STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **16**, September 2026, pp 2137–2158.
Published online in International Academic Press (www.IAPress.org)

A New Efficient Biased Estimator in Overdispersed Count Regression for Effectively Handling Multicollinearity in COVID-19 Data from Saudi Arabia

Eslam Hussam [1], Ali T. Hammad [2,*], Ramy Aldallal [3], and Ahmed M. Gemeay[2]

[1] *Department of Accounting, College of Business Administration in Hawtat Bani Tamim, Prince Sattam bin Abdulaziz University, Saudi Arabia; e.hafez@psau.edu.sa*
[2] *Department of Mathematics, Faculty of Science, Tanta University, Tanta 31527, Egypt; ali.taha@science.tanta.edu.eg; ahmed.gemeay@science.tanta.edu.eg*
[3] *Department of Management, College of Business Administration in Hawtat Bani Tamim, Prince Sattam bin Abdulaziz University, Saudi Arabia, r.eldallal@psau.edu.sa*

Abstract The negative binomial regression model is a common choice for count data, offering flexibility through an added dispersion parameter that accounts for excess variability. But when explanatory variables are highly correlated, maximum likelihood estimation breaks down: variances inflate and coefficient estimates become unstable, undermining the validity of inference. This study addresses that problem by proposing a new biased estimator for negative binomial regression, one that combines the shrinkage properties of the ridge and adjusted Liu estimators within a unified two-parameter framework. We establish the estimator's theoretical properties by deriving its scalar and matrix mean squared error, along with formal dominance conditions against the maximum likelihood estimator and several competing biased alternatives. Its finite-sample performance is evaluated through an extensive Monte Carlo simulation across a range of sample sizes, predictor correlations, numbers of explanatory variables, and overdispersion levels. We also apply the estimator to Saudi Arabia COVID-19 mortality data, where it produces more narrower confidence intervals, stable coefficients, and substantially lower mean squared error than existing alternatives. Overall, these results provide support to the designed estimator as an effective tool for negative binomial regression under multicollinearity.

Keywords Mean squared error, Modified Liu-Ridge estimator, Multicollinearity, Monte Carlo simulation, Negative binomial regression.

AMS 2010 subject classifications 62J12, 62J07, 62H12, 62P10.

DOI: 10.19139/soic-2310-5070-3781

1. Introduction

Modeling count data presents distinct challenges owing to their typically highly asymmetric distributions; unlike Gaussian data, count records are not continuous or symmetric. Count data also tends to violate one of the central assumptions of traditional linear regression, namely that errors are normally distributed, a violation that would otherwise produce biased estimates. To ensure that count data models are constructed as effectively as possible, the earliest models developed for count data relied on the Poisson regression model (PRM) [1, 2]. The PRM rests on a very specific assumption, namely that the mean and variance of the count data are equal; that is, PRM models assume equidispersion. PRM models are widely used across count data collected in health, economics, and the social sciences; however, in nearly all such collections, the variance of the data tends to substantially exceed the

*Correspondence to: Ali T. Hammad (Email: ali.taha@science.tanta.edu.eg). Department of Mathematics, Faculty of Science, Tanta University, Tanta 31527, Egypt.

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

mean, violating the equidispersion condition. The PRM tends to fail under these circumstances, where standard errors are severely underestimated, and the resulting inferences are often misleading. The negative binomial regression model (NBRM) offered a further extension of the Poisson framework, and while it retains most of the structure and assumptions of the PRM, it introduces additional flexibility by allowing the variance to exceed the mean. In settings where equidispersion cannot be justified, the NBRM has been successfully adopted across a wide range of applications, from biostatistics to education [3, 4]. The NBRM uses the logarithm of the expected response as a linear function of the independent variables to model count data [5].

The NBRM assumes predictors are linearly independent. This rarely holds in practice: predictors move together, share underlying causes, and produce multicollinearity. Frisch [6] was the first to formally examine the issue, showing how correlated predictors distort estimation. Under multicollinearity, ML estimation of the negative binomial coefficients grows sensitive to these correlations, and the resulting estimates become unstable with inflated variances. Standard errors rise, confidence intervals widen, and coefficients can flip signs or lose significance, leaving individual predictors difficult to interpret or test reliably [7, 8].

Several biased estimation techniques address multicollinearity in linear and generalized linear models. Ridge regression, introduced by Hoerl and Kennard [9], was the first; it adds a small positive constant to the variance-covariance matrix to stabilize the coefficient estimates, and the idea was later extended beyond the linear model [10]. The Liu estimator took a related approach, using a single shrinkage parameter in a linear formulation rather than the ridge constant [11]. Further biased estimators developed from these two lines include the new ridge-type estimator [12], the modified ridge-type estimator [13], the Liu-type estimator [14], the modified one-parameter Liu estimator [15], the Dawoud-Kibria estimator [16], the modified Liu-ridge-type estimator [17], the new adjusted ridge-type estimator [18], the new modification of the ridge-type estimator [19], the modified jackknife Liu-type estimator [20], and the Özkale-Kaciranlar estimator [21], among others. Most aim to improve on their predecessors' estimation efficiency in specific regression settings.

Biased estimators have also been developed specifically within the NBRM to combat multicollinearity. Mansson [24, 25] initiated this line of work, introducing both ridge and Liu estimators adapted to the model. Asar [26] then proposed a Liu-type estimator for the model, and Kandemir Çetinkaya and Kaçıranlar [27] contributed the Özkale-Kaçıranlar estimator as a further shrinkage alternative. More recently, Koc and Koc [28] adapted the Kibria-Lukman estimator to the model and introduced its jackknife variant; Ashraf et al. [29] proposed a Stein estimator, Jabur et al. [30] introduced a jackknifed Liu-type estimator, Alghamdi et al. [32] suggested a new bias-reduced estimator, and Oranye et al. [31] developed a new biasing parameter for the jackknife Kibria-Lukman estimator, among others.

This study proposes a new modification to the Liu-Ridge framework for the NBRM, combining the shrinkage properties of both approaches into a single biased estimator for settings with high, though imperfect, multicollinearity among explanatory variables. We first establish the estimator's theoretical foundations, deriving its scalar mean squared error (SMSE) and matrix mean squared error (MMSE) and comparing these analytically against the ML estimator and other competing biased estimators. We also present and evaluate several methods for selecting the optimal biasing parameter. An extensive Monte Carlo simulation then assesses finite-sample performance under varying correlation, sample size, and overdispersion, using simulated mean squared error and mean absolute error as evaluation criteria. Finally, we apply the estimator to real healthcare data from Saudi Arabia, where correlated predictors test its performance against existing alternatives.

The remainder of the paper is organized as follows. Section 2 provides background on the NBRM, its standard estimation procedure, and the existing biased estimators that motivate this work. Section 3 presents the methodological core: the derivation of the proposed estimator, a theoretical comparison against its main competitors, and a discussion of alternative criteria for selecting the biasing parameter. Section 4 covers the simulation framework, detailing the Monte Carlo design and the main results, with particular attention to performance across different degrees of multicollinearity. Section 5 illustrates the estimator's practical behavior using real healthcare data. Section 6 concludes, noting the study's main limitations and directions for further work.

2. Statistical Methodology

The NBRM offers greater flexibility than the PRM when modeling count data that display excess variability. For a sample of n observations $y_1, y_2, \ldots, y_n$, the probability mass function of the negative binomial distribution can be expressed as:

$$f(y_i|x_i) = \frac{\Gamma(\psi^{-1} + y_i)}{\Gamma(\psi^{-1})\Gamma(1 + y_i)} \left(\frac{\psi^{-1}}{\psi^{-1} + \mu_i}\right)^{\psi^{-1}} \left(\frac{\mu_i}{\psi^{-1} + \mu_i}\right)^{y_i}, \quad y_i = 0, 1, 2, \ldots, \; i = 1, 2, \ldots, n, \tag{1}$$

where ψ represents the overdispersion parameter (defined as the inverse of τ), μ_i denotes the mean of the distribution for the i-th observation, and $\Gamma(\cdot)$ is the gamma function. The conditional mean and variance of the response variable are given, respectively, by:

$$\mathrm{E}(y_i|\mu_i, \tau_i) = \mu_i, \qquad \mathrm{Var}(y_i|\mu_i, \tau_i) = \mu_i(1 + \psi\mu_i).$$

Estimation of the coefficient vector β is typically carried out through ML estimation, where the mean parameter μ_i is linked to the explanatory variables through a logarithmic link function, $\log(\mu_i) = x_i^\top \beta$, where x_i is a $p \times 1$ column vector containing the values of the predictors for the i-th observation, i.e., $x_i = (x_{i1}, x_{i2}, \ldots, x_{ip})^\top$ and β is the $p \times 1$ vector of unknown regression coefficients to be estimated, i.e., $\beta = (\beta_0, \beta_1, \ldots, \beta_{p-1})^\top$. The corresponding log-likelihood function takes the following form:

$$\ell(\beta, \psi) = \sum_{i=1}^{n} \left[y_i \log\left(\frac{\mu_i}{\psi^{-1} + \mu_i}\right) + \log\Gamma(\psi^{-1} + y_i) + \psi^{-1} \log\left(\frac{\psi^{-1}}{\psi^{-1} + \mu_i}\right) - \log\Gamma(\psi^{-1}) - \log\Gamma(y_i + 1) \right]. \tag{2}$$

Taking the derivative of this expression with respect to β_j gives the score function:

$$\frac{\partial \ell(\beta, \psi)}{\partial \beta_j} = \sum_{i=1}^{n} \left(\frac{y_i - \mu_i}{1 + \psi\mu_i}\right) x_{ij} = 0, \quad j = 1, 2, \ldots, p. \tag{3}$$

Because this system of equations lacks a closed-form solution, parameter estimates must be obtained numerically. The iterative reweighted least squares algorithm is commonly employed for this purpose, with each iteration updating the estimate according to:

$$\hat{\beta}^{r+1} = \hat{\beta}^{r} + (I(\hat{\beta}^{r}))^{-1} S(\hat{\beta}^{r}), \tag{4}$$

where $I(\beta^r) = X^\top G^r X$ is the Fisher information matrix of dimension $p \times p$, evaluated at β^r and $S(\hat{\beta}^r)$ denotes the score function, with G^r representing an $n \times n$ diagonal weight matrix. Iterations continue until the change between successive estimates, $\|\hat{\beta}^r - \hat{\beta}^{r+1}\|$, falls below a pre-specified tolerance ϵ. Under this scheme, the ML estimates can be written iteratively as:

$$\hat{\beta}^{r+1} = (X^\top \hat{G}^r X)^{-1} X^\top \hat{G}^r s^r. \tag{5}$$

Upon convergence, the final ML estimate is obtained as:

$$\hat{\beta}_{\text{NBML}} = P^{-1} X^\top \hat{G} \hat{s}^{*}, \tag{6}$$

where $P = X^\top \hat{G} X$, $\hat{s}^*$ representing the converged adjusted response vector, $\hat{s}_i^* = \log(\hat{\mu}_i) + \frac{y_i - \hat{\mu}_i}{\hat{\mu}_i}$, and $\hat{G} = \operatorname{diag}\left(\frac{\hat{\mu}_i}{1+\psi\hat{\mu}_i}\right)$. The corresponding variance-covariance matrix of the estimator follows directly as:

$$\operatorname{Cov}(\hat{\beta}_{\text{NBML}}) = P^{-1}. \tag{7}$$

To derive the MMSE and the SMSE of the ML estimator ($\hat{\beta}_{\text{NBML}}$), the spectral decomposition of the matrix P is applied, expressed as $P = V\Theta^{-1}V^\top$, where V is the orthogonal matrix whose columns correspond to the eigenvectors of P, and Θ is a diagonal matrix containing its associated eigenvalues, $\Theta = \operatorname{diag}(\theta_1, \theta_2, \ldots, \theta_p)$. Using this decomposition, the MMSE and SMSE of the ML estimator can be obtained as follows:

$$\operatorname{MMSE}(\hat{\beta}_{\text{NBML}}) = V\Theta^{-1}V^\top, \tag{8}$$

$$\operatorname{SMSE}(\hat{\beta}_{\text{NBML}}) = \sum_{j=1}^{p} \frac{1}{\theta_j}. \tag{9}$$

When the predictor variables exhibit severe multicollinearity, the resulting weighted matrix becomes ill-conditioned, which in turn substantially inflates the variance of the ML. This inflated variance makes the regression coefficients unstable and unrealistically large, undermining the model's reliability and interpretability.

In addressing the challenges posed by multicollinearity within NBRM, a variety of biased estimators have been proposed in the literature. These approaches rely on the introduction of shrinkage parameters to enhance the stability of the coefficient estimates. Månsson [24] was among the first to explore this domain, introducing the negative binomial ridge (NBR) estimator, defined as:

$$\hat{\beta}_{\text{NBR}} = (P + kI)^{-1} P \hat{\beta}_{\text{NBML}}, \quad k > 0, \tag{10}$$

where the ridge parameter $k > 0$ dictates the intensity of the penalization and the estimator converges to the conventional ML as the parameter k approaches zero. The MMSE and SMSE metrics for the NBR estimator can be expressed as follows:

$$\operatorname{MMSE}(\hat{\beta}_{\text{NBR}}) = V\Theta_k^{-1}\Theta\Theta_k^{-1}V^\top + k^2 V\Theta_k^{-1}\alpha\alpha^\top\Theta_k^{-1}V^\top, \tag{11}$$

$$\operatorname{SMSE}(\hat{\beta}_{\text{NBR}}) = \sum_{j=1}^{p} \frac{\theta_j}{(\theta_j + k)^2} + k^2 \sum_{j=1}^{p} \frac{\alpha_j^2}{(\theta_j + k)^2}, \tag{12}$$

where $\Theta_k = \operatorname{diag}(\theta_1 + k, \theta_2 + k, ..., \theta_p + k)$ and $\alpha = V^\top \hat{\beta}_{\text{NBML}}$.

Expanding on this framework, Månsson [25] later developed the negative binomial Liu (NBL) estimator, which employs an alternative shrinkage formulation:

$$\hat{\beta}_{\text{NBL}} = (P + I)^{-1}(P + dI)\hat{\beta}_{\text{NBML}}, \quad 0 < d < 1, \tag{13}$$

where d is the Liu parameter typically restricted to the unit interval $(0, 1)$, with the NBL reduced to NBML when d tends to unity. The MMSE and SMSE corresponding to the NBL estimator are given below:

$$\operatorname{MMSE}(\hat{\beta}_{\text{NBL}}) = V\Theta_1^{-1}\Theta_d\Theta^{-1}\Theta_d\Theta_1^{-1}V^\top + (d-1)^2 V\Theta_1^{-1}\alpha\alpha^\top\Theta_1^{-1}V^\top, \tag{14}$$

$$\operatorname{SMSE}(\hat{\beta}_{\text{NBL}}) = \sum_{j=1}^{p} \frac{(\theta_j + d)^2}{\theta_j(\theta_j + 1)^2} + (d-1)^2 \sum_{j=1}^{p} \frac{\alpha_j^2}{(\theta_j + 1)^2}, \tag{15}$$

where $\Theta_1 = \text{diag}(\theta_1 + 1, \theta_2 + 1, ..., \theta_p + 1)$ and $\Theta_d = \text{diag}(\theta_1 + d, \theta_2 + d, ..., \theta_p + d)$.

Lukman et al. [15] introduced an alternative penalized estimator called the adjusted Liu estimator, which is defined for the NBRM, known as the negative binomial adjusted Liu (NBAL) estimator. This estimator is formulated as:

$$\hat{\beta}_{\text{NBAL}} = (P + I)^{-1}(P - dI)\hat{\beta}_{\text{NBML}}, \quad 0 < d < 1, \tag{16}$$

where the tuning parameter d is conventionally bounded within the unit interval $(0, 1)$. The corresponding MMSE and SMSE for this estimator are presented as follows:

$$\text{MMSE}(\hat{\beta}_{\text{NBAL}}) = V\Theta_1^{-1}\Theta_{-d}\Theta^{-1}\Theta_{-d}\Theta_1^{-1}V^\top + (1 + d)^2 V\Theta_1^{-1}\alpha\alpha^\top\Theta_1^{-1}V^\top, \tag{17}$$

$$\text{SMSE}(\hat{\beta}_{\text{NBAL}}) = \sum_{j=1}^{p} \frac{(\theta_j - d)^2}{\theta_j(\theta_j + 1)^2} + (d + 1)^2 \sum_{j=1}^{p} \frac{\alpha_j^2}{(\theta_j + 1)^2}, \tag{18}$$

where $\Theta_{-d} = \text{diag}(\theta_1 - d, \theta_2 - d, ..., \theta_p - d)$.

Asar [26] extended the methodology further by proposing the negative binomial Liu-type (NBLT) estimator, a generalized two-parameter model, which is defined as follows:

$$\hat{\beta}_{\text{NBLT}} = (P + kI)^{-1}(P - dI)\hat{\beta}_{\text{NBML}}, \quad -\infty < d < \infty,\ k > 0, \tag{19}$$

in which k and d jointly determine the degree of shrinkage. This formulation is sufficiently general to reduce to the NBR when d approaches zero and reducing to the NBML when both parameters equal zero. The corresponding MMSE and SMSE expressions for the NBLT estimator can be formulated as follows:

$$\text{MMSE}(\hat{\beta}_{\text{NBLT}}) = V\Theta_k^{-1}\Theta_{-d}\Theta^{-1}\Theta_{-d}\Theta_k^{-1}V^\top + (d + k)^2 V\Theta_k^{-1}\alpha\alpha^\top\Theta_k^{-1}V^\top, \tag{20}$$

$$\text{SMSE}(\hat{\beta}_{\text{NBLT}}) = \sum_{j=1}^{p} \frac{(\theta_j - d)^2}{\theta_j(\theta_j + k)^2} + (d + k)^2 \sum_{j=1}^{p} \frac{\alpha_j^2}{(\theta_j + k)^2}. \tag{21}$$

Kandemir Çetinkaya and Kaçıranlar [27] extended this line of work by applying the Özkale-Kaçıranlar estimator [21] to the negative binomial setting, resulting in the negative binomial Özkale-Kaçıranlar (NBOK) estimator:

$$\hat{\beta}_{\text{NBOK}} = (P + kI)^{-1}(P + kdI)\hat{\beta}_{\text{NBML}}, \quad k > 0,\ 0 < d < 1, \tag{22}$$

where k and d denote shrinkage parameters. The estimator is quite flexible: it reduces to the NBL as $k \to 1$, to the NBR as $d \to 0$, and to the NBML when both parameters equal zero simultaneously. The MMSE and SMSE associated with the NBOK estimator are expressed below:

$$\text{MMSE}(\hat{\beta}_{\text{NBOK}}) = V\Theta_k^{-1}\Theta_{kd}\Theta^{-1}\Theta_{kd}\Theta_k^{-1}V^\top + k^2(d - 1)^2 V\Theta_k^{-1}\alpha\alpha^\top\Theta_k^{-1}V^\top, \tag{23}$$

$$\text{SMSE}(\hat{\beta}_{\text{NBOK}}) = \sum_{j=1}^{p} \frac{(\theta_j + kd)^2}{\theta_j(\theta_j + k)^2} + k^2(d - 1)^2 \sum_{j=1}^{p} \frac{\alpha_j^2}{(\theta_j + k)^2}, \tag{24}$$

where $\Theta_{kd} = \text{diag}(\theta_1 + kd, \theta_2 + kd, ..., \theta_p + kd)$.

3. Proposed Estimator

This section introduces a new biased estimator for the NBRM under multicollinearity, along with its main statistical properties. We compare our estimator's theoretical performance and efficiency against several existing estimators in the presence of multicollinearity and present methods for selecting the biasing parameter for both the proposed estimator and its competitors.

3.1. Negative binomial Adjusted Liu-Ridge estimator

Existing biased estimators for the NBRM each improve on ML estimation, but each is also limited by its own shrinkage structure, which may not handle every degree of multicollinearity severity. This study introduces a new estimator, the negative binomial adjusted Liu-Ridge (NBALR) estimator, combining the penalization strategies of the ridge and adjusted Liu estimators into a single framework. The ridge component reduces coefficient variance by perturbing the information matrix; the adjusted Liu component allows finer control over the shrinkage magnitude through a linear parameter. Combining these two mechanisms gives the NBALR estimator greater flexibility and produces more stable estimates under moderate to severe multicollinearity. The NBALR estimator is defined as follows:

$$\begin{aligned}\hat{\beta}_{\text{NBALR}} &= (P+I)^{-1}(P-dI)\hat{\beta}_{\text{NBR}} \\ &= (P+I)^{-1}(P-dI)(P+kI)^{-1}P\hat{\beta}_{\text{NBML}},\end{aligned} \tag{25}$$

where $k>0$ and d are the ridge and adjusted Liu parameters, respectively. The estimator reduces to the NBAL estimator when $k \to 0$, collapses to the NBR estimator when $d \to -1$, and coincides with the NBML estimator in the limiting case where both parameters vanish.

Using the unbiasedness of the maximum likelihood estimator, i.e., $\mathrm{E}(\hat{\beta}_{\text{NBML}}) = \beta$ and $\mathrm{Cov}(\hat{\beta}_{\text{NBML}}) = P^{-1}$, the bias vector of the NBALR estimator is obtained as

$$\begin{aligned}\mathrm{Bias}(\hat{\beta}_{\text{NBALR}}) &= \mathrm{E}(\hat{\beta}_{\text{NBALR}}) - \beta \\ &= (P+I)^{-1}(P-dI)(P+kI)^{-1}P\beta \\ &= -V\big[(d+k+1)\Theta + kI\big]\Theta_1^{-1}\Theta_k^{-1}\alpha.\end{aligned} \tag{26}$$

The covariance matrix of the NBALR estimator follows directly from the linear transformation of $\hat{\beta}_{\text{NBML}}$:

$$\begin{aligned}\mathrm{Cov}(\hat{\beta}_{\text{NBALR}}) &= \mathrm{E}\left(\Big[\hat{\beta}_{\text{NBALR}} - E(\hat{\beta}_{\text{NBALR}})\Big]\Big[\hat{\beta}_{\text{NBALR}} - E(\hat{\beta}_{\text{NBALR}})\Big]^{\top}\right) \\ &= V\Theta\Theta_d\Theta_1^{-1}\Theta_k^{-1}\Theta^{-1}\Theta\Theta_d\Theta_1^{-1}\Theta_k^{-1}V^{\top}.\end{aligned} \tag{27}$$

Thus, the MMSE and the SMSE of the NBALR estimator are obtained using the covariance matrix and bias vector as follows:

$$\begin{aligned}\mathrm{MMSE}(\hat{\beta}_{\text{NBALR}}) &= \mathrm{Cov}(\hat{\beta}_{\text{NBALR}}) + \mathrm{Bias}(\hat{\beta}_{\text{NBALR}})\,\mathrm{Bias}(\hat{\beta}_{\text{NBALR}})^{\top} \\ &= V\Theta\Theta_d\Theta_1^{-1}\Theta_k^{-1}\Theta_d\Theta_1^{-1}\Theta_k^{-1}V^{\top} \\ &\quad + V\big[(d+k+1)\Theta + kI\big]\Theta_1^{-1}\Theta_k^{-1}\alpha\alpha^{\top}\Theta_k^{-1}\Theta_1^{-1}\big[(d+k+1)\Theta + kI\big]V^{\top} \\ &= V\Theta\Theta_d\Theta_1^{-1}\Theta_k^{-1}\Theta_d\Theta_1^{-1}\Theta_k^{-1}V^{\top} + V\big[(d+k+1)\Theta + kI\big]^2\Theta_1^{-1}\Theta_k^{-1}\alpha\alpha^{\top}\Theta_k^{-1}\Theta_1^{-1}V^{\top},\end{aligned}$$

which can be compactly written as

$$\mathrm{MMSE}(\hat{\beta}_{\text{NBALR}}) = V\left[\Theta_A\Theta^{-1}\Theta_A + (\Theta_A - I)\alpha\alpha^{\top}(\Theta_A - I)\right]V^{\top}. \tag{28}$$

where $\Theta_A = \Theta\Theta_d\Theta_1^{-1}\Theta_k^{-1}$.

Finally, the SMSE is obtained by taking the trace of the MMSE:

$$\begin{aligned}\mathrm{SMSE}(\hat{\beta}_{\text{NBALR}}) &= \mathrm{tr}\big(\mathrm{MMSE}(\hat{\beta}_{\text{NBALR}})\big) \\ &= \sum_{j=1}^{p}\frac{\theta_j(\theta_j - d)^2}{(\theta_j+1)^2(\theta_j+k)^2} + \sum_{j=1}^{p}\frac{[\theta_j(d+k+1)+k]^2}{(\theta_j+1)^2(\theta_j+k)^2}\alpha_j^2.\end{aligned} \tag{29}$$

The derived MSE separates into a variance component, which decreases with k, and a squared-bias component, governed by both k and d. This dual-penalization structure enhances adaptability, enabling the NBALR estimator to achieve a superior bias-variance trade-off compared to existing estimators.

3.2. *The superiority of the proposed estimator*

Lemma 3.1
Let B be a positive definite matrix, c a positive scalar, and γ a non-zero vector. Then the matrix $cB - \gamma\gamma^\top$ is positive definite if and only if $\gamma^\top B^{-1}\gamma < c$ [33].

Theorem 3.1
The proposed estimator $\hat{\beta}_{\text{NBALR}}$ dominates $\hat{\beta}_{\text{NBML}}$ in the MMSE sense, that is, $\text{MMSE}(\hat{\beta}_{\text{NBML}}) - \text{MMSE}(\hat{\beta}_{\text{NBALR}}) > 0$, if and only if $L^\top \left(\Theta^{-1} - \Theta_A\Theta^{-1}\Theta_A\right)^{-1} L < 1$, where $L = \text{Bias}(\hat{\beta}_{\text{NBALR}})$ and $k > 0$.

Proof
The dominance condition is established by evaluating the discrepancy between the MMSE of the NBML estimator and that of the proposed estimator. Substituting the expressions from Eq. (8) and Eq. (28), this difference is given by:

$$\text{MMSE}(\hat{\beta}_{\text{NBML}}) - \text{MMSE}(\hat{\beta}_{\text{NBALR}}) = V\left(\Theta^{-1} - \Theta_A\Theta^{-1}\Theta_A\right)V^\top - LL^\top. \tag{30}$$

The corresponding scalar form is:

$$\text{SMSE}(\hat{\beta}_{\text{NBML}}) - \text{SMSE}(\hat{\beta}_{\text{NBALR}}) = V\,\text{diag}\left(\frac{1}{\theta_j} - \frac{\theta_j(\theta_j - d)^2}{(\theta_j + 1)^2(\theta_j + k)^2}\right)V^\top - LL^\top. \tag{31}$$

A necessary condition for the positive definiteness of the matrix difference in Eq. (30) is that the diagonal matrix $D = \Theta^{-1} - \Theta_A\Theta^{-1}\Theta_A$ is itself positive definite. This requires each diagonal element to satisfy:

$$\frac{1}{\theta_j} - \frac{\theta_j(\theta_j - d)^2}{(\theta_j + 1)^2(\theta_j + k)^2} > 0,$$

which simplifies to:

$$(\theta_j + 1)^2(\theta_j + k)^2 > \theta_j^2(\theta_j - d)^2, \quad \forall j.$$

Under the constraint $k > 0$, this inequality is satisfied for all j, confirming that D is positive definite. The proof is then completed by invoking Lemma 3.1 with $B = D$, $c = 1$, and $\gamma = L$, from which the required condition $L^\top D^{-1}L < 1$ follows directly. □

Theorem 3.2
The proposed estimator $\hat{\beta}_{\text{NBALR}}$ dominates $\hat{\beta}_{\text{NBR}}$ in the MMSE sense, that is, $\text{MMSE}(\hat{\beta}_{\text{NBR}}) - \text{MMSE}(\hat{\beta}_{\text{NBALR}}) > 0$, if and only if $\omega^\top D_1^{-1}\omega < 1$, where $L = \text{Bias}(\hat{\beta}_{\text{NBALR}})$, $L_1 = \text{Bias}(\hat{\beta}_{\text{NBR}})$, $\omega = L - L_1$, $D_1 = \Theta_k^{-1}\Theta\Theta_k^{-1} - \Theta_A\Theta^{-1}\Theta_A$, $0 < d < 1$, and $k > 0$.

Proof
The dominance condition is established by evaluating the discrepancy between the MMSE of the NBR estimator and that of the proposed estimator. Substituting the expressions from Eq. (11) and Eq. (28), this difference is given by:

$$\text{MMSE}(\hat{\beta}_{\text{NBR}}) - \text{MMSE}(\hat{\beta}_{\text{NBALR}}) = V\left(\Theta_k^{-1}\Theta\Theta_k^{-1} - \Theta_A\Theta^{-1}\Theta_A\right)V^\top + L_1L_1^\top - LL^\top. \tag{32}$$

The corresponding scalar form is:

$$\text{SMSE}(\hat{\beta}_{\text{NBR}}) - \text{SMSE}(\hat{\beta}_{\text{NBALR}}) = V\,\text{diag}\left(\frac{\theta_j}{(\theta_j + k)^2} - \frac{\theta_j(\theta_j - d)^2}{(\theta_j + 1)^2(\theta_j + k)^2}\right)V^\top + L_1L_1^\top - LL^\top. \tag{33}$$

A necessary condition for the positive definiteness of the matrix difference in Eq. (32) is that the diagonal matrix $D_1 = \Theta_k^{-1}\Theta\Theta_k^{-1} - \Theta_A\Theta^{-1}\Theta_A$ is positive definite. This requires each diagonal element to satisfy:

$$\frac{\theta_j}{(\theta_j + k)^2} - \frac{\theta_j(\theta_j - d)^2}{(\theta_j + 1)^2(\theta_j + k)^2} > 0,$$

which simplifies to:

$$\theta_j(\theta_j+1)^2 > \theta_j(\theta_j-d)^2, \quad \forall\, j.$$

Under the constraint $k>0$ and $0<d<1$, this inequality is satisfied for all j, confirming that D_1 is positive definite. The remaining term $L_1L_1^\top - LL^\top$ can be rewritten as $\omega\omega^\top + 2\omega L_1^\top$, where $\omega = L - L_1$. The proof is then completed by invoking Lemma 3.1 with $B = D_1$, $c=1$, and $\gamma=\omega$, from which the required condition $\omega^\top D_1^{-1}\omega < 1$ follows directly, ensuring the overall matrix difference in Eq. (32) is positive definite. □

Theorem 3.3
The proposed estimator $\hat{\beta}_{\text{NBALR}}$ dominates $\hat{\beta}_{\text{NBL}}$ in the MMSE sense, that is, $\text{MMSE}(\hat{\beta}_{\text{NBL}}) - \text{MMSE}(\hat{\beta}_{\text{NBALR}}) > 0$, if and only if $\omega^\top D_2^{-1}\omega < 1$, where $L = \text{Bias}(\hat{\beta}_{\text{NBALR}})$, $L_2 = \text{Bias}(\hat{\beta}_{\text{NBL}})$, $\omega = L - L_2$, $D_2 = \Theta_1^{-1}\Theta_d\Theta^{-1}\Theta_d\Theta_1^{-1} - \Theta_A\Theta^{-1}\Theta_A$, $0<d<1$, and $k>0$.

Proof
The dominance condition is established by evaluating the discrepancy between the MMSE of the NBL estimator and that of the proposed estimator. Substituting the expressions from Eq. (14) and Eq. (28), this difference is given by:

$$\text{MMSE}(\hat{\beta}_{\text{NBL}}) - \text{MMSE}(\hat{\beta}_{\text{NBALR}}) = V\left(\Theta_1^{-1}\Theta_d\Theta^{-1}\Theta_d\Theta_1^{-1} - \Theta_A\Theta^{-1}\Theta_A\right)V^\top + L_2L_2^\top - LL^\top. \tag{34}$$

The corresponding scalar form is:

$$\text{SMSE}(\hat{\beta}_{\text{NBL}}) - \text{SMSE}(\hat{\beta}_{\text{NBALR}}) = V\,\text{diag}\left(\frac{(\theta_j+d)^2}{\theta_j(\theta_j+1)^2} - \frac{\theta_j(\theta_j-d)^2}{(\theta_j+1)^2(\theta_j+k)^2}\right)V^\top + L_2L_2^\top - LL^\top. \tag{35}$$

A necessary condition for the positive definiteness of the matrix difference in Eq. (34) is that the diagonal matrix $D_2 = \Theta_1^{-1}\Theta_d\Theta^{-1}\Theta_d\Theta_1^{-1} - \Theta_A\Theta^{-1}\Theta_A$, is positive definite. This requires each diagonal element to satisfy:

$$\frac{(\theta_j+d)^2}{\theta_j(\theta_j+1)^2} - \frac{\theta_j(\theta_j-d)^2}{(\theta_j+1)^2(\theta_j+k)^2} > 0,$$

which simplifies to:

$$(\theta_j+d)^2(\theta_j+k)^2 > \theta_j^2(\theta_j-d)^2, \quad \forall\, j.$$

Under the constraint $k>0$ and $0<d<1$, this inequality is satisfied for all j, confirming that D_2 is positive definite. The remaining term $L_2L_2^\top - LL^\top$ can be rewritten as $\omega\omega^\top + 2\omega L_2^\top$, where $\omega = L - L_2$. The proof is then completed by invoking Lemma 3.1 with $B = D_2$, $c=1$, and $\gamma=\omega$, from which the required condition $\omega^\top D_2^{-1}\omega < 1$ follows directly, ensuring the overall matrix difference in Eq. (34) is positive definite. □

Theorem 3.4
The proposed estimator $\hat{\beta}_{\text{NBALR}}$ dominates $\hat{\beta}_{\text{NBAL}}$ in the MMSE sense, that is, $\text{MMSE}(\hat{\beta}_{\text{NBAL}}) - \text{MMSE}(\hat{\beta}_{\text{NBALR}}) > 0$, if and only if $\omega^\top D_3^{-1}\omega < 1$, where $L = \text{Bias}(\hat{\beta}_{\text{NBALR}})$, $L_3 = \text{Bias}(\hat{\beta}_{\text{NBAL}})$, $\omega = L - L_3$, $D_3 = \Theta_1^{-1}\Theta_{-d}\Theta^{-1}\Theta_{-d}\Theta_1^{-1} - \Theta_A\Theta^{-1}\Theta_A$, $0<d<1$, and $k>0$.

Proof
The dominance condition is established by evaluating the discrepancy between the MMSE of the NBAL estimator and that of the proposed estimator. Substituting the expressions from Eq. (17) and Eq. (28), this difference is given by:

$$\text{MMSE}(\hat{\beta}_{\text{NBAL}}) - \text{MMSE}(\hat{\beta}_{\text{NBALR}}) = V\left(\Theta_1^{-1}\Theta_{-d}\Theta^{-1}\Theta_{-d}\Theta_1^{-1} - \Theta_A\Theta^{-1}\Theta_A\right)V^\top + L_3L_3^\top - LL^\top. \tag{36}$$

The corresponding scalar form is:

$$\text{SMSE}(\hat{\beta}_{\text{NBAL}}) - \text{SMSE}(\hat{\beta}_{\text{NBALR}}) = V\,\text{diag}\left(\frac{(\theta_j-d)^2}{\theta_j(\theta_j+1)^2} - \frac{\theta_j(\theta_j-d)^2}{(\theta_j+1)^2(\theta_j+k)^2}\right)V^\top + L_3L_3^\top - LL^\top. \tag{37}$$

A necessary condition for the positive definiteness of the matrix difference in Eq. (36) is that the diagonal matrix $D_3 = \Theta_1^{-1}\Theta_{-d}\Theta^{-1}\Theta_{-d}\Theta_1^{-1} - \Theta_A\Theta^{-1}\Theta_A$, is positive definite. This requires each diagonal element to satisfy:

$$\frac{(\theta_j - d)^2}{\theta_j(\theta_j+1)^2} - \frac{\theta_j(\theta_j - d)^2}{(\theta_j+1)^2(\theta_j+k)^2} > 0,$$

which simplifies to:

$$(\theta_j - d)^2(\theta_j+k)^2 > \theta_j^2(\theta_j - d)^2, \quad \forall j.$$

Under the constraint $k > 0$ and $0 < d < 1$, this inequality is satisfied for all j, confirming that D_3 is positive definite. The remaining term $L_3L_3^\top - LL^\top$ can be rewritten as $\omega\omega^\top + 2\omega L_3^\top$, where $\omega = L - L_3$. The proof is then completed by invoking Lemma 3.1 with $B = D_3$, $c = 1$, and $\gamma = \omega$, from which the required condition $\omega^\top D_3^{-1}\omega < 1$ follows directly, ensuring the overall matrix difference in Eq. (36) is positive definite. □

Theorem 3.5
The proposed estimator $\hat{\beta}_{\text{NBALR}}$ dominates $\hat{\beta}_{\text{NBLT}}$ in the MMSE sense, that is, $\text{MMSE}(\hat{\beta}_{\text{NBLT}}) - \text{MMSE}(\hat{\beta}_{\text{NBALR}}) > 0$, if and only if $\omega^\top D_4^{-1}\omega < 1$, where $L = \text{Bias}(\hat{\beta}_{\text{NBALR}})$, $L_4 = \text{Bias}(\hat{\beta}_{\text{NBLT}})$, $\omega = L - L_4$, $D_4 = \Theta_k^{-1}\Theta_{-d}\Theta^{-1}\Theta_{-d}\Theta_k^{-1} - \Theta_A\Theta^{-1}\Theta_A$, $0 < d < 1$, and $k > 0$.

Proof
The dominance condition is established by evaluating the discrepancy between the MMSE of the NBLT estimator and that of the proposed estimator. Substituting the expressions from Eq. (20) and Eq. (28), this difference is given by:

$$\text{MMSE}(\hat{\beta}_{\text{NBLT}}) - \text{MMSE}(\hat{\beta}_{\text{NBALR}}) = V\left(\Theta_k^{-1}\Theta_{-d}\Theta^{-1}\Theta_{-d}\Theta_k^{-1} - \Theta_A\Theta^{-1}\Theta_A\right)V^\top + L_4L_4^\top - LL^\top. \tag{38}$$

The corresponding scalar form is:

$$\text{SMSE}(\hat{\beta}_{\text{NBLT}}) - \text{SMSE}(\hat{\beta}_{\text{NBALR}}) = V\,\text{diag}\left(\frac{(\theta_j - d)^2}{\theta_j(\theta_j+k)^2} - \frac{\theta_j(\theta_j - d)^2}{(\theta_j+1)^2(\theta_j+k)^2}\right)V^\top + L_4L_4^\top - LL^\top. \tag{39}$$

A necessary condition for the positive definiteness of the matrix difference in Eq. (38) is that the diagonal matrix $D_4 = \Theta_k^{-1}\Theta_{-d}\Theta^{-1}\Theta_{-d}\Theta_k^{-1} - \Theta_A\Theta^{-1}\Theta_A$, is positive definite. This requires each diagonal element to satisfy:

$$\frac{(\theta_j - d)^2}{\theta_j(\theta_j+k)^2} - \frac{\theta_j(\theta_j - d)^2}{(\theta_j+1)^2(\theta_j+k)^2} > 0,$$

which simplifies to:

$$(\theta_j - d)^2(\theta_j+1)^2 > \theta_j^2(\theta_j - d)^2, \quad \forall j.$$

Under the constraint $k > 0$ and $0 < d < 1$, this inequality is satisfied for all j, confirming that D_4 is positive definite. The remaining term $L_4L_4^\top - LL^\top$ can be rewritten as $\omega\omega^\top + 2\omega L_4^\top$, where $\omega = L - L_4$. The proof is then completed by invoking Lemma 3.1 with $B = D_4$, $c = 1$, and $\gamma = \omega$, from which the required condition $\omega^\top D_4^{-1}\omega < 1$ follows directly, ensuring the overall matrix difference in Eq. (38) is positive definite. □

Theorem 3.6
The proposed estimator $\hat{\beta}_{\text{NBALR}}$ dominates $\hat{\beta}_{\text{NBOK}}$ in the MMSE sense, that is, $\text{MMSE}(\hat{\beta}_{\text{NBOK}}) - \text{MMSE}(\hat{\beta}_{\text{NBALR}}) > 0$, if and only if $\omega^\top D_5^{-1}\omega < 1$, where $L = \text{Bias}(\hat{\beta}_{\text{NBALR}})$, $L_5 = \text{Bias}(\hat{\beta}_{\text{NBOK}})$, $\omega = L - L_5$, $D_5 = \Theta_k^{-1}\Theta_{kd}\Theta^{-1}\Theta_{kd}\Theta_k^{-1} - \Theta_A\Theta^{-1}\Theta_A$, $0 < d < 1$, and $k > 0$.

Proof
The dominance condition is established by evaluating the discrepancy between the MMSE of the NBOK estimator and that of the proposed estimator. Substituting the expressions from Eq. (23) and Eq. (28), this difference is given

by:

$$\mathrm{MMSE}(\hat{\beta}_{\mathrm{NBOK}}) - \mathrm{MMSE}(\hat{\beta}_{\mathrm{NBALR}}) = V\left(\Theta_k^{-1}\Theta_{kd}\Theta^{-1}\Theta_{kd}\Theta_k^{-1} - \Theta_A\Theta^{-1}\Theta_A\right)V^\top + L_5L_5^\top - LL^\top. \quad (40)$$

The corresponding scalar form is:

$$\mathrm{SMSE}(\hat{\beta}_{\mathrm{NBOK}}) - \mathrm{SMSE}(\hat{\beta}_{\mathrm{NBALR}}) = V\,\mathrm{diag}\left(\frac{(\theta_j + kd)^2}{\theta_j(\theta_j + k)^2} - \frac{\theta_j(\theta_j - d)^2}{(\theta_j + 1)^2(\theta_j + k)^2}\right)V^\top + L_5L_5^\top - LL^\top. \quad (41)$$

A necessary condition for the positive definiteness of the matrix difference in Eq. (40) is that the diagonal matrix $D_5 = \Theta_k^{-1}\Theta_{kd}\Theta^{-1}\Theta_{kd}\Theta_k^{-1} - \Theta_A\Theta^{-1}\Theta_A$, is positive definite. This requires each diagonal element to satisfy:

$$\frac{(\theta_j + kd)^2}{\theta_j(\theta_j + k)^2} - \frac{\theta_j(\theta_j - d)^2}{(\theta_j + 1)^2(\theta_j + k)^2} > 0,$$

which simplifies to:

$$(\theta_j + kd)^2(\theta_j + 1)^2 > \theta_j^2(\theta_j - d)^2, \quad \forall j.$$

Under the constraint $k > 0$ and $0 < d < 1$, this inequality is satisfied for all j, confirming that D_5 is positive definite. The remaining term $L_5L_5^\top - LL^\top$ can be rewritten as $\omega\omega^\top + 2\omega L_5^\top$, where $\omega = L - L_5$. The proof is then completed by invoking Lemma 3.1 with $B = D_5$, $c = 1$, and $\gamma = \omega$, from which the required condition $\omega^\top D_5^{-1}\omega < 1$ follows directly, ensuring the overall matrix difference in Eq. (40) is positive definite. □

3.3. Selection of biasing parameters

The NBALR estimator's performance depends heavily on how the biasing parameters k and d are chosen. Unlike the ML estimator, which is uniquely determined, NBALR is a family of estimators indexed by these shrinkage parameters. The choice of k and d therefore needs to follow a well-defined optimality criterion. In this study, we adopt the SMSE as the measure of estimation accuracy, and we seek the pair (k, d) that minimizes this quantity.

Recalling the SMSE function derived in Eq. (29), to determine the optimal values of the shrinkage parameters k and d, we minimize the SMSE of the proposed estimator with respect to each parameter in turn.

Differentiating $\mathrm{SMSE}(\hat{\beta}_{\mathrm{NBALR}})$ with respect to k and setting the result equal to zero gives:

$$\frac{\partial\,\mathrm{SMSE}}{\partial k} = -2\sum_{j=1}^{p}\frac{\theta_j(\theta_j - d)^2}{(\theta_j + 1)^2(\theta_j + k)^3} + 2\sum_{j=1}^{p}\frac{\alpha_j^2\,\theta_j(\theta_j - d)\left[\theta_j(d + k + 1) + k\right]}{(\theta_j + 1)^2(\theta_j + k)^3} = 0. \quad (42)$$

Differentiating $\mathrm{SMSE}(\hat{\beta}_{\mathrm{NBALR}})$ with respect to d and setting the result equal to zero gives:

$$\frac{\partial\,\mathrm{SMSE}}{\partial d} = -2\sum_{j=1}^{p}\frac{\theta_j(\theta_j - d)}{(\theta_j + 1)^2(\theta_j + k)^2} + 2\sum_{j=1}^{p}\frac{\alpha_j^2\,\theta_j\left[\theta_j(d + k + 1) + k\right]}{(\theta_j + 1)^2(\theta_j + k)^2} = 0. \quad (43)$$

Finally, we get the values of the shrinkage parameters k and d, which are given by:

$$k = \frac{1}{p}\sum_{j=1}^{p}\frac{\theta_j - d - \alpha_j^2\,\theta_j(d + 1)}{\alpha_j^2(\theta_j + 1)}, \qquad d = \frac{1}{p}\sum_{j=1}^{p}\frac{\theta_j(1 - \alpha_j^2) - \alpha_j^2\,k(\theta_j + 1)}{\alpha_j^2\,\theta_j + 1}.$$

In the present study, the estimation of the shrinkage parameters k and d proceeds through a structured sequential procedure. As an initial value for k, defined as:

$$\hat{k} = \min\left(\sqrt{\frac{1}{\hat{\alpha}_j^2}}\right).$$

This initial value $\hat{k}$ is then used to calculate the value of d as follows:

$$\hat{d} = \frac{1}{p} \min \left(\frac{\theta_j \left(1 - \hat{\alpha}_j^2\right) - \hat{\alpha}_j^2 \, \hat{k} \, (\theta_j + 1)}{\hat{\alpha}_j^2 \, \theta_j + 1} \right).$$

The resulting estimate $\hat{d}$ is subsequently used to compute the updated estimate of k as follows:

$$\hat{k} = \frac{1}{p} \min \left(\frac{\theta_j - \hat{d} - \hat{\alpha}_j^2 \, \theta_j \left(\hat{d} + 1\right)}{\hat{\alpha}_j^2 \, (\theta_j + 1)} \right).$$

4. Monte Carlo Simulation Study

A Monte Carlo simulation experiment is carried out to examine the finite-sample behavior of the proposed estimator against a set of established competitors under controlled conditions. The response variable is drawn from a negative binomial distribution, NB$(\mu_i, \mu_i + \vartheta\mu_i^2)$, where the mean μ_i is related to the explanatory variables through a log-linear model:

$$\mu_i = \exp(\beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip}), \quad i = 1, 2, \ldots, n. \tag{44}$$

In line with common practice in the simulation literature [24, 26], the regression coefficients are normalized so that $\sum_{j=1}^{p} \beta_j^2 = 1$, with the equality constraint $\beta_0 = \beta_1 = \cdots = \beta_{p-1}$ imposed to maintain consistency across all experimental configurations.

Following Hammad et al. [22] and Hussam et al. [23], the multicollinearity among the predictors is induced through the following structure:

$$x_{ij} = (1 - \rho^2)^{\frac{1}{2}} \, e_{ij} + \rho \, e_{ip}, \quad i = 1, 2, \ldots, n, \quad j = 1, 2, \ldots, p, \tag{45}$$

where e_{ij} are independently generated from a standard normal distribution, and the parameter ρ governs the strength of the linear dependence among the regressors, with larger values of ρ corresponding to more severe multicollinearity.

Estimator performance is evaluated using the simulated mean squared error (MSE) and simulated mean absolute error (MAE), which are defined as

$$\text{Simulated MSE} = \frac{1}{q} \sum_{i=1}^{q} (\hat{\beta}_i - \beta)'(\hat{\beta}_i - \beta), \tag{46}$$

$$\text{Simulated MAE}(\hat{\beta}) = \frac{1}{q} \sum_{i=1}^{q} \sum_{j=1}^{p} |\hat{\beta}_{ij} - \beta_j|, \tag{47}$$

with $\hat{\beta}_i$ denoting the coefficient vector obtained at replication i, β the true parameter vector, and q the total number of Monte Carlo replications. This criterion accounts for both the variance and the bias of each estimator, making it well suited for comparing estimation accuracy across methods. The simulation design covers conditions likely to affect the estimators' relative performance, including sample size, number of predictors, correlation level, and dispersion. Table 1 summarizes the full set of design factors and their levels.

The shrinkage parameters for each competing estimator are selected based on well-established expressions from the existing literature. For the NBR [24], the ridge parameter is estimated as $\hat{k} = 1/\sum_{j=1}^{p} \hat{\alpha}_j^2$. The Liu parameter

Table 1. Levels of the design factors employed in the Monte Carlo simulation.

Factor	**Symbol**	**Levels**
Number of replicates	q	2000
Dispersion parameter	ϑ	1, 2
Number of predictors	p	4, 7, 10, 15
Sample size	n	50, 100, 150, 200, 300, 400
Correlation level	ρ	0.80, 0.85, 0.90, 0.95, 0.99

Table 2. Performance comparison of estimators in the NBRM based on simulated MSE and MAE when $p = 4$ and $\vartheta = 1$.

		NBML	NBR	NBL	NBAL	NBLT	NBOK	NBALR	NBML	NBR	NBL	NBAL	NBLT	NBOK	NBALR
ρ	n	-	$\hat{k}$	$\hat{d}$	$\hat{d}$	$\hat{k}, \hat{d}$	$\hat{k}, \hat{d}$	$\hat{k}, \hat{d}$	-	$\hat{k}$	$\hat{d}$	$\hat{d}$	$\hat{k}, \hat{d}$	$\hat{k}, \hat{d}$	$\hat{k}, \hat{d}$
		MSE							MAE						
0.80	50	0.82403	0.68481	0.70505	0.70372	0.58804	0.57202	0.26189	1.56884	1.42469	1.45122	1.45378	1.32029	1.29852	0.85002
	100	0.34705	0.30759	0.32282	0.32232	0.28709	0.27284	0.17173	1.03581	0.97446	0.99733	0.99795	0.94609	0.91745	0.73197
	150	0.21151	0.19468	0.20286	0.20245	0.18840	0.17905	0.13569	0.80382	0.77113	0.78613	0.78606	0.76078	0.73987	0.64701
	200	0.15367	0.14402	0.14894	0.14874	0.14144	0.13492	0.11105	0.68633	0.66473	0.67512	0.67512	0.65977	0.64380	0.58605
	300	0.09422	0.09034	0.09234	0.09232	0.08972	0.08660	0.07730	0.53805	0.52708	0.53246	0.53257	0.52559	0.51631	0.48869
	400	0.07866	0.07566	0.07739	0.07731	0.07520	0.07273	0.06564	0.49176	0.48260	0.48754	0.48751	0.48138	0.47353	0.45068
0.85	50	0.72428	0.60899	0.62919	0.62867	0.52948	0.51505	0.24084	1.46755	1.33875	1.36453	1.36752	1.25063	1.22645	0.81571
	100	0.47337	0.40799	0.42890	0.42840	0.36629	0.35143	0.18498	1.19135	1.10388	1.13195	1.13335	1.05309	1.02404	0.74409
	150	0.26743	0.24171	0.25255	0.25230	0.23065	0.21841	0.15229	0.89905	0.85425	0.87253	0.87306	0.83759	0.81198	0.68139
	200	0.18797	0.17326	0.18055	0.18029	0.16855	0.15946	0.12306	0.76027	0.73023	0.74440	0.74447	0.72185	0.70113	0.61868
	300	0.12963	0.12228	0.12614	0.12600	0.12064	0.11530	0.09740	0.62466	0.60703	0.61578	0.61577	0.60362	0.58986	0.54373
	400	0.09905	0.09431	0.09695	0.09683	0.09343	0.08973	0.07837	0.54623	0.53335	0.54009	0.54000	0.53129	0.52072	0.48777
0.90	50	1.52615	1.19506	1.19389	1.18405	0.95427	0.93648	0.33800	2.11278	1.84612	1.85615	1.85947	1.62824	1.61867	0.86819
	100	0.64399	0.53962	0.56356	0.56290	0.46717	0.45367	0.20001	1.38394	1.26157	1.29015	1.29295	1.18033	1.15390	0.75473
	150	0.41335	0.35799	0.38011	0.37800	0.32469	0.30968	0.16799	1.09663	1.01895	1.04950	1.04874	0.97690	0.94702	0.70120
	200	0.31421	0.27839	0.29468	0.29317	0.26087	0.24626	0.15456	0.96199	0.90488	0.92977	0.92888	0.88032	0.85120	0.67935
	300	0.18115	0.16662	0.17388	0.17357	0.16207	0.15308	0.11715	0.73107	0.70131	0.71537	0.71544	0.69337	0.67267	0.59135
	400	0.13644	0.12778	0.13286	0.13247	0.12555	0.11949	0.09841	0.63585	0.61561	0.62676	0.62635	0.61115	0.59579	0.54224
0.95	50	3.15221	2.28508	2.08763	2.02977	1.85672	1.69253	0.65444	2.95154	2.44866	2.36087	2.36277	2.13447	2.06167	1.13224
	100	1.16745	0.91991	0.94108	0.93612	0.74350	0.72533	0.24852	1.83284	1.60929	1.63630	1.63929	1.43331	1.41737	0.70748
	150	0.75217	0.60742	0.64117	0.63776	0.49939	0.49039	0.16504	1.47933	1.32149	1.35875	1.35996	1.20455	1.18363	0.65494
	200	0.52438	0.44057	0.46604	0.46463	0.38535	0.36943	0.16385	1.22005	1.11417	1.14629	1.14777	1.04987	1.01823	0.67346
	300	0.34982	0.30615	0.32436	0.32315	0.28311	0.26716	0.15511	1.00059	0.93481	0.96117	0.96110	0.90456	0.87326	0.67076
	400	0.28747	0.25489	0.27048	0.26904	0.24024	0.22545	0.14393	0.91437	0.86051	0.88525	0.88432	0.83944	0.80943	0.65183
0.99	50	11.59989	7.97699	5.48232	5.31840	6.77989	5.65652	2.15688	5.57149	4.40096	3.62019	3.67814	3.93899	3.55901	1.81224
	100	6.18524	4.24728	3.30669	3.22103	3.50017	2.98709	1.04968	4.05992	3.22091	2.85354	2.88499	2.80876	2.60441	1.28918
	150	3.64632	2.58466	2.29473	2.23927	2.06944	1.84315	0.62466	3.13285	2.54543	2.41660	2.42432	2.18485	2.09250	1.01649
	200	3.01363	2.13252	2.01391	1.94444	1.67702	1.51971	0.48559	2.82410	2.30075	2.25532	2.24736	1.95805	1.89255	0.90072
	300	2.06543	1.48152	1.48089	1.43992	1.13284	1.05987	0.31416	2.34178	1.92975	1.94216	1.93799	1.62834	1.59833	0.72742
	400	1.38998	1.04317	1.08558	1.06482	0.79909	0.77888	0.21220	1.94121	1.65318	1.69451	1.69164	1.42404	1.41021	0.58958

for the NBL [25] is chosen as $\hat{d} = \max\left(0, \min\left((\hat{\alpha}_j^2 - 1)/(1/\theta_j + \hat{\alpha}_j^2)\right)\right)$. For the NBAL [15], the adjusted Liu parameter is estimated as $\hat{d} = \max\left(0, \min\left((\theta_j(1 - \hat{\alpha}_j^2))/(1 + \theta_j \hat{\alpha}_j^2)\right)\right)$. For the NBLT [26], both parameters are jointly determined as $\hat{k} = \min\left((\theta_j - d(1 - \theta_j \hat{\alpha}_j^2))/(\theta_j \hat{\alpha}_j^2)\right)$ and $\hat{d} = \max\left(0, \min\left(\hat{\alpha}_j^2/(1/\theta_j + \hat{\alpha}_j^2)\right)\right)$. Finally, for the NBOK [27], the parameters are given by $\hat{k} = p\hat{\sigma}^2 / \sum_{j=1}^{p}(\hat{\alpha}_j^2 - d(\hat{\alpha}_j^2 + \hat{\sigma}^2/\theta_j))$ and $\hat{d} = \sum_{j=1}^{p}((k\hat{\alpha}_j^2 - \hat{\sigma}^2)/(\theta_j + k)^2) / \sum_{j=1}^{p}(k(\hat{\alpha}_j^2 \theta_j + \hat{\sigma}^2)/\theta_j(\theta_j + k)^2)$.

Tables 2–9 and Figures 1, 2 reveal several consistent patterns across the simulation settings. Most notably, the proposed NBALR estimator achieves the lowest MSE and MAE across nearly all configurations of sample size, number of predictors, correlation level, and dispersion parameter, supporting the theoretical properties established in the previous section.

We consider the effects of sample size, multicollinearity severity, number of predictors, and dispersion parameter in turn. Sample size (n) shows the expected pattern: MSE and MAE decrease monotonically from $n = 30$ to $n = 400$ for all estimators, with performance improving as n grows. The gap between NBALR and its competitors is largest at small sample sizes, confirming NBALR's advantage in that regime.

Table 3. Performance comparison of estimators in the NBRM based on simulated MSE and MAE when $p = 4$ and $\vartheta = 2$.

		NBML	NBR	NBL	NBAL	NBLT	NBOK	NBALR	NBML	NBR	NBL	NBAL	NBLT	NBOK	NBALR
ρ	n	-	$\hat{k}$	$\hat{d}$	$\hat{d}$	$\hat{k},\hat{d}$	$\hat{k},\hat{d}$	$\hat{k},\hat{d}$	-	$\hat{k}$	$\hat{d}$	$\hat{d}$	$\hat{k},\hat{d}$	$\hat{k},\hat{d}$	$\hat{k},\hat{d}$
					MSE							MAE			
0.80	50	0.52973	0.45748	0.47734	0.47643	0.41149	0.39655	0.21532	1.25657	1.16520	1.19103	1.19249	1.11039	1.08316	0.79525
	100	0.21706	0.19934	0.20657	0.20670	0.19319	0.18317	0.13873	0.81559	0.78168	0.79490	0.79576	0.77173	0.74954	0.65530
	150	0.12897	0.12174	0.12546	0.12536	0.12016	0.11488	0.09738	0.63427	0.61641	0.62503	0.62513	0.61303	0.59908	0.55242
	200	0.09511	0.09102	0.09311	0.09312	0.09035	0.08708	0.07734	0.54309	0.53146	0.53707	0.53727	0.52981	0.52007	0.49075
	300	0.06169	0.06004	0.06094	0.06093	0.05986	0.05842	0.05453	0.43729	0.43153	0.43448	0.43453	0.43101	0.42584	0.41183
	400	0.05135	0.05006	0.05083	0.05079	0.04993	0.04879	0.04575	0.39618	0.39129	0.39404	0.39398	0.39086	0.38645	0.37455
0.85	50	0.44452	0.38858	0.40453	0.40514	0.35517	0.34048	0.19508	1.15853	1.08080	1.10282	1.10554	1.03903	1.01022	0.76359
	100	0.29075	0.26022	0.27212	0.27204	0.24654	0.23294	0.15514	0.93187	0.88147	0.90042	0.90129	0.86165	0.83431	0.68528
	150	0.16483	0.15381	0.15909	0.15892	0.15072	0.14351	0.11605	0.70698	0.68314	0.69396	0.69404	0.67751	0.66021	0.59578
	200	0.12290	0.11646	0.12020	0.11990	0.11496	0.11029	0.09446	0.60817	0.59236	0.60090	0.60058	0.58924	0.57687	0.53554
	300	0.07901	0.07612	0.07770	0.07764	0.07571	0.07332	0.06638	0.49517	0.48625	0.49085	0.49081	0.48518	0.47751	0.45510
	400	0.06690	0.06476	0.06599	0.06593	0.06449	0.06268	0.05757	0.45169	0.44461	0.44843	0.44837	0.44386	0.43764	0.42005
0.90	50	0.95977	0.77987	0.80404	0.80096	0.64621	0.63650	0.24703	1.68462	1.50771	1.53466	1.53791	1.37165	1.35471	0.78761
	100	0.36434	0.31960	0.33519	0.33441	0.29583	0.28112	0.16629	1.04045	0.97318	0.99587	0.99662	0.94163	0.91197	0.70521
	150	0.25740	0.23155	0.24350	0.24272	0.22055	0.20813	0.14267	0.86951	0.82455	0.84448	0.84426	0.80825	0.78206	0.65205
	200	0.18802	0.17258	0.18013	0.17989	0.16776	0.15818	0.12041	0.75301	0.72190	0.73645	0.73652	0.71339	0.69195	0.60679
	300	0.11850	0.11214	0.11568	0.11548	0.11076	0.10603	0.09058	0.59319	0.57736	0.58559	0.58546	0.57447	0.56184	0.52087
	400	0.08426	0.08066	0.08288	0.08270	0.08009	0.07716	0.06864	0.50450	0.49391	0.50003	0.49975	0.49244	0.48347	0.45682
0.95	50	2.04098	1.53202	1.48286	1.46279	1.23792	1.17453	0.44591	2.38177	2.02689	2.00727	2.01260	1.77687	1.74856	0.95560
	100	0.68738	0.56934	0.60285	0.59912	0.48214	0.46919	0.18799	1.40568	1.27300	1.31108	1.31098	1.17841	1.15299	0.70460
	150	0.46376	0.39603	0.41986	0.41813	0.35406	0.33747	0.16691	1.15842	1.06755	1.09878	1.09907	1.01670	0.98424	0.69376
	200	0.31354	0.27568	0.29104	0.29054	0.25765	0.24195	0.14733	0.95337	0.89342	0.91692	0.91753	0.86854	0.83721	0.65897
	300	0.21417	0.19458	0.20423	0.20351	0.18756	0.17653	0.12732	0.78756	0.75077	0.76797	0.76750	0.73957	0.71562	0.61180
	400	0.17292	0.15889	0.16623	0.16578	0.15471	0.14581	0.11170	0.70521	0.67628	0.69067	0.69040	0.66899	0.64837	0.57064
0.99	50	7.18140	5.05233	4.04176	3.82141	4.23687	3.64126	1.42782	4.38999	3.51673	3.12906	3.13246	3.10784	2.88016	1.51431
	100	3.59647	2.53045	2.20790	2.16778	2.02257	1.82353	0.60423	3.10412	2.51775	2.36616	2.38491	2.15995	2.08288	1.01921
	150	2.23246	1.64396	1.59067	1.55860	1.28525	1.21542	0.39576	2.47513	2.07349	2.05503	2.05445	1.77090	1.74543	0.82069
	200	1.78735	1.31972	1.33940	1.30916	1.01996	0.96956	0.31234	2.20481	1.85071	1.87662	1.87304	1.57815	1.55969	0.72260
	300	1.31659	0.99424	1.04684	1.02033	0.76444	0.74891	0.21819	1.87647	1.59937	1.64926	1.64235	1.37780	1.36755	0.57337
	400	0.86200	0.68195	0.73405	0.72190	0.54890	0.53347	0.16256	1.52688	1.34325	1.39680	1.39200	1.20608	1.17939	0.58208

Table 4. Performance comparison of estimators in the NBRM based on simulated MSE and MAE when $p = 7$ and $\vartheta = 1$.

		NBML	NBR	NBL	NBAL	NBLT	NBOK	NBALR	NBML	NBR	NBL	NBAL	NBLT	NBOK	NBALR
ρ	n	-	$\hat{k}$	$\hat{d}$	$\hat{d}$	$\hat{k},\hat{d}$	$\hat{k},\hat{d}$	$\hat{k},\hat{d}$	-	$\hat{k}$	$\hat{d}$	$\hat{d}$	$\hat{k},\hat{d}$	$\hat{k},\hat{d}$	$\hat{k},\hat{d}$
					MSE							MAE			
0.80	50	1.66648	1.44412	1.37356	1.37918	1.22687	1.01009	0.59621	2.83034	2.62374	2.57248	2.58151	2.40232	2.17124	1.58987
	100	0.69828	0.62815	0.63489	0.63612	0.56910	0.47327	0.32914	1.85435	1.75618	1.76668	1.76994	1.67886	1.51860	1.27262
	150	0.43744	0.40371	0.41062	0.41152	0.38479	0.32327	0.26722	1.47584	1.41697	1.42910	1.43134	1.38689	1.26600	1.15653
	200	0.30524	0.28612	0.29145	0.29175	0.27811	0.23862	0.21174	1.22208	1.18303	1.19371	1.19484	1.16852	1.08006	1.02165
	300	0.20502	0.19520	0.19858	0.19873	0.19247	0.16966	0.15878	1.01305	0.98865	0.99671	0.99736	0.98252	0.92216	0.89364
	400	0.15740	0.15139	0.15369	0.15377	0.15006	0.13525	0.12936	0.88598	0.86903	0.87521	0.87564	0.86565	0.82182	0.80460
0.85	50	2.81401	2.34601	2.08934	2.08437	1.93608	1.54031	0.88848	3.66566	3.32599	3.15676	3.16623	2.98011	2.66047	1.93042
	100	0.89792	0.79532	0.79582	0.79723	0.69823	0.57899	0.36340	2.08823	1.95995	1.96313	1.96731	1.84251	1.66116	1.31230
	150	0.62625	0.56339	0.57055	0.57168	0.51716	0.42447	0.30598	1.75753	1.66535	1.67671	1.67964	1.60167	1.44234	1.23255
	200	0.42093	0.38782	0.39556	0.39600	0.36961	0.30902	0.25478	1.44179	1.38298	1.39634	1.39808	1.35358	1.23254	1.12476
	300	0.27600	0.25971	0.26445	0.26479	0.25360	0.21860	0.19709	1.15924	1.12435	1.13424	1.13537	1.11269	1.03116	0.98227
	400	0.20928	0.19921	0.20272	0.20285	0.19642	0.17296	0.16169	1.01766	0.99300	1.00123	1.00184	0.98677	0.92568	0.89654
0.90	50	3.21321	2.68839	2.35422	2.34482	2.21481	1.76248	1.00731	3.88617	3.52377	3.31276	3.32174	3.14457	2.80266	1.98150
	100	1.48634	1.28528	1.24135	1.24259	1.06473	0.88785	0.46623	2.66589	2.46530	2.43036	2.43686	2.23913	2.02265	1.40164
	150	0.91534	0.80776	0.80600	0.80855	0.70495	0.58069	0.34827	2.11200	1.97931	1.97995	1.98522	1.85696	1.66883	1.29277
	200	0.65514	0.58893	0.59631	0.59723	0.53741	0.44115	0.30947	1.78194	1.68713	1.69857	1.70136	1.61832	1.45535	1.22772
	300	0.40892	0.37689	0.38407	0.38473	0.35999	0.30055	0.24948	1.41210	1.35530	1.36787	1.36978	1.32800	1.20962	1.10824
	400	0.30614	0.28636	0.29206	0.29238	0.27827	0.23719	0.20959	1.22058	1.18040	1.19168	1.19284	1.16558	1.07423	1.01381
0.95	50	6.74411	5.39743	3.96543	3.92604	4.41520	3.27584	1.82077	5.57620	4.93005	4.23188	4.25302	4.36081	3.74832	2.61559
	100	3.19364	2.63749	2.25170	2.26062	2.05603	1.65330	0.75065	3.90682	3.51932	3.26253	3.28280	3.05667	2.73420	1.65644
	150	1.87266	1.57978	1.47871	1.48251	1.25418	1.03125	0.47182	2.96638	2.70590	2.62492	2.63595	2.39659	2.15431	1.35017
	200	1.37913	1.18207	1.14418	1.14855	0.96869	0.79509	0.38375	2.55895	2.35765	2.32428	2.33357	2.13428	1.91259	1.27763
	300	0.77703	0.68850	0.69235	0.69411	0.61198	0.49782	0.31372	1.93337	1.81630	1.82249	1.82702	1.72051	1.53720	1.22901
	400	0.59543	0.53550	0.54291	0.54412	0.49261	0.40071	0.28677	1.68734	1.59789	1.60959	1.61272	1.53878	1.37777	1.17533
0.99	50	41.24259	32.46885	15.90615	15.13113	27.68827	19.45392	11.99676	13.50862	11.76933	8.01641	8.04479	10.65982	8.80391	6.55491
	100	14.35715	11.18524	6.02265	6.09770	8.91402	6.37203	3.07755	8.08994	7.00557	5.07221	5.19928	6.11254	5.10282	3.21318
	150	8.21119	6.52392	4.30101	4.31887	5.14059	3.81995	1.77785	6.14086	5.39214	4.36052	4.42005	4.67168	4.00287	2.43347
	200	5.55645	4.39010	3.14513	3.19816	3.35902	2.52837	1.07328	5.07359	4.44152	3.74753	3.81599	3.79229	3.26796	1.88456
	300	3.88754	3.13154	2.50985	2.52620	2.37847	1.86339	0.77894	4.22584	3.74482	3.35471	3.38813	3.18618	2.81174	1.59955
	400	2.77057	2.26757	1.98716	1.99130	1.70406	1.37701	0.53586	3.57995	3.20618	3.00509	3.02209	2.73223	2.44466	1.29455

Table 5. Performance comparison of estimators in the NBRM based on simulated MSE and MAE when $p = 7$ and $\vartheta = 2$.

		NBML	NBR	NBL	NBAL	NBLT	NBOK	NBALR	NBML	NBR	NBL	NBAL	NBLT	NBOK	NBALR
ρ	n	-	$\hat{k}$	$\hat{d}$	$\hat{d}$	$\hat{k}, \hat{d}$	$\hat{k}, \hat{d}$	$\hat{k}, \hat{d}$	-	$\hat{k}$	$\hat{d}$	$\hat{d}$	$\hat{k}, \hat{d}$	$\hat{k}, \hat{d}$	$\hat{k}, \hat{d}$
					MSE							MAE			
0.80	50	1.00732	0.89056	0.88415	0.88599	0.78360	0.65040	0.41871	2.21054	2.07432	2.07107	2.07564	1.95013	1.76312	1.40064
	100	0.42971	0.39759	0.40447	0.40516	0.37949	0.32133	0.26810	1.45826	1.40221	1.41453	1.41641	1.37361	1.25939	1.15636
	150	0.26006	0.24616	0.25033	0.25065	0.24134	0.21060	0.19314	1.13613	1.10513	1.11411	1.11516	1.09552	1.02163	0.98051
	200	0.18501	0.17708	0.17981	0.17997	0.17506	0.15621	0.14774	0.96278	0.94211	0.94902	0.94963	0.93738	0.88535	0.86247
	300	0.12422	0.12035	0.12184	0.12191	0.11968	0.10978	0.10636	0.78913	0.77689	0.78138	0.78174	0.77500	0.74244	0.73137
	400	0.09926	0.09677	0.09782	0.09784	0.09642	0.08984	0.08788	0.70646	0.69769	0.70119	0.70138	0.69659	0.67260	0.66554
0.85	50	1.60062	1.36079	1.30433	1.30272	1.13821	0.92840	0.54071	2.76738	2.54393	2.50207	2.50633	2.31881	2.08581	1.53300
	100	0.55616	0.50663	0.51387	0.51501	0.47269	0.39443	0.30458	1.65278	1.57620	1.58784	1.59061	1.52787	1.38759	1.22608
	150	0.36133	0.33490	0.34138	0.34186	0.32251	0.27171	0.23240	1.33985	1.28953	1.30148	1.30305	1.26807	1.16056	1.07769
	200	0.25310	0.23889	0.24331	0.24359	0.23405	0.20279	0.18518	1.11650	1.08471	1.09423	1.09522	1.07491	0.99950	0.95738
	300	0.16880	0.16186	0.16428	0.16443	0.16026	0.14342	0.13642	0.91941	0.90047	0.90673	0.90738	0.89656	0.84808	0.82824
	400	0.12853	0.12448	0.12610	0.12614	0.12374	0.11338	0.10978	0.79594	0.78337	0.78814	0.78842	0.78135	0.74790	0.73647
0.90	50	1.89482	1.60136	1.51430	1.51074	1.31014	1.07285	0.58365	3.01685	2.76114	2.69665	2.70006	2.48052	2.23644	1.57670
	100	0.81694	0.72303	0.72940	0.72952	0.63726	0.52528	0.33073	1.99715	1.87508	1.88400	1.88665	1.76840	1.59046	1.26387
	150	0.54740	0.49816	0.50669	0.50734	0.46485	0.38515	0.29477	1.64234	1.56593	1.57960	1.58158	1.51813	1.37505	1.21148
	200	0.37792	0.34991	0.35679	0.35712	0.33589	0.28239	0.23893	1.35608	1.30471	1.31700	1.31843	1.28165	1.17176	1.08406
	300	0.23236	0.21965	0.22377	0.22397	0.21561	0.18723	0.17196	1.06868	1.03914	1.04835	1.04919	1.03069	0.95972	0.92210
	400	0.17958	0.17160	0.17454	0.17463	0.16966	0.15053	0.14233	0.94327	0.92230	0.92965	0.93015	0.91767	0.86454	0.84186
0.95	50	4.61050	3.67052	2.86987	2.86871	2.96610	2.20471	1.20208	4.55195	4.01982	3.58005	3.60845	3.52503	3.04972	2.09267
	100	2.08958	1.77668	1.65570	1.65623	1.41965	1.18173	0.56325	3.14933	2.88441	2.78952	2.79840	2.55964	2.31693	1.48352
	150	1.11943	0.96935	0.95623	0.95837	0.81595	0.66913	0.35076	2.32222	2.15346	2.14164	2.14789	1.98208	1.77550	1.26990
	200	0.79999	0.70562	0.71001	0.71133	0.62349	0.50633	0.31404	1.95055	1.82712	1.83414	1.83832	1.72490	1.53852	1.21590
	300	0.46256	0.42242	0.43114	0.43142	0.39923	0.32843	0.26155	1.49045	1.42353	1.43769	1.43928	1.38809	1.25413	1.12636
	400	0.35779	0.33175	0.33827	0.33870	0.31974	0.26818	0.22985	1.30848	1.25942	1.27126	1.27282	1.23912	1.13150	1.05231
0.99	50	26.18753	20.27638	11.06681	10.53046	17.12038	11.88964	7.30297	10.76618	9.30213	6.73293	6.73981	8.37198	6.88791	5.12519
	100	8.93962	7.05023	4.48260	4.53597	5.59231	4.09050	1.99961	6.44481	5.63380	4.46532	4.54330	4.90599	4.16146	2.64060
	150	4.57199	3.65708	2.81260	2.83850	2.83903	2.16621	1.00106	4.57537	4.03524	3.54135	3.58517	3.46413	3.01804	1.83421
	200	3.33274	2.70773	2.27759	2.28209	2.06603	1.63619	0.69261	3.94438	3.51701	3.22970	3.25124	3.01412	2.67159	1.54443
	300	2.28616	1.88141	1.69422	1.70406	1.42742	1.15805	0.44676	3.26329	2.93648	2.79053	2.80925	2.53103	2.26435	1.24927
	400	1.50824	1.26156	1.20483	1.21016	0.98248	0.79932	0.29908	2.64585	2.40529	2.35350	2.36556	2.12273	1.88940	1.08504

Table 6. Performance comparison of estimators in the NBRM based on simulated MSE and MAE when $p = 10$ and $\vartheta = 1$.

		NBML	NBR	NBL	NBAL	NBLT	NBOK	NBALR	NBML	NBR	NBL	NBAL	NBLT	NBOK	NBALR
ρ	n	-	$\hat{k}$	$\hat{d}$	$\hat{d}$	$\hat{k}, \hat{d}$	$\hat{k}, \hat{d}$	$\hat{k}, \hat{d}$	-	$\hat{k}$	$\hat{d}$	$\hat{d}$	$\hat{k}, \hat{d}$	$\hat{k}, \hat{d}$	$\hat{k}, \hat{d}$
					MSE							MAE			
0.80	50	2.88039	2.57535	2.37118	2.35873	2.20178	1.79356	1.22528	4.41114	4.15641	4.00131	3.99770	3.81238	3.43187	2.75008
	100	1.10592	1.00753	0.99602	0.99635	0.89608	0.72979	0.52316	2.75909	2.62893	2.61594	2.61829	2.48485	2.22505	1.88746
	150	0.71997	0.66473	0.66641	0.66748	0.61827	0.50110	0.41120	2.22691	2.13765	2.14081	2.14383	2.06641	1.84929	1.68314
	200	0.46384	0.43763	0.44124	0.44222	0.42308	0.35280	0.32537	1.78961	1.73759	1.74472	1.74717	1.71066	1.55776	1.50063
	300	0.31471	0.30043	0.30370	0.30408	0.29479	0.25233	0.24283	1.46867	1.43462	1.44212	1.44339	1.42231	1.31369	1.29139
	400	0.24096	0.23194	0.23438	0.23464	0.22924	0.20057	0.19682	1.28367	1.25935	1.26571	1.26661	1.25262	1.17090	1.16127
0.85	50	4.07416	3.48713	2.93671	2.92538	2.93294	2.23057	1.54459	5.18744	4.78030	4.41823	4.42423	4.32602	3.78481	3.06566
	100	1.55992	1.40391	1.34989	1.35556	1.20241	0.98361	0.63786	3.25766	3.08317	3.03008	3.03811	2.85376	2.56146	2.03900
	150	0.87215	0.80005	0.79697	0.79818	0.73026	0.59050	0.45605	2.43707	2.33116	2.32798	2.33134	2.23362	1.99372	1.76178
	200	0.70533	0.65362	0.65533	0.65644	0.61112	0.49723	0.41406	2.19968	2.11514	2.11848	2.12143	2.05026	1.83756	1.68615
	300	0.41973	0.39710	0.40091	0.40151	0.38560	0.32296	0.30187	1.69494	1.64802	1.65572	1.65750	1.62594	1.48430	1.43915
	400	0.29928	0.28572	0.28894	0.28927	0.28071	0.23988	0.23167	1.43196	1.39903	1.40666	1.40776	1.38774	1.28147	1.26164
0.90	50	6.38729	5.57799	4.52640	4.44804	4.73502	3.69229	2.58031	6.47317	6.02133	5.43810	5.41937	5.47906	4.83707	3.90710
	100	2.45484	2.15632	1.96887	1.97313	1.75812	1.42596	0.83972	4.05518	3.78650	3.62530	3.63642	3.40138	3.04800	2.26512
	150	1.19864	1.08658	1.06107	1.06537	0.95543	0.77148	0.52848	2.84794	2.70691	2.67816	2.68564	2.54534	2.26762	1.87943
	200	0.98757	0.89992	0.88953	0.89208	0.80889	0.64958	0.47640	2.59051	2.46881	2.45631	2.46154	2.34729	2.08560	1.79369
	300	0.65669	0.60884	0.61080	0.61229	0.57324	0.46274	0.39291	2.12389	2.04378	2.04746	2.05073	1.98708	1.77806	1.64659
	400	0.46478	0.43742	0.44122	0.44209	0.42225	0.34923	0.32063	1.76104	1.70780	1.71534	1.71744	1.68040	1.52407	1.46560
0.95	50	13.13664	11.39840	7.69693	7.53289	9.62333	7.38161	5.02209	9.23435	8.54333	6.98448	6.96977	7.74549	6.74825	5.34248
	100	4.31271	3.75869	3.14078	3.14481	2.99078	2.42425	1.37571	5.34098	4.96181	4.54112	4.55959	4.36791	3.92551	2.78661
	150	2.65327	2.31843	2.07114	2.07930	1.82793	1.49088	0.76007	4.22350	3.93096	3.71986	3.73514	3.46899	3.11068	2.09880
	200	2.16875	1.90340	1.74690	1.75656	1.52386	1.23678	0.63996	3.76970	3.51993	3.37729	3.39276	3.14379	2.80867	1.94426
	300	1.21147	1.09481	1.07006	1.07322	0.95418	0.76695	0.50282	2.84234	2.69666	2.66767	2.67423	2.52490	2.24232	1.82143
	400	0.90206	0.82270	0.81461	0.81767	0.74545	0.59296	0.44329	2.44620	2.33273	2.32254	2.32851	2.22725	1.97092	1.71509
0.99	50	69.00163	58.79192	25.31776	23.99210	50.20832	36.68295	25.70111	20.85886	19.07140	12.11313	12.10447	17.36259	14.70205	11.83107
	100	24.11506	20.18832	9.97768	9.88992	16.36631	11.94132	7.25755	12.48205	11.30369	7.78676	7.89641	10.00132	8.46552	6.27157
	150	14.37814	12.12861	7.04666	6.95294	9.61760	7.23325	4.07235	9.62291	8.76404	6.61505	6.65934	7.66153	6.60684	4.62926
	200	9.60961	8.09049	5.17076	5.20904	6.28550	4.76003	2.47310	7.90949	7.19411	5.70176	5.77768	6.22697	5.38061	3.58597
	300	5.78984	4.92857	3.63829	3.67471	3.76724	2.94166	1.40427	6.12155	5.60892	4.80922	4.85979	4.82681	4.24747	2.70326
	400	4.56982	3.89390	3.04835	3.07866	2.91605	2.31928	1.02814	5.47766	5.02595	4.44162	4.48273	4.28317	3.80806	2.29477

Multicollinearity severity (ρ) proves to be one of the most influential factors. As ρ rises from 0.80 to 0.99, MSE and MAE increase sharply for all estimators, but the NBML and the two single-parameter shrinkage estimators (NBR, NBL) degrade more severely than NBALR as the design matrix approaches singularity. At $\rho = 0.99$,

Table 7. Performance comparison of estimators in the NBRM based on simulated MSE and MAE when $p = 10$ and $\vartheta = 2$.

		NBML	NBR	NBL	NBAL	NBLT	NBOK	NBALR	NBML	NBR	NBL	NBAL	NBLT	NBOK	NBALR
ρ	n	-	$\hat{k}$	$\hat{d}$	$\hat{d}$	$\hat{k},\hat{d}$	$\hat{k},\hat{d}$	$\hat{k},\hat{d}$	-	$\hat{k}$	$\hat{d}$	$\hat{d}$	$\hat{k},\hat{d}$	$\hat{k},\hat{d}$	$\hat{k},\hat{d}$
					MSE							MAE			
0.80	50	1.61823	1.45001	1.39817	1.39662	1.24824	1.01815	0.69100	3.30444	3.12226	3.07406	3.07545	2.89537	2.60027	2.12021
	100	0.65299	0.60797	0.61140	0.61213	0.57331	0.47125	0.40430	2.12667	2.05052	2.05672	2.05889	1.99548	1.80048	1.67495
	150	0.43801	0.41323	0.41739	0.41797	0.39988	0.33425	0.30901	1.73207	1.68192	1.69040	1.69205	1.65693	1.51109	1.45776
	200	0.27133	0.26074	0.26354	0.26380	0.25716	0.22420	0.21884	1.36976	1.34257	1.34955	1.35045	1.33414	1.24430	1.23106
	300	0.18898	0.18304	0.18481	0.18496	0.18164	0.16199	0.16049	1.14241	1.12433	1.12952	1.13013	1.12037	1.05775	1.05365
	400	0.14446	0.14093	0.14209	0.14217	0.14027	0.12805	0.12775	0.99590	0.98368	0.98753	0.98791	0.98161	0.93779	0.93714
0.85	50	2.52450	2.15845	1.96021	1.94685	1.80588	1.38940	0.95226	4.09050	3.77249	3.61344	3.61338	3.42309	3.00660	2.43879
	100	0.89354	0.82002	0.81536	0.81749	0.74857	0.60917	0.47345	2.48151	2.37430	2.36953	2.37383	2.27449	2.03829	1.80411
	150	0.51641	0.48412	0.48826	0.48914	0.46424	0.38324	0.34502	1.88124	1.82072	1.82846	1.83077	1.78580	1.61769	1.54046
	200	0.40548	0.38365	0.38756	0.38813	0.37292	0.31273	0.29295	1.66845	1.62260	1.63082	1.63239	1.60172	1.46394	1.42111
	300	0.24508	0.23607	0.23857	0.23877	0.23331	0.20471	0.20085	1.29188	1.26775	1.27418	1.27493	1.26099	1.18001	1.17028
	400	0.17949	0.17415	0.17587	0.17597	0.17295	0.15506	0.15400	1.11372	1.09707	1.10222	1.10268	1.09361	1.03529	1.03236
0.90	50	3.59469	3.12357	2.74698	2.72353	2.59698	2.05134	1.36026	4.88066	4.52993	4.26397	4.25918	4.08570	3.63131	2.86568
	100	1.41593	1.26842	1.22661	1.23006	1.09735	0.88495	0.59461	3.09353	2.92227	2.87844	2.88521	2.72070	2.42624	1.97442
	150	0.67764	0.62836	0.63040	0.63161	0.58997	0.47910	0.40410	2.14835	2.06774	2.07152	2.07440	2.00867	1.80195	1.66441
	200	0.58487	0.54533	0.54849	0.54955	0.51805	0.42347	0.37033	1.99209	1.92252	1.92835	1.93101	1.87785	1.69084	1.58897
	300	0.38819	0.36800	0.37176	0.37229	0.35855	0.30154	0.28455	1.62534	1.58223	1.59011	1.59165	1.56351	1.43143	1.39409
	400	0.26057	0.25001	0.25283	0.25303	0.24666	0.21362	0.20874	1.33493	1.30751	1.31459	1.31534	1.29949	1.20843	1.19611
0.95	50	7.49780	6.47272	4.94415	4.86952	5.33301	4.12453	2.64080	6.99414	6.46087	5.63959	5.63044	5.78445	5.07316	3.88824
	100	2.56191	2.26246	2.07305	2.07252	1.84439	1.51149	0.88321	4.13928	3.87713	3.71610	3.72332	3.48279	3.13700	2.31255
	150	1.59114	1.42199	1.35976	1.36443	1.19368	0.97088	0.57451	3.26806	3.08232	3.01763	3.02627	2.82901	2.52860	1.93196
	200	1.26671	1.13127	1.09644	1.10182	0.96960	0.76999	0.48124	2.90322	2.73917	2.69997	2.70935	2.54494	2.24783	1.78065
	300	0.67666	0.62601	0.62831	0.62930	0.58682	0.47219	0.39509	2.13769	2.05420	2.05852	2.06126	1.99344	1.77836	1.63570
	400	0.52808	0.49300	0.49654	0.49758	0.47128	0.38324	0.34119	1.88989	1.82502	1.83174	1.83433	1.78761	1.60631	1.52223
0.99	50	39.18735	32.69168	15.21626	14.70222	27.25224	19.50612	12.94784	15.79085	14.27458	9.47406	9.53192	12.83221	10.75892	8.45366
	100	14.15002	11.79993	6.96525	6.93442	9.46255	6.95053	4.13863	9.58496	8.67274	6.58495	6.65252	7.64561	6.50315	4.82025
	150	7.95026	6.68956	4.50965	4.56678	5.22615	3.94474	2.12140	7.14704	6.51470	5.34334	5.41010	5.66200	4.91236	3.38123
	200	5.55784	4.72930	3.55929	3.58727	3.63630	2.85489	1.47138	6.03665	5.52681	4.79517	4.83679	4.75171	4.20002	2.77837
	300	3.31927	2.85941	2.41571	2.43746	2.18646	1.75532	0.80783	4.65139	4.29394	3.94572	3.97665	3.71572	3.31025	2.06829
	400	2.75544	2.39774	2.11048	2.12420	1.85339	1.50641	0.69384	4.23847	3.93709	3.69763	3.71808	3.43982	3.08011	1.95091

Table 8. Performance comparison of estimators in the NBRM based on simulated MSE and MAE when $p = 15$ and $\vartheta = 1$.

		NBML	NBR	NBL	NBAL	NBLT	NBOK	NBALR	NBML	NBR	NBL	NBAL	NBLT	NBOK	NBALR
ρ	n	-	$\hat{k}$	$\hat{d}$	$\hat{d}$	$\hat{k},\hat{d}$	$\hat{k},\hat{d}$	$\hat{k},\hat{d}$	-	$\hat{k}$	$\hat{d}$	$\hat{d}$	$\hat{k},\hat{d}$	$\hat{k},\hat{d}$	$\hat{k},\hat{d}$
					MSE							MAE			
0.80	50	6.41869	5.81342	4.90476	4.82361	5.05767	3.78053	3.14721	7.87802	7.48200	6.91091	6.86845	6.89819	5.97386	5.31500
	100	2.08249	1.93337	1.84872	1.84863	1.68410	1.33432	1.03083	4.58485	4.41184	4.31896	4.32142	4.11490	3.64072	3.17973
	150	1.27401	1.18928	1.16500	1.16748	1.07783	0.83721	0.71333	3.57907	3.45540	3.42200	3.42699	3.29578	2.88755	2.67754
	200	0.86203	0.81627	0.81331	0.81416	0.77259	0.61004	0.58012	2.95558	2.87482	2.87012	2.87234	2.80079	2.47895	2.42760
	300	0.51939	0.49796	0.49984	0.50053	0.48513	0.39481	0.38071	2.29616	2.24789	2.25223	2.25411	2.22054	2.01889	1.99917
	400	0.41294	0.39819	0.40050	0.40082	0.39114	0.32500	0.32430	2.03160	1.99452	2.00026	2.00133	1.97782	1.82835	1.79956
0.85	50	7.17407	6.46077	5.29415	5.20378	5.51502	4.10200	3.29002	8.35266	7.91043	7.19419	7.15023	7.23161	6.25042	5.48314
	100	2.64383	2.45649	2.31858	2.31374	2.10171	1.70237	1.25931	5.13954	4.94692	4.80982	4.80824	4.56623	4.08651	3.47896
	150	1.76712	1.64097	1.57933	1.58185	1.43823	1.12644	0.87719	4.18781	4.03223	3.96054	3.96549	3.78024	3.32558	2.93535
	200	1.22616	1.14722	1.12638	1.12849	1.04547	0.81324	0.70123	3.49468	3.37673	3.34755	3.35204	3.22987	2.82749	2.64086
	300	0.66888	0.63583	0.63570	0.63676	0.61110	0.48398	0.47820	2.59059	2.52493	2.52510	2.52767	2.47818	2.19855	2.19308
	400	0.53346	0.51068	0.51241	0.51306	0.49686	0.40143	0.40093	2.30701	2.25647	2.26030	2.26218	2.22747	2.01584	1.99708
0.90	50	11.94586	10.98612	8.74072	8.44159	9.73052	7.51345	6.43698	10.73489	10.25628	9.14976	9.02856	9.54941	8.35705	7.56225
	100	3.90820	3.59628	3.22851	3.21899	3.00703	2.40393	1.72847	6.22405	5.95914	5.65175	5.65067	5.41841	4.82734	4.00150
	150	2.41351	2.22635	2.08659	2.08892	1.87109	1.48866	1.03655	4.89688	4.69431	4.54744	4.55432	4.30092	3.80310	3.14994
	200	1.51276	1.40565	1.35972	1.36471	1.24716	0.96389	0.77248	3.89650	3.75262	3.69383	3.70171	3.54190	3.09282	2.78098
	300	1.08941	1.02046	1.00505	1.00772	0.94074	0.72371	0.64489	3.28968	3.18145	3.15819	3.16358	3.06021	2.66921	2.53361
	400	0.72783	0.68889	0.68670	0.68824	0.65743	0.51201	0.49895	2.69043	2.61636	2.61263	2.61611	2.55927	2.25066	2.23077
0.95	50	28.47495	25.96740	17.02902	15.99752	23.04150	17.22740	14.95994	16.47240	15.67162	12.66163	12.36234	14.60232	12.57299	11.48323
	100	9.57846	8.63917	6.47412	6.41792	7.18985	5.42078	4.06600	9.66224	9.14533	7.91171	7.90338	8.25489	7.14430	6.02899
	150	4.62533	4.21582	3.59363	3.60642	3.38759	2.68672	1.73613	6.73247	6.41155	5.92028	5.93963	5.70613	5.06027	3.93183
	200	3.25337	2.97254	2.66291	2.67100	2.40310	1.90319	1.20595	5.67448	5.41372	5.12794	5.14169	4.85152	4.29200	3.33984
	300	2.18799	2.00247	1.86380	1.87302	1.66989	1.28617	0.86338	4.64274	4.43366	4.28035	4.29450	4.05031	3.52450	2.87530
	400	1.54131	1.42474	1.36986	1.37511	1.24853	0.95038	0.73125	3.88475	3.73061	3.66014	3.66916	3.50038	3.02897	2.67335
0.99	50	130.54641	116.10792	55.50960	49.61428	102.54816	73.08145	64.83764	34.94297	32.83664	22.24580	21.35320	30.56908	25.73994	23.96605
	100	44.83962	40.35893	18.99829	18.53735	33.70550	24.93101	18.62420	20.71336	19.56998	13.18999	13.19853	17.70513	15.11241	12.68553
	150	22.87513	20.50379	10.81555	10.92472	16.50758	12.24555	8.22626	14.84010	13.99376	10.06569	10.19041	12.41351	10.61628	8.34487
	200	18.04067	16.11344	8.96535	9.07456	12.78594	9.45377	6.04158	13.26028	12.48280	9.22523	9.34865	10.99524	9.39125	7.17652
	300	10.17908	9.11444	6.13008	6.20646	7.08415	5.35850	3.20292	9.94679	9.37868	7.66588	7.74440	8.17812	7.07641	5.20150
	400	8.11528	7.25642	5.15505	5.22477	5.60693	4.24218	2.48622	8.87423	8.36422	7.03918	7.10694	7.26750	6.29924	4.57107

the performance gap widens substantially: in Table 4 ($n = 50$, $p = 7$, $\vartheta = 1$), the NBML's MSE reaches 41.243, compared to just 11.997 for the proposed estimator.

The number of predictors (p) also affects performance. As p increases from 4 to 10 to 15, MSE and MAE rise for all estimators, reflecting the greater complexity of multicollinearity with more predictors. NBALR remains the

Table 9. Performance comparison of estimators in the NBRM based on simulated MSE and MAE when $p = 15$ and $\vartheta = 2$.

		NBML	NBR	NBL	NBAL	NBLT	NBOK	NBALR	NBML	NBR	NBL	NBAL	NBLT	NBOK	NBALR
ρ	n	-	$\hat{k}$	$\hat{d}$	$\hat{d}$	$\hat{k}, \hat{d}$	$\hat{k}, \hat{d}$	$\hat{k}, \hat{d}$	-	$\hat{k}$	$\hat{d}$	$\hat{d}$	$\hat{k}, \hat{d}$	$\hat{k}, \hat{d}$	$\hat{k}, \hat{d}$
					MSE							MAE			
0.80	50	3.81185	3.40673	3.00943	2.98978	2.89135	2.13378	1.68578	6.10031	5.76449	5.44995	5.44007	5.27383	4.54186	3.96576
	100	1.18268	1.11035	1.09651	1.09685	1.02186	0.80540	0.71397	3.43512	3.32515	3.30601	3.30787	3.19468	2.81786	2.66296
	150	0.73085	0.69266	0.69240	0.69315	0.66116	0.52246	0.50749	2.69901	2.62676	2.62699	2.62891	2.56995	2.27704	2.25287
	200	0.47324	0.45490	0.45712	0.45759	0.44498	0.36606	0.37379	2.18705	2.14358	2.14889	2.15028	2.12138	1.94382	1.91972
	300	0.28971	0.28149	0.28327	0.28348	0.27871	0.24802	0.23921	1.70996	1.68559	1.69076	1.69154	1.67783	1.58372	1.55408
	400	0.23185	0.22624	0.22762	0.22778	0.22474	0.20410	0.19675	1.53008	1.51148	1.51596	1.51657	1.50674	1.43630	1.40947
0.85	50	4.10336	3.63045	3.12320	3.11002	2.99093	2.19042	1.64370	6.28217	5.90013	5.49765	5.49637	5.31485	4.55888	3.88979
	100	1.48789	1.39280	1.36336	1.36303	1.25588	0.99734	0.83736	3.85331	3.72513	3.68821	3.68919	3.54330	3.13884	2.88520
	150	1.01482	0.95540	0.94638	0.94776	0.89006	0.69805	0.63842	3.16892	3.07377	3.06139	3.06430	2.97362	2.62220	2.52213
	200	0.70235	0.66801	0.66848	0.66918	0.64109	0.51085	0.50210	2.66195	2.59510	2.59652	2.59829	2.54493	2.26458	2.25217
	300	0.36671	0.35398	0.35610	0.35646	0.34860	0.30054	0.29101	1.91782	1.88403	1.88955	1.89074	1.87055	1.73706	1.70686
	400	0.30572	0.29675	0.29863	0.29885	0.29358	0.25994	0.25076	1.75202	1.72605	1.73136	1.73218	1.71737	1.61643	1.58617
0.90	50	6.63145	6.01668	5.08144	4.98931	5.15033	3.91470	3.13820	7.98595	7.58259	6.97864	6.93450	6.94600	6.04452	5.30524
	100	2.10939	1.94474	1.84558	1.84686	1.67346	1.31278	0.99024	4.54767	4.35943	4.25099	4.25552	4.04141	3.55524	3.07347
	150	1.34303	1.25405	1.22634	1.22797	1.13426	0.88281	0.74658	3.66869	3.54135	3.50403	3.50786	3.37380	2.95552	2.73024
	200	0.86237	0.81441	0.80997	0.81135	0.76993	0.60159	0.57180	2.93828	2.85428	2.84712	2.85026	2.77914	2.44794	2.39605
	300	0.62586	0.59559	0.59610	0.59701	0.57443	0.45556	0.45407	2.50108	2.43922	2.44050	2.44288	2.39816	2.13369	2.13041
	400	0.42019	0.40406	0.40612	0.40657	0.39639	0.33471	0.32505	2.05438	2.01423	2.01926	2.02067	1.99613	1.83458	1.80491
0.95	50	15.66246	14.07256	9.96188	9.56608	12.19031	8.87989	7.42991	12.23226	11.55006	9.71794	9.58056	10.61972	9.04776	8.10703
	100	5.38804	4.85135	4.05875	4.03513	3.97547	3.03149	2.18777	7.25272	6.86474	6.28689	6.28146	6.15853	5.37036	4.45286
	150	2.54237	2.33694	2.16538	2.17141	1.95825	1.53978	1.06921	5.01914	4.80395	4.62890	4.63868	4.39196	3.86782	3.18696
	200	1.82409	1.68430	1.60751	1.61146	1.45522	1.12726	0.83876	4.24467	4.07359	3.98216	3.98966	3.79089	3.31204	2.86021
	300	1.26310	1.17291	1.14463	1.14779	1.05820	0.80179	0.67378	3.53446	3.40322	3.36364	3.36955	3.23892	2.80327	2.58483
	400	0.89840	0.84458	0.83678	0.83881	0.79282	0.60826	0.56928	2.97654	2.88447	2.87191	2.87616	2.79917	2.44135	2.37338
0.99	50	77.14998	67.00142	32.25520	29.99253	58.36568	39.94166	35.74113	26.67075	24.74774	16.87539	16.53786	22.82082	18.87671	17.71033
	100	27.28580	24.41999	13.41046	13.14159	20.21921	14.88182	11.03738	16.17634	15.24370	11.14824	11.14774	13.73228	11.70230	9.83112
	150	12.47140	11.13494	7.10964	7.18697	8.77259	6.55218	4.17641	10.99428	10.35681	8.25381	8.33292	9.09652	7.83895	6.01413
	200	10.29441	9.22812	6.26706	6.32675	7.27344	5.49833	3.46905	10.00158	9.43782	7.75267	7.82192	8.28641	7.17450	5.45567
	300	5.72389	5.14944	4.07150	4.10543	3.97146	3.08427	1.79277	7.45801	7.05097	6.26574	6.30649	6.12547	5.37627	3.87840
	400	4.41046	3.96509	3.27699	3.30954	3.04189	2.36499	1.31525	6.52317	6.16971	5.61011	5.64684	5.36193	4.70762	3.33431

best-performing estimator throughout, and its advantage over the alternatives grows with p; at $p = 15$ (Tables 8–9), NBALR's MSE values fall well below those of the NBML and even the NBOK, its closest biased competitor.

The dispersion parameter (ϑ) has comparatively little effect: comparing Tables 2–5 with 6–9 at matched sample size and correlation level shows only a minor MSE reduction at $\vartheta = 2$ relative to $\vartheta = 1$. Estimator rankings are unaffected by ϑ, with NBALR remaining best across both values.

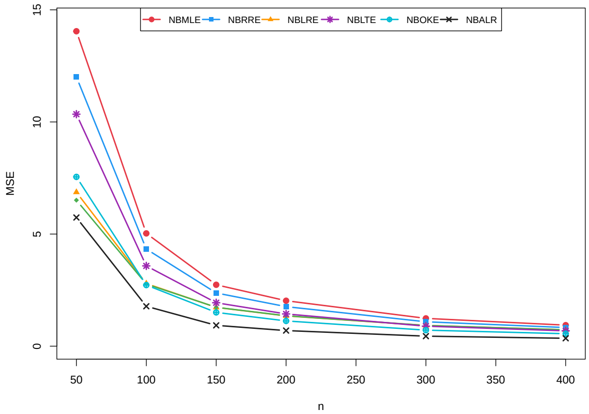
The NBML and several of its single- or two-parameter variants (NBR, NBL, NBAL, NBLT) share a common weakness: their performance deteriorates under high multicollinearity, particularly when combined with a large number of predictors. The NBML performs worst overall, with its largest MSE and MAE values occurring at $\rho = 0.90$ or 0.95. This confirms that ML estimation is highly sensitive to multicollinearity. The NBR and NBL estimators improve on the NBML but remain vulnerable to the same problem, performing poorly when p is large or ρ is severe.

The graphical results in Figures 1 and 2 support these patterns: all estimators approach their best performance as n grows large, with the NBALR curve lying below all others throughout. Across the full range of simulation settings, NBALR performs as well as or better than every alternative, with its advantage most pronounced under severe multicollinearity, small to moderate sample sizes, and a large number of predictors.

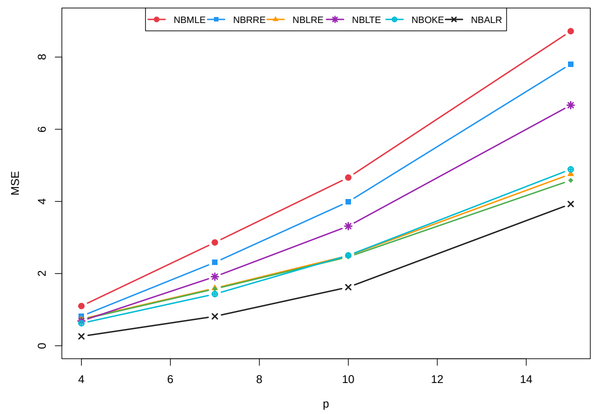
In summary, the simulation results show that ridge-type estimators for the NBRM, particularly NBOK and NBALR, improve substantially on the NBML under high multicollinearity or a large number of predictors, with NBALR performing best overall and its advantage growing under severe multicollinearity. These results support NBALR as the preferred estimator in practical applications.

5. Application

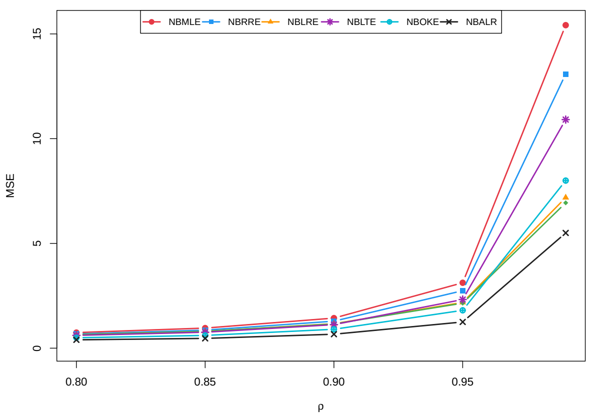
To test the proposed estimator's practical relevance beyond simulation, we apply it to COVID-19 surveillance data from Saudi Arabia (KSA). The dataset, available on Kaggle (`https://www.kaggle.com/datasets/fahdahalalyan/covid19-ksa`), contains 149 observations of epidemiological metrics collected during the pandemic. The response variable, number of recorded deaths y, is a non-negative count likely to show



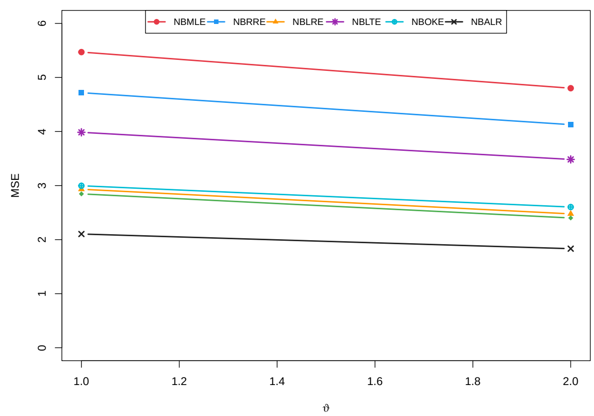
(a) MSE with sample size.



(b) MSE with the number of explanatory variables.



(c) MSE with multicollinearity level.



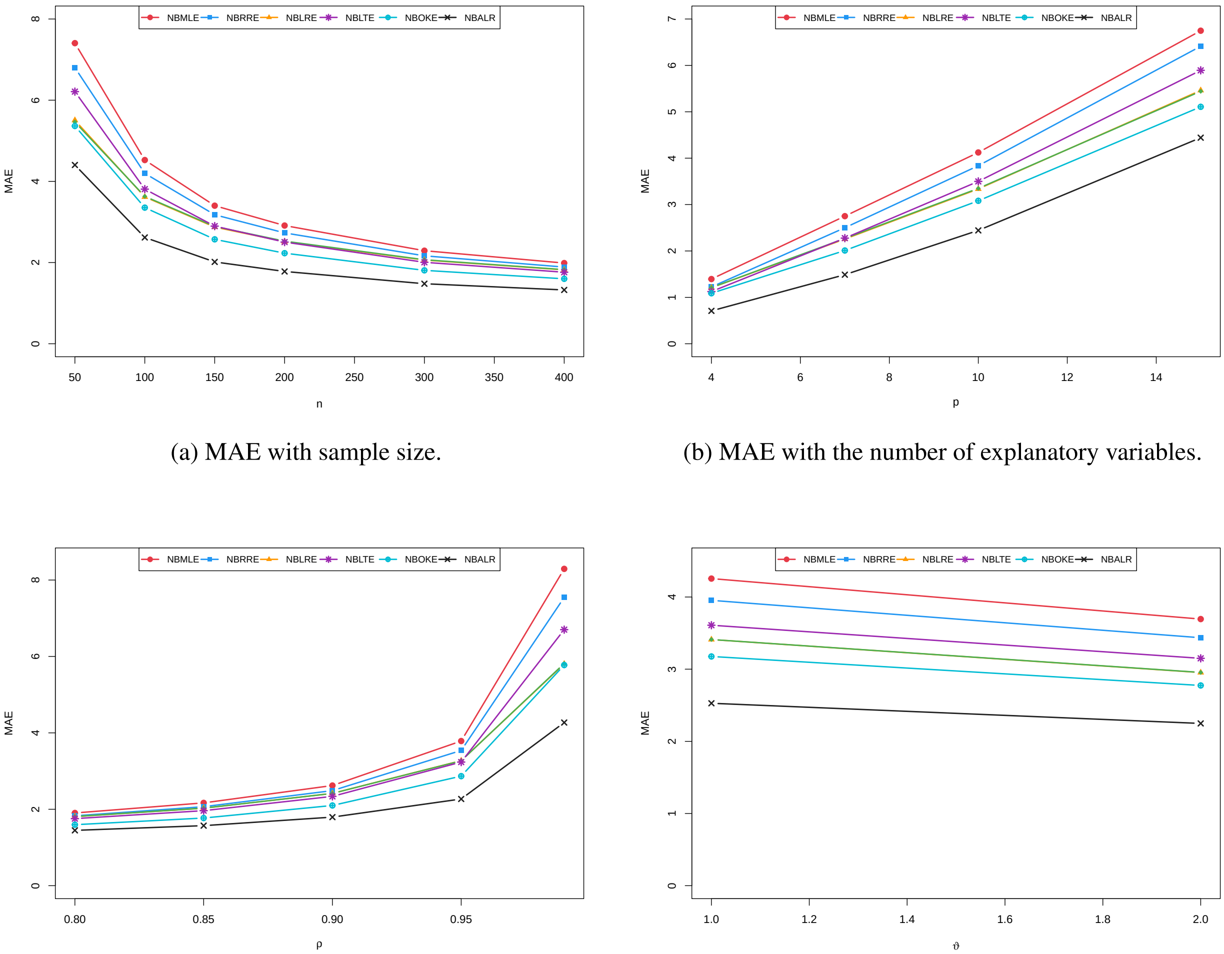
(d) MSE with values of the dispersion parameter.

Figure 1. Simulated MSE for estimators in the NBRM under different simulation factors.

overdispersion, making the negative binomial regression model a suitable choice. Four explanatory variables are included: cumulative confirmed cases (x_1), number of recovered individuals (x_2), active case count (x_3), and a set of derived epidemiological indicators (x_4). Together, these variables reflect the pandemic's progression across its main observable dimensions.

The response variable shows pronounced overdispersion: its variance (26,194.48) is far larger than its mean (41.92), yielding a variance-to-mean ratio of 624.88. This is corroborated by a Pearson dispersion statistic of 43.80, substantially exceeding the value of 1 that would hold under Poisson equidispersion.

Before estimation, we compared three candidate models for the mortality count data: the PRM, the NBRM, and the Poisson-inverse Gaussian regression model (PIGRM), using log-likelihood, AIC, and BIC. The PRM performed poorly (log-likelihood $= -2444.649$, AIC $= 4899.298$, BIC $= 4914.318$), consistent with count data of this kind rarely meeting the Poisson model's equidispersion assumption. Both overdispersed alternatives fit considerably better: the NBRM gave a log-likelihood of -411.258, AIC $= 834.516$, and BIC $= 852.540$, compared to -421.844, 855.687, and 873.711 for the PIGRM. Since the NBRM outperforms the PIGRM on all three criteria, it is used as



(a) MAE with sample size.

(b) MAE with the number of explanatory variables.

(c) MAE with multicollinearity level.

(d) MAE with values of the dispersion parameter.

Figure 2. Simulated MAE for estimators in the NBRM under different simulation factors.

the modeling framework for the remaining analyses, and these results confirm significant overdispersion in the data.

Before applying the estimators, we checked for multicollinearity among the explanatory variables. The correlation plot in Figure 3 gave an initial signal, confirmed by two formal diagnostics. The condition number (CN) came out to 16,456.09, far above the standard threshold of $1{,}000$ for severe multicollinearity. Variance inflation factors (VIFs) pointed in the same direction: x_1 and x_2 registered $13{,}044.12$ and $12{,}976.75$, both far past the acceptable limit of 10, while x_3 and x_4 sat at 4.18 and 1.41, showing no cause for concern. The high correlation between x_1 and x_2 makes sense, since confirmed and recovered case counts both track the same underlying pandemic trend. This level of multicollinearity undermines the NBML estimator and motivates the shrinkage estimator proposed here.

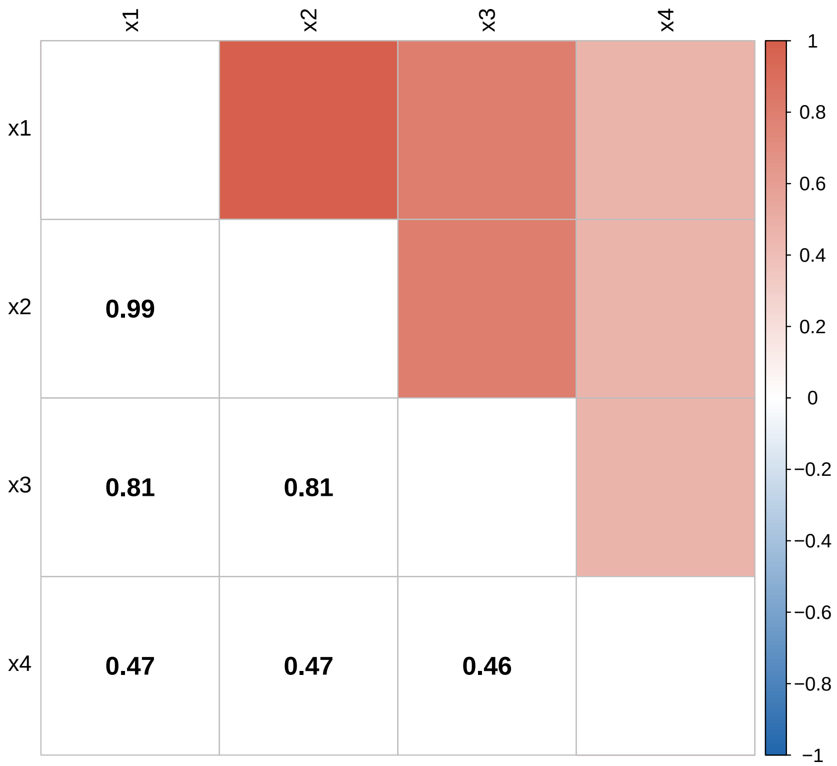


Figure 3. Correlation matrix among predictors in the dataset.

Table 10. Estimated coefficients, standard errors, and 95% confidence intervals for each estimator in the NBRM applied to the Saudi Arabia COVID-19 dataset.

Estimator	β_0	β_1	β_2	β_3	β_4	SMSE
	-0.33726	0.01347	-0.01355	-0.01248	0.01011	0.08585
NBML	(0.29276)	(0.00249)	(0.00254)	(0.01123)	(0.00141)	
	$[-0.91106, 0.23654]$	$[0.00860, 0.01834]$	$[-0.01853, -0.00857]$	$[-0.03448, 0.00953]$	$[0.00735, 0.01287]$	
	-0.19278	0.01334	-0.01341	-0.01301	0.00959	0.04902
NBR	(0.16734)	(0.00248)	(0.00253)	(0.01118)	(0.00112)	
	$[-0.52077, 0.13521]$	$[0.00849, 0.01819]$	$[-0.01837, -0.00845]$	$[-0.03491, 0.00890]$	$[0.00740, 0.01179]$	
	-0.31063	0.01345	-0.01352	-0.01257	0.01002	0.07356
NBL	(0.26965)	(0.00248)	(0.00254)	(0.01122)	(0.00135)	
	$[-0.83913, 0.21787]$	$[0.00858, 0.01831]$	$[-0.01850, -0.00855]$	$[-0.03456, 0.00941]$	$[0.00737, 0.01266]$	
	-0.19233	0.01334	-0.01341	-0.01301	0.00959	0.04902
NBAL	(0.16696)	(0.00248)	(0.00253)	(0.01118)	(0.00112)	
	$[-0.51956, 0.13490]$	$[0.00849, 0.01819]$	$[-0.01837, -0.00845]$	$[-0.03492, 0.00890]$	$[0.00740, 0.01178]$	
	-0.19233	0.01334	-0.01341	-0.01301	0.00959	0.04902
NBLT	(0.16696)	(0.00248)	(0.00253)	(0.01118)	(0.00112)	
	$[-0.51956, 0.13490]$	$[0.00849, 0.01819]$	$[-0.01837, -0.00845]$	$[-0.03493, 0.00890]$	$[0.00740, 0.01178]$	
	-0.18938	0.01172	-0.01177	-0.00789	0.00925	0.04899
NBOK	(0.16441)	(0.00218)	(0.00223)	(0.00741)	(0.00104)	
	$[-0.51161, 0.13285]$	$[0.00746, 0.01598]$	$[-0.01613, -0.00740]$	$[-0.02241, 0.00662]$	$[0.00721, 0.01130]$	
	-0.26768	0.01341	-0.01348	-0.01273	0.00986	0.00969
NBALR	(0.23237)	(0.00248)	(0.00253)	(0.01120)	(0.00126)	
	$[-0.72311, 0.18775]$	$[0.00855, 0.01827]$	$[-0.01845, -0.00852]$	$[-0.03469, 0.00922]$	$[0.00740, 0.01233]$	

Note: Standard errors are reported in parentheses; 95% confidence intervals are reported in square brackets.

Table 10 presents the estimation results for all estimators applied to the Saudi Arabian COVID-19 mortality data. The estimators differ mainly in their standard errors and the width of their confidence intervals. Ranking the estimators by SMSE, the proposed NBALR estimator achieves the lowest value at 0.00969, a substantial improvement over the NBML and all competing estimators. The NBOK follows at 0.04899, with the NBRRE and NBLTE close behind at 0.04902. The NBLRE records an SMSE of 0.07356, while the NBML performs worst at 0.08585, again reflecting the classical multicollinearity problem. The low SMSE of the NBALR points to strong performance relative to its competitors.

Turning to the coefficient estimates, the NBML produces the largest standard errors, most notably for the intercept ($\beta_0 = -0.33726,\ SE = 0.29276$) and for active cases ($\beta_3 = -0.01248,\ SE = 0.01123$). Its 95% confidence interval for β_0, $[-0.91106,\ 0.23654]$, contains zero and is therefore statistically insignificant, a direct consequence of the variance inflation caused by multicollinearity. The NBALR, by contrast, produces a standard error of 0.23237 for β_0 and a tighter confidence interval of $[-0.72311,\ 0.18775]$, a clear improvement over the NBML. The same pattern holds across the remaining coefficients, where NBALR consistently yields sharper confidence intervals than the NBML.

These real-data results align with the simulation findings and support the conclusion that the NBALR estimator produces more stable and reliable coefficient estimates than its competitors under severe multicollinearity.

6. Conclusion

While the NBRM is often used for modeling overdispersed count data, its conventional ML estimator can perform poorly when the regression variables are highly correlated. This high multicollinearity leads to large standard errors, unstable coefficient estimates, and inaccurate inferential testing. Many biased estimators have been developed in the literature to alleviate this issue, but no single estimator has proven optimal across all conditions in practice. In response to this gap, this article develops a new two-parameter biased estimator, called NBALR, for the NBRM. The proposed estimator combines, within a general framework, the shrinkage mechanisms of both the ridge and Liu estimators to address the problem of multicollinearity. By utilizing the shrinkage effects of both estimators, the proposed estimator achieves a better bias-variance trade-off than each estimator alone, especially under high multicollinearity, yielding more precise and stable parameter estimates. The asymptotic properties of the proposed NBALR estimator are established in terms of its MMSE and SMSE. These results are verified experimentally through a comprehensive Monte Carlo simulation across various sample sizes, numbers of covariates, multicollinearity levels, and overdispersion parameters. Furthermore, the application of the NBALR estimator to real COVID-19 mortality data from Saudi Arabia demonstrates its practical utility. However, this study has certain limitations. The NBALR estimator was investigated only within the context of the NBRM and not for more complex count data models such as hurdle and zero-inflated regression, the presented estimator is not designed to handle data containing outliers, and optimal selection criteria for the shrinkage parameters could also be investigated in future research as an extension of the works published by ALbazzaz et al. [34], Omara [35], and Kamel et al. [36].

CRediT authorship contribution statement

Eslam Hussam: Conceptualization, Methodology, Formal analysis, Validation, Investigation, Writing – original draft, Writing – review & editing, Funding acquisition. **Ali T. Hammad:** Conceptualization, Methodology, Formal analysis, Validation, Investigation, Writing – original draft, Writing – review & editing. **Ramy Aldallal:** Conceptualization, Methodology, Formal analysis, Validation, Investigation, Writing – original draft, Writing – review & editing. **Ahmed M. Gemeay:** Conceptualization, Methodology, Formal analysis, Validation, Investigation, Writing – original draft, Writing – review & editing.

Conflicts of Interest

The authors declare no conflict of interest.

Data Availability Statement

The dataset used in this study was available at https://www.kaggle.com/datasets/fahdahalalyan/covid19-ksa or upon request.

Acknowledgements

The authors extend their appreciation to Prince Sattam bin Abdulaziz University for funding this research work through the project number (PSAU/2025/02/32755).

REFERENCES

1. A. C. Cameron, and P. K. Trivedi, *Regression analysis of count data*, Cambridge University Press, vol. 53, 2013.
2. J. M. Hilbe, *Modeling count data*, Cambridge University Press, 2014.
3. W. Zhang, Q. Liu, J. Ni, J. Wang, and Z. Gong, *Negative binomial regression analysis of factors influencing the number of distinct mosquito species in Zhejiang Province, China, 2023*, Scientific Reports, vol. 15, no. 1, pp. 10433, 2025.
4. A. Ramadan, N. Chamidah, I. N. Budiantara, N. R. A. A. Siregar, and D. Aydin, *Modeling Pulmonary Tuberculosis Case Based on HIV and AIDS Cases in Indonesia Using Negative Binomial Regression Least Square Spline*, Statistics, Optimization & Information Computing, 2026.
5. J. M. Hilbe, *Negative binomial regression*, Cambridge University Press, 2011.
6. R. Frisch, *Statistical Confluence Analysis by Means of Complete Regression Systems*, The Economic Journal, vol. 45, no. 180, pp. 741–742, 1934.
7. P. Singh, S. Singh, and M. Paprzycki, *Detection and elimination of multicollinearity in regression analysis*, International Journal of Knowledge-Based and Intelligent Engineering Systems, vol. 27, no. 1, pp. 105–111, 2023.
8. J. H. Kim, *Multicollinearity and misleading statistical results*, Korean Journal of Anesthesiology, vol. 72, no. 6, pp. 558–569, 2019.
9. A. E. Hoerl, and R. W. Kennard, *Ridge regression: Biased estimation for nonorthogonal problems*, Technometrics, vol. 12, no. 1, pp. 55–67, 1970.
10. K. C. Arum, S. C. Ndukwe, H. E. Oranye, and O. B. Sule, *Comparative analysis of ridge and principal component regression in addressing multicollinearity*, FUDMA Journal of Sciences, vol. 9, no. 1, pp. 240–245, 2025.
11. L. Kejian, *A new class of biased estimate in linear regression*, Communications in Statistics-Theory and Methods, vol. 22, no. 2, pp. 393–402, 1993.
12. B. G. Kibria, and A. F. Lukman, *A new ridge-type estimator for the linear regression model: simulations and applications*, Scientifica, vol. 2020, no. 1, pp. 9758378, 2020.
13. A. F. Lukman, K. Ayinde, S. Binuomote, and O. A. Clement, *Modified ridge-type estimator to combat multicollinearity: Application to chemical data*, Journal of Chemometrics, vol. 33, no. 5, pp. e3125, 2019.
14. K. Liu, *Using Liu-type estimator to combat collinearity*, Communications in Statistics-Theory and Methods, vol. 32, no. 5, pp. 1009–1020, 2003.
15. A. F. Lukman, B. G. Kibria, K. Ayinde, and S. L. Jegede, *Modified one-parameter Liu estimator for the linear regression model*, Modelling and Simulation in Engineering, vol. 2020, no. 1, pp. 9574304, 2020.
16. I. Dawoud, and B. G. Kibria, *A new biased estimator to combat the multicollinearity of the Gaussian linear regression model*, Stats, vol. 3, no. 4, pp. 526–541, 2020.
17. M. Aslam, and S. Ahmad, *The modified Liu-ridge-type estimator: a new class of biased estimators to address multicollinearity*, Communications in Statistics-Simulation and Computation, vol. 51, no. 11, pp. 6591–6609, 2022.
18. F. M. Alghamdi, A. M. Gemeay, G. A. Abd-Elmougod, E. Hussam, and A. T. Hammad, *A new regularized Poisson regression estimator for effectively modelling multicollinear count data*, Statistics, pp. 1–33, 2026.
19. A. T. Hammad, I. Elbatal, O. A. Alqasem, E. M. Almetwally, A. M. Haleeb, A. M. Yousuf, *et al.*, *New Robust Ridge-Type Estimators for Beta Regression Under Multicollinearity and Outliers with Application to Breast Cancer Data*, Applied Mathematics in Science and Engineering, vol. 33, no. 1, pp. 2590432, 2025.
20. A. M. Algboory, A. N. Alkhateeb, and Z. Y. Algamal, *A Modified Jackknife Liu-Type Estimator for the Gamma Regression Models Data*, Statistics, Optimization & Information Computing, vol. 14, no. 5, pp. 2142–2154, 2025.
21. M. R. Özkale, and S. Kaciranlar, *The restricted and unrestricted two-parameter estimators*, Communications in Statistics-Theory and Methods, vol. 36, no. 15, pp. 2707–2725, 2007.
22. A. T. Hammad, E. H. Hafez, U. Shahzad, E. Yıldırım, E. M. Almetwally, and B. M. Golam Kibria, *New Modified Liu Estimators to Handle the Multicollinearity in the Beta Regression Model: Simulation and Applications*, Modern Journal of Statistics, vol. 1, no. 1, pp. 58–79, 2025.

23. E. Hussam, A. M. A. Gemeay, M. M. Abd El-Raouf, R. Aldallal, M. S. Mohamed, and A. T. Hammad, *A Novel Shrinkage Class for the Negative Binomial Regression Model: Theory, Simulation, and Healthcare Application in Saudi Arabia*, Mathematical and Computational Applications, vol. 31, no. 4, pp. 136, 2026.
24. K. Månsson, *On ridge estimators for the negative binomial regression model*, Economic Modelling, vol. 29, no. 2, pp. 178–184, 2012.
25. K. Månsson, *Developing a Liu estimator for the negative binomial regression model: method and application*, Journal of Statistical Computation and Simulation, vol. 83, no. 9, pp. 1773–1780, 2013.
26. Y. Asar, *Liu-type negative binomial regression: A comparison of recent estimators and applications*, in Trends and Perspectives in Linear Statistical Inference: Proceedings of the LINSTAT2016 meeting held 22-25 August 2016 in Istanbul, Turkey, pp. 23–39, Springer, 2018.
27. M. Kandemir Cetinkaya, and S. Kaciranlar, *Improved two-parameter estimators for the negative binomial and Poisson regression models*, Journal of Statistical Computation and Simulation, vol. 89, no. 14, pp. 2645–2660, 2019.
28. T. Koç and H. Koç, *A new effective jackknifing estimator in the negative binomial regression model*, Symmetry, vol. 15, no. 12, pp. 2107, 2023.
29. B. Ashraf, M. Amin, W. Emam, Y. Tashkandy, and M. Faisal, *Negative binomial regression model estimation using Stein approach: methods, simulation, and applications*, Journal of Mathematics, vol. 2025, no. 1, pp. 9134821, 2025.
30. D. M. Jabur, N. K. Rashad, and Z. Y. Algamal, *Jackknifed Liu-type estimator in the negative binomial regression model*, International Journal of Nonlinear Analysis and Applications, vol. 13, no. 1, pp. 2675–2684, 2022.
31. H. E. Oranye, K. C. Arum, F. I. Ugwuowo, S. O. Bashiru, T. E. Ugah, T. M. Adegoke, et al., *Selection of a new biasing parameter for the jackknife Kibria–Lukman estimator for the negative binomial regression model*, NIPES Journal of Science and Technology Research, Special Issue, vol. 7, no. 1, pp. 257–266, 2025.
32. F. M. Alghamdi, G. A. Abd-Elmougod, M. A. El-Qurashi, E. M. Almetwally, A. M. Gemeay, and A. T. Hammad, *A Bias-Reduced Estimator for Negative Binomial Regression with an Application to* CO_2 *Emissions Data*, Chiang Mai Journal of Science, vol. 52, no. 6, pp. 1–26, 2025.
33. R. W. Farebrother, *Further results on the mean square error of ridge regression*, Journal of the Royal Statistical Society: Series B (Methodological), vol. 38, no. 3, pp. 248–250, 1976.
34. Z. ALbazzaz, F. Younus, and Z. Y. Algamal, *Dawoud–Kibria Estimator for the Cox Regression Model*, Computational Journal of Mathematical and Statistical Sciences, vol. 5, no. 1, pp. 266–274, 2026. doi: 10.21608/cjmss.2026.432419.1280.
35. T. Omara, *Robust Lasso Estimator for the Liu-Type Regression Model and Its Applications*, Statistics, Optimization & Information Computing, vol. 13, no. 5, pp. 1819–1831, 2025.
36. A. Kamel, M. Abonazel, and A. Youssef, *A Novel Robust M-Estimator for the Random-Coefficients Regression Model: Simulation and Its Application in Energy Management Systems*, Computational Journal of Mathematical and Statistical Sciences, vol. 4, no. 2, pp. 738–767, 2025. doi: 10.21608/cjmss.2025.405428.1232.