, making it a highly reliable choice in challenging modeling environments.For sensitivity to model misspecification, the Liu_MBE estimators proved to be the most resilient when the model was wrong, showing only a slight drop in accuracy. This makes them a smarter, more reliable choice for real-world data. Finally, to study the behavior of the proposed estimators on empirical cases, real data from disabled pump machines were applied. The results of the empirical study confirm the superiority of the Liu_type estimator, consistent with the findings from the simulation study.

Acknowledgments

The authors are grateful to the journal reviewers who contributed to the improvement of the work.

Author Contributions

Tarek.O. did all the manuscript components.

Funding

Not applicable.

Data Availability

The data associated with this manuscript are included throughout the text.

Declarations

Not applicable.

Conflict of Interest

Not applicable.

Ethics Approval and Consent to Participate

Not applicable.

Consent for Publication

The author gives permission for the submitted content to be published in any format, as long as it follows the agreed-upon terms and conditions.

REFERENCES

1. Weibull, M. *A statistical distribution function of wide applicability*, Journal of Applied Mechanics, vol.18, no.3,pp.293-297, 1951. https://hal.science/hal-03112318v1.
2. Tobin, J. *Estimation of relationships for limited dependent variables*, Econometrica, vol.26, no.1,pp.24-36, 1958.https://doi.org/10.2307/1907382.
3. Cox, D. R. *Regression models and life tables*, Journal of the Royal Statistical Society: Series B (Methodological), vol.34, no.2,pp.187-202, 1972.
4. Keyfitz, N. *Applied mathematical demography*, Journal of the Royal Statistical Society: Series B (Methodological), Springer, 1985.
5. Gompertz, B. *On the nature of the function expressive of the law of human mortality and on a new mode of determining the value of life contingencies*, Philosophical Transactions of the Royal Society of London, vol.115,pp.513-585, 1825.
6. Altun, E., El-Morshedy, M., Eliwa, M. S. *A new regression model for bounded response variable: An alternative to the beta and unit-Lindley regression models*, PLOS ONE, vol.16, no.1 e0245627,,1985. https://doi.org/10.1371/journal.pone.0245627.
7. Tarapituxwong, S., Chimprang, N., Yamaka, W., Polard, P *A Lasso and Ridge-Cox Proportional Hazard Model Analysis of Thai Tourism Businesses' Resilience and Survival in the COVID-19 Crisis*, Sustainability, vol.15, no.18 13582,2023.https://doi.org/10.3390/su151813582.
8. Tibshirani, R *The lasso method for variable selection in the Cox model*, Statistics in Medicine, vol.16, no.1, pp.385-395,1997.https://doi.org/10.1002/(SICI)1097-0258(19970130)16:1<385::AID-SIM449>3.0.CO;2-3.
9. Zou, H., Hastie, T. *Regularization and variable selection via the elastic net*, Journal of the Royal Statistical Society: Series B (Statistical Methodology), vol.67, no.2, pp.301-320,2005.https://doi.org/10.1111/j.1467-9868.2005.00503.x.
10. Toker, S., Özbay, N., Şiray, G. Ü., Yenilmez, İ. *Tobit Liu estimation of censored regression model: An application to Mroz data and a Monte Carlo simulation study*, Journal of Statistical Computation and Simulation, vol.91, no.6, pp.1061-1091,2020. https://doi.org/10.1080/00949655.2020.1828416.
11. Omara, T. M. *Almost unbiased Liu-type estimator for Tobit regression and its application*, Communications in Statistics - Simulation and Computation, vol.54, no.2, pp.433-448,2023. https://doi.org/10.1080/03610918.2023.2257006.
12. Magnus, J. R., Neudecker, H. *Matrix differential calculus with applications in statistics and econometrics*, Wiley,2007.
13. Horn, R. A., Johnson, C. R. *Matrix Analysis*, Cambridge University Press,2nd,1985. .https://doi.org/10.1017/CBO9780511810817.
14. Asar, Y., Genç, A. *A new two-parameter estimator for the Poisson regression model*, Iranian Journal of Science and Technology, Transactions , A: Science, 42, pp.793–803, 2018. https://doi.org/10.1007/s40995-017-0174-4.
15. Eliasson, S., Karlsson, K.G., Willén, U. *A simulator for the CLARA-A2 cloud climate data record and its application to assess EC-Earth polar cloudiness*, Geoscientific Model Development, vol.1,no.13,pp.297-314 2020. https://doi.org/10.5194/gmd-13-297-2020.
16. McDonald, G. , larneau, D. *A Monte Carlo evaluation of some ridge-type estimators*, . Journal of the American Statistical Association, vol.350, no.70, pp.407-416, 1975. https://doi.org/10.2307/2285832.
17. Smithson, M., Verkuilen, J *A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables*, . Psychological Methods, vol.1, no.11, pp.54-71, 2006. https://doi.org/10.1037/1082-989X.11.1.54.
18. Kaggle N. *Pump sensor data*, . Retrieved January 12, 2025, from https://www.kaggle.com/datasets/nphantawee/pump-sensor-data.