Available online at http://scik.org
Commun. Math. Biol. Neurosci. 2025, 2025:53
https://doi.org/10.28919/cmbn/9149
ISSN: 2052-2541

A NEW ESTIMATOR OF THE GAMMA REGRESSION MODEL: THEORY, SIMULATION, AND APPLICATION TO BODY FAT DATA

MOHAMMED NAJI AL-GHAMDI[1], MOHAMED R. ABONAZEL[2,*], ISSAM DAWOUD[3],
ZAKARIYA YAHYA ALGAMAL[4], ABEER R. AZAZY[5]

[1]Department of General Studies, Technical College of Telecom and Information - Technical and Vocational Training Corporation (TVTC), Jeddah, Saudi Arabia

[2]Department of Applied Statistics and Econometrics, Faculty of Graduate Studies for Statistical Research, Cairo University, Giza, Egypt

[3]Department of Mathematics, Al-Aqsa University, Gaza, Palestine

[4]Department of Statistics and Informatics, University of Mosul, Mosul, Iraq

[5]Al-Alsun Higher Institute for Tourism, Hotels and Computer, Cairo, Egypt

Abstract: The gamma regression model, a specialized generalized linear model, effectively analyzes continuous, positive, and potentially skewed dependent variables when standard linear regression assumptions are violated. However, this model remains susceptible to the multicollinearity problem. Therefore, in this paper we propose a novel two-parameter shrinkage estimator (GDK) to address this problem, theoretically establishing its relationship with existing estimators through formal theorems. Through comprehensive Monte Carlo simulations examining various collinearity scenarios and a real-world numerical application, we demonstrate the GDK estimator's superior performance via mean squared error comparisons. Both simulation results and empirical evidence confirm that our proposed estimator outperforms existing alternatives, offering improved reliability for gamma regression analyses when multicollinearity is present.

[*]Corresponding author
E-mail address: mabonazel@cu.edu.eg
Received January 23, 2025

Keywords: body fat; gamma Kibria-Lukman estimator; gamma Özkale-Kaçiranlar estimator; gamma ridge regression estimator; maximum likelihood; multicollinearity.

2020 AMS Subject Classification: 62J07, 62J12, 62J20.

1. INTRODUCTION

The suitable model for finding the relationship between a skewed dependent variable having a gamma distribution with one or more explanatory variables is called the gamma regression model (GRM). This model is used in modeling real data in different fields such as the medical sciences, and the automobile insurance claim [1]. It is better to have a GRM in case of the positively skewed dependent variable having a gamma distribution with a given set of explanatory variables, see [2], [3], [4]. As we know in practice, the independence assumption of the explanatory variables in the linear regression models rarely exists, so multicollinearity also exists in the GRM that means the maximum likelihood estimator (MLE) is unstable and has a high variance [5]. Consequently, various authors proposed biased and unbiased estimators to handle the problem of multicollinearity. The popular alternative to MLE if there is multicollinearity in the linear regression model is the ridge estimator of Hoerl and Kennard [6]. This estimator is extended to the generalized linear models (GLM), see, e.g., [7], [8], [9], [10], [11], [12], [13], [14], where the gamma ridge regression estimator is denoted by the gamma ridge regression (GRR) estimator. Also, as an extension of the recent one parameter estimator proposed by [15], then Lukman et al. [16] obtained the gamma version of this estimator and denoted by the gamma Kibria-Lukman (GKL) estimator. Then, some of the two-parameter estimators are given to cope with multicollinearity in the linear regression model, such as [17], [18], [19] . Those proposed two-parameter estimators have shown that they are giving better results than the one-parameter estimators in several cases as the estimator obtained by [19]. As a result, Amin et al. [1] proposed the gamma two-parameter estimator of [19] and denoted by gamma Özkale-Kaçiranlar (GOK) estimator. Recently Dawoud and Kibria [17] proposed an efficient estimator called the Dawoud-Kibria (DK) estimator. Since the DK estimator is an efficient estimator among different two-parameter estimators in the linear model and there is no so analysis of the two-parameter estimators for the GRM, it is reasonable to introduce the gamma DK (GDK) estimator. This paper focuses on extending and developing the estimator proposed by [17] for the GRM.

2. STATISTICAL METHODOLOGY

Consider the dependent variable y_i has a gamma distribution with the parameters of the non-negative shape a and the nonnegative scale b, then the probability density function is

$$f(y_i) = \frac{y_i^{a-1} e^{-\frac{-y_i}{b}}}{\Gamma(a) b^a}, y_i \geq 0 \quad (1)$$

where $E(y_i) = ab = \theta_i$ and $Var(y_i) = ab^2 = \frac{\theta_i^2}{a}, \theta_i = e^{x_i' \beta}$. The log-likelihood function of equation (1) is given as:

$$l(\beta) = \sum_{i=1}^{n} [(a-1) \ln(y_i) - y_i/b - a \ln(b) - \ln(\Gamma(a))] \quad (2)$$

Since Equation (2) is not linear for β, we get it iteratively using the Fisher Scoring approach as:

$$\beta^{t+1} = \beta^t + I^{-1}(\beta^t) S(\beta^t) \quad (3)$$

where t is the iteration degree, $S(\beta) = \frac{\partial l(\beta)}{\partial \beta}$ and $I^{-1}(\beta) = \left(-E(\partial^2 l(\beta)/\partial \beta \partial \beta')\right)^{-1}$. The final step of the estimated parameter is:

$$\hat{\beta}_{MLE} = (X' \widehat{W} X)^{-1} X' \widehat{W} \hat{z} \quad (4)$$

where $\widehat{W} = diag(\hat{\theta}_i^2)$ matrix, $\hat{\theta}_i = exp(x_i' \hat{\beta}_{MLE})$ and $\hat{z}$ is the vector in ith element, $\hat{z} = \hat{\theta}_i + \frac{y_i - \hat{\theta}_i}{\hat{\theta}_i^2}$. Note that: $\widehat{W}$ and $\hat{z}$ are obtained iteratively (see [20], [21]).

Now, consider $\Gamma = \mathrm{diag}(\gamma_1, \ldots, \gamma_p) = (U'X'\widehat{W}XU)$, and $\alpha = U'\beta$, where $\gamma_1 \geq \cdots \geq \gamma_p \geq 0$, and the matrix $(X'\widehat{W}X)$ eigenvectors are the columns of matrix U. Then, the mean squared error matrix (MSEM) criterion as well as the mean squared error (MSE) of any estimator $\tilde{\beta}$ are given as:

$$MSEM(\tilde{\beta}) = Var(\tilde{\beta}) + \left(Bias(\tilde{\beta})\right)\left(Bias(\tilde{\beta})\right)', \quad (5)$$

$$MSE(\tilde{\beta}) = trace\left(MSEM(\tilde{\beta})\right). \quad (6)$$

Then the MSEM and MSE of $\hat{\beta}_{ML}$ as [22], [23]:

$$MSEM(\hat{\beta}_{MLE}) = Cov(\hat{\beta}_{ML}) = \phi (X' \widehat{W} X)^{-1} \quad (7)$$

where $\phi = \frac{1}{n-p} \sum_{i=1}^{n} \left(\frac{(y_i - \theta_i)^2}{\theta_i^2}\right)$.

$$MSE(\hat{\beta}_{MLE}) = \phi \sum_{j=1}^{p} \frac{1}{\gamma_j} \quad (8)$$

As it was written earlier, multicollinearity existence among the explanatory variables in the data indicates that there are some unfavorable effects on the estimates of regression parameter as the large variances. So, the MLE is not a good estimator in this case. Therefore, we should review some measures that are the remedy to multicollinearity and call the gamma biased regression estimators.

2.1 Gamma Ridge Regression (GRR) Estimator

Segerstedt [7] extended the ridge estimator to the GLM and the GRR estimator for the GRM is written as follows:

$$\hat{\beta}_{GRR} = (X'\widehat{W}X + kI_p)^{-1}X'\widehat{W}\hat{z} \qquad k > 0. \tag{9}$$

The MSEM and the MSE of $\hat{\beta}_{GRR}$ are

$$MSEM(\hat{\beta}_{GRR}) = \phi UL^{-1}\Gamma L^{-1}U' + (UL^{-1}\Gamma U' - I_p)\alpha\alpha'(UL^{-1}\Gamma U' - I_p)', \tag{10}$$

$$MSE(\hat{\beta}_{GRR}) = \phi \sum_{j=1}^{p} \frac{\gamma_j}{L_j^2} + k^2 \sum_{j=1}^{p} \frac{\alpha_j^2}{L_j^2} \tag{11}$$

where $L = (\Gamma + kI_p)$ and $L_j = (\gamma_j + k)$.

2.2 Gamma Kibria-Lukman (GKL) Estimator

Lukman et al. [16] obtained the GKL estimator as follows:

$$\hat{\beta}_{GKL} = (X'\widehat{W}X + kI_p)^{-1}(X'\widehat{W}X - kI_p)\hat{\beta}_{MLE}, \; k > 0, \tag{12}$$

The MSEM and the MSE of $\hat{\beta}_{GKL}$ are

$$MSEM(\hat{\beta}_{GKL}) = \phi UL^{-1}N\Gamma^{-1}NL^{-1}U' + (UL^{-1}NU' - I_p)\alpha\alpha'(UL^{-1}NU' - I_p)', \tag{13}$$

$$MSE(\hat{\beta}_{GKL}) = \phi \sum_{j=1}^{p} \frac{N_j^2}{\gamma_j L_j^2} + 4k^2 \sum_{j=1}^{p} \frac{\alpha_j^2}{L_j^2} \tag{14}$$

where $N = (\Gamma - kI_p)$ and $N_j = (\gamma_j - k)$.

2.3 Gamma Özkale-Kaçiranlar (GOK) Estimator

Amin et al. [1] obtained the GOK estimator as:

$$\hat{\beta}_{GOK} = (X'\widehat{W}X + kI_p)^{-1}(X'\widehat{W}X + kdI_p)\hat{\beta}_{MLE}, k > 0, 0 < d < 1 \tag{15}$$

The MSEM and MSE of $\hat{\beta}_{GOK}$ are

$$MSEM(\hat{\beta}_{GOK}) = \phi UL^{-1}G\Gamma^{-1}GL^{-1}U' + (UL^{-1}GU' - I_p)\alpha\alpha'(UL^{-1}GU' - I_p)', \tag{16}$$

$$MSE(\hat{\beta}_{GOK}) = \phi \sum_{j=1}^{p} \frac{G_j^2}{\gamma_j L_j^2} + (1-d)^2 k^2 \sum_{j=1}^{p} \frac{\alpha_j^2}{L_j^2} \tag{17}$$

where $G = (\Gamma + kdI_p)$ and $G_j = (\gamma_j + kd)$.

2.4 Proposed Gamma Dawoud-Kibria (GDK) Estimator

Because of the goodness of the DK estimator among the two parameter estimators in the linear regression model and the extensions of the two-parameter estimators to the area of generalized linear models, we introduce the GDK estimator as follows:

$$\hat{\beta}_{GDK} = (X'\widehat{W}X + k(1+d)I_p)^{-1}(X'\widehat{W}X - k(1+d)I_p)\hat{\beta}_{MLE}, k > 0, 0 < d < 1 \quad (18)$$

We give the MSEM and the MSE of the proposed $\hat{\beta}_{GDK}$ as follows:

$$MSEM(\hat{\beta}_{GDK}) = \phi U M^{-1} R \Gamma^{-1} R M^{-1} U' + (UM^{-1}RU' - I_p)\alpha\alpha'(UM^{-1}RU' - I_p)', \quad (19)$$

$$MSE(\hat{\beta}_{GDK}) = \phi \sum_{j=1}^{p} \frac{R_j^2}{\gamma_j M_j^2} + 4k^2(1+d)^2 \sum_{j=1}^{p} \frac{\alpha_j^2}{M_j^2}, \quad (20)$$

where $M = (\Gamma + k(1+d)I_p)$, $R = (\Gamma - k(1+d)I_p)$, $M_j = (\gamma_j + k(1+d))$ and $R_j = (\gamma_j - k(1+d))$.

2.5 Superiority of the Proposed GDK Estimator

This section presents four theorems that demonstrate the superiority of our proposed GDK estimator compared to the MLE, GRR, GKL, and GOK estimators.

Theorem 1: Compare the proposed GDK estimator with MLE: If $4k^2(1+d)^2 \sum_{j=1}^{p} \gamma_j \alpha_j^2 < \phi \sum_{j=1}^{p} (M_j^2 - R_j^2)$, then $MSE(\hat{\beta}_{GDK}) < MSE(\hat{\beta}_{MLE})$.

Proof: The MSE difference between the MLE and the GDK estimators is written as

$$\Delta_1 = MSE(\hat{\beta}_{GDK}) - MSE(\hat{\beta}_{MLE}) = \phi \sum_{j=1}^{p} \left[\frac{R_j^2 - M_j^2 + (1/\phi)4k^2(1+d)^2\gamma_j\alpha_j^2}{\gamma_j M_j^2} \right]. \quad (21)$$

In case of $R_j^2 - M_j^2 + (1/\phi)4k^2(1+d)^2\gamma_j\alpha_j^2 < 0$ in the equation (21), implies that $4k^2(1+d)^2 \sum_{j=1}^{p} \gamma_j \alpha_j^2 < \phi \sum_{j=1}^{p} (M_j^2 - R_j^2)$, then $MSE(\hat{\beta}_{GDK}) < MSE(\hat{\beta}_{MLE})$. That means, the GDK estimator is better than the MLE estimator if $4k^2(1+d)^2 \sum_{j=1}^{p} \gamma_j \alpha_j^2 < \phi \sum_{j=1}^{p} (M_j^2 - R_j^2)$.

Theorem 2: Compare the proposed GDK estimator with GRR estimator: If $\phi \sum_{j=1}^{p} (R_j^2 L_j^2 - \gamma_j^2 M_j^2) < k^2 \sum_{j=1}^{p} \gamma_j \alpha_j^2 (M_j^2 - 4(1+d)^2 L_j^2)$, then $MSE(\hat{\beta}_{GDK}) < MSE(\hat{\beta}_{GRR})$.

Proof: The MSE difference between the GRR and the GDK estimators is written as

$$\Delta_2 = MSE(\hat{\beta}_{GDK}) - MSE(\hat{\beta}_{GRR}) = \phi \sum_{j=1}^{p} \left[\frac{R_j^2 L_j^2 - \gamma_j^2 M_j^2 - (1/\phi)k^2\gamma_j\alpha_j^2(M_j^2 - 4(1+d)^2 L_j^2)}{\gamma_j L_j^2 M_j^2} \right]. \quad (22)$$

In case of $R_j^2 L_j^2 - \gamma_j^2 M_j^2 - (1/\phi)k^2\gamma_j\alpha_j^2(M_j^2 - 4(1+d)^2 L_j^2) < 0$ in the equation (22), implies that $\phi \sum_{j=1}^{p} (R_j^2 L_j^2 - \gamma_j^2 M_j^2) < k^2 \sum_{j=1}^{p} \gamma_j \alpha_j^2 (M_j^2 - 4(1+d)^2 L_j^2)$, then $MSE(\hat{\beta}_{GDK}) <$

$MSE(\hat{\beta}_{GRR})$. That means, the GDK estimator is better than the GRR estimator if $\phi \sum_{j=1}^{p}\left(R_j^2 L_j^2 - \gamma_j^2 M_j^2\right) < k^2 \sum_{j=1}^{p} \gamma_j \alpha_j^2 (M_j^2 - 4(1+d)^2 L_j^2)$.

Theorem 3: Compare the proposed GDK estimator with GKL estimator: If $\phi \sum_{j=1}^{p}\left(R_j^2 L_j^2 - N_j^2 M_j^2\right) < 4k^2 \sum_{j=1}^{p} \gamma_j \alpha_j^2 (M_j^2 - (1+d)^2 L_j^2)$, then $MSE(\hat{\beta}_{GDK}) < MSE(\hat{\beta}_{GKL})$

Proof: The MSE difference between the GKL and the GDK estimators is written as

$$\Delta_3 = MSE(\hat{\beta}_{GDK}) - MSE(\hat{\beta}_{GKL}) = \phi \sum_{j=1}^{p}\left[\frac{R_j^2 L_j^2 - N_j^2 M_j^2 - (1/\phi)4k^2 \gamma_j \alpha_j^2 (M_j^2 - (1+d)^2 L_j^2)}{\gamma_j L_j^2 M_j^2}\right]. \quad (23)$$

In case of $R_j^2 L_j^2 - N_j^2 M_j^2 - (1/\phi)4k^2 \gamma_j \alpha_j^2 (M_j^2 - (1+d)^2 L_j^2) < 0$ in the equation (23), implies that $\phi \sum_{j=1}^{p}\left(R_j^2 L_j^2 - N_j^2 M_j^2\right) < 4k^2 \sum_{j=1}^{p} \gamma_j \alpha_j^2 (M_j^2 - (1+d)^2 L_j^2)$, then $MSE(\hat{\beta}_{GDK}) < MSE(\hat{\beta}_{GKL})$. That means, the GDK estimator is better than the GKL estimator if $\phi \sum_{j=1}^{p}\left(R_j^2 L_j^2 - N_j^2 M_j^2\right) < 4k^2 \sum_{j=1}^{p} \gamma_j \alpha_j^2 (M_j^2 - (1+d)^2 L_j^2)$.

Theorem 4: Compare the proposed GDK estimator with GOK estimator: If $\phi \sum_{j=1}^{p}\left(R_j^2 L_j^2 - G_j^2 M_j^2\right) < k^2 \sum_{j=1}^{p} \gamma_j \alpha_j^2 ((1-d)^2 M_j^2 - 4(1+d)^2 L_j^2)$, then $MSE(\hat{\beta}_{GDK}) < MSE(\hat{\beta}_{GOK})$

Proof: The MSE difference between the GOK and the GDK estimators is written as

$$\Delta_4 = MSE(\hat{\beta}_{GDK}) - MSE(\hat{\beta}_{GOK}) = \phi \sum_{j=1}^{p}\left[\frac{R_j^2 L_j^2 - G_j^2 M_j^2 - (1/\phi)k^2 \gamma_j \alpha_j^2 ((1-d)^2 M_j^2 - 4(1+d)^2 L_j^2)}{\gamma_j L_j^2 M_j^2}\right]. \quad (24)$$

In case of $R_j^2 L_j^2 - G_j^2 M_j^2 - (1/\phi)k^2 \gamma_j \alpha_j^2 ((1-d)^2 M_j^2 - 4(1+d)^2 L_j^2) < 0$ in the equation (24), implies that $\phi \sum_{j=1}^{p}\left(R_j^2 L_j^2 - G_j^2 M_j^2\right) < k^2 \sum_{j=1}^{p} \gamma_j \alpha_j^2 ((1-d)^2 M_j^2 - 4(1+d)^2 L_j^2)$, then $MSE(\hat{\beta}_{GDK}) < MSE(\hat{\beta}_{GOK})$. That means, the GDK estimator is better than the GOK estimator if $\phi \sum_{j=1}^{p}\left(R_j^2 L_j^2 - G_j^2 M_j^2\right) < k^2 \sum_{j=1}^{p} \gamma_j \alpha_j^2 ((1-d)^2 M_j^2 - 4(1+d)^2 L_j^2)$.

2.6 Selection of biasing parameters (k and d)

We give the biasing parameters' estimators of the four mentioned estimators (GRR, GKL, GOK, and GDK).

- Following [24], $\hat{k}$ of the GRR estimator is written as

$$\hat{k}_{GRR} = \frac{p\hat{\phi}}{\sum_{j=1}^{p} \hat{\alpha}_j^2}. \quad (25)$$

- Following [16], $\hat{k}$ of the GKL estimator is written as

$$\hat{k}_{GKL} = \frac{p\,\hat{\phi}}{\sum_{j=1}^{p}\left(\frac{\hat{\phi}}{\gamma_j}+2\hat{\alpha}_j^2\right)}. \tag{26}$$

- Following [19], [1], $\hat{d}_{GOK}$ and $\hat{k}_{GOK}$ of the GOK estimator are written as

$$\hat{d}_{GOK} = \frac{1}{2} min\left(\frac{\hat{\alpha}_j^2}{\frac{\hat{\phi}}{\gamma_j}+\hat{\alpha}_j^2}\right)_{j=1}^{p} ; \tag{27}$$

$$\hat{k}_{GOK} = \frac{p\,\hat{\phi}}{\sum_{j=1}^{p}\left[\hat{\alpha}_j^2-\hat{d}_{GOK}\left(\frac{\hat{\phi}}{\gamma_j}+\hat{\alpha}_j^2\right)\right]}. \tag{28}$$

- Following [17], [25], [26], [27], [28], we can suggest four different $\hat{k}$ of the proposed GDK estimator as follows:

$$\hat{k}_{GDK.1} = \left(\hat{k}_{GRR}\right)^{1/p} \tag{29}$$

$$\hat{k}_{GDK.2} = \left(min\left(\frac{\hat{\phi}}{(1+\hat{d}_{GOK})\left(\frac{\hat{\phi}}{\gamma_j}+2\hat{\alpha}_j^2\right)}\right)\right)^{1/p} \tag{30}$$

$$\hat{k}_{GDK.3} = \left(\frac{p\hat{\phi}}{(1+\hat{d}_{GOK})\sum_{j=1}^{p}\left(\frac{\hat{\phi}}{\gamma_j}+2\hat{\alpha}_j^2\right)}\right)^{1/p} \tag{31}$$

$$\hat{k}_{GDK.4} = \left(\frac{1}{p}\sum_{j=1}^{p}\left(\frac{\hat{\phi}}{(1+\hat{d}_{GOK})\left(\frac{\hat{\phi}}{\gamma_j}+2\hat{\alpha}_j^2\right)}\right)\right)^{1/p} \tag{32}$$

The following section of our simulation study will employ the suggested estimators of the biasing parameters (k and d) for the four biased estimators (GRR, GKL, GOK, and GDK).

3. MONTE CARLO SIMULATION STUDY

This simulation study aims to assess the effectiveness of the proposed GDK estimator within the gamma regression model in addressing multicollinearity. The GDK estimator's performance is compared against other estimators (MLE, GRR, GKL, and GOK) using mean squared error (MSE) under various conditions. R software was used to conduct our simulation study, see [29], based on "*Gammareg*" and "*gamlss*" packages.

3.1 Simulation Design

We conducted simulation experiments using different values of n, p, ρ, and ϕ as follows:

The explanatory variables were generated as follows [25], [26], [27], [30]:

$$x_{ij} = \xi_{ij}(1-\rho^2)^{1/2} + \rho\xi_{ip-1}, \quad i = 1, \dots, n; j = 1, \dots, p-1 \tag{33}$$

where ξ_{ij} are generated from standard normal and ρ is the correlation between $(p-1)$ explanatory variables (assuming the generated model has an intercept). Since we are interested in the effect of multicollinearity, the correlation value can be considered the most important value.

The response variable (y_i) is generated as follows:

$$y_i \sim gamma(\theta_i, \phi); \ \theta_i = exp(x_i'\beta); \ i = 1, \dots, n. \tag{34}$$

We choose the values of β vector (with intercept) under the assumption that $\sum_{j=1}^{p} \beta_j^2 = 1; \beta_1 = \cdots = \beta_p$ according to [25], [27], [28], [30], [31].

We used the simulated MSE criterion to evaluate the performance of the proposed GDK estimator and other estimators by using the different combination of n, p, ρ, ϕ and the output data (y_i) is repeated 1000 times to calculate the MSE criterion as follows:

$$MSE(\hat{\beta}) = \frac{1}{1000} \sum_{t=1}^{1000} (\hat{\beta}_{(t)} - \beta)'(\hat{\beta}_{(t)} - \beta) \tag{35}$$

where $\hat{\beta}_{(t)}$ is the estimate of β in the t^{th} replication.

The simulation study examined several key factors across different values: Sample sizes (n) were tested at 50, 75, 100, 150, and 200 observations. Correlation levels (ρ) ranged from 0.80 to 0.99, including 0.80, 0.85, 0.90, 0.95, and 0.99. Regression parameters (p) were evaluated at three levels: 3, 6, and 9 predictors. Dispersion (ϕ) was considered at three values: 0.75, 1.00, and 1.25. This experimental design allowed for comprehensive evaluation across varying sample sizes, multicollinearity conditions, model complexities, and dispersion levels.

3.2 Simulation Results

The findings of the simulation study are summarized in Tables 1 to 9:

- Tables 1–3 display MSE values for the MLE, GRR, GKL, GOK, and GDK estimators when $p = 3$, across three levels of dispersion ($\phi = 0.75$, 1, and 1.25).
- Tables 4–6 present the same comparison for $p = 6$ under varying ϕ values.
- Tables 7–9 extend the analysis to $p = 9$, again evaluating performance at $\phi = 0.75$, 1, and 1.25.

In all tables, the best-performing estimator (lowest MSE) is highlighted in bold for easy identification.

In Table 1 ($\phi = 0.75$), we find that proposed GDK estimators (GDK.1–GDK.4) demonstrate superior performance over traditional estimators (MLE, GRR, GKL, GOK) across varying multicollinearity levels (ρ) and sample sizes (n). Notably, GDK.1 and GDK.4 consistently achieve the lowest MSE, particularly under high multicollinearity ($\rho \geq 0.95$). While the performance gap diminishes with larger sample sizes, GDK estimators remain advantageous in small-to-moderate samples, a common challenge in applied statistics.

In Table 2 ($\phi = 1$), GDK estimators maintain their dominance, with GDK.1 emerging as the most robust across ρ and n. In extreme multicollinearity ($\rho = 0.99$), GDK.4 excels, especially for small samples, suggesting its suitability for highly collinear, limited-data scenarios.

In Table 3 ($\phi = 1.25$), the results further validate GDK estimators' effectiveness. GDK.1 remains the top performer overall, while GDK.2 and GDK.3 occasionally outperform others in high ρ and large n. This consistency across dispersion levels highlights their adaptability to diverse real-world data conditions.

For Tables 4–6 ($p = 6$), the trends observed for $p = 3$ persist, with GDK estimators consistently outperforming traditional estimators under varying ϕ and ρ. Also, in Tables 7–9 ($p = 9$), similar results are achieved, confirming the robustness of GDK estimators even as the number of regression parameters increases. Their sustained superiority underscores their reliability in higher-dimensional settings.

In summary, these simulation results provide strong evidence for the efficacy of the proposed GDK estimators, particularly in challenging scenarios involving high multicollinearity and smaller sample sizes. Their consistent outperformance of established estimators across various conditions suggests that they represent a valuable addition to the toolbox of applied statisticians, especially when dealing with complex, highly correlated explanatory variables.

Figures 1 to 4 present a summary of the findings from our simulation study to assess the performance of the four proposed estimators (GDK.1, GDK.2, GDK.3, and GDK.4) under different values of $n, \rho, p,$ and ϕ in terms of MSE and identify the best one of them in most simulation cases.

TABLE 1: MSEs of MLE, GRR, GKL, GOK, and GDK estimators when $p = 3, \phi = 0.75$

n	ρ	MLE	GRR	GKL	GOK	GDK.1	GDK.2	GDK.3	GDK.4
50	0.80	0.24246	0.22967	0.22936	0.22998	**0.20873**	0.21500	0.21244	0.20905
	0.85	0.42733	0.38382	0.38311	0.38578	**0.31693**	0.32802	0.32215	0.32143
	0.90	0.43149	0.39605	0.39670	0.39732	0.36960	0.36049	**0.36033**	0.36505
	0.95	1.12203	0.94173	0.93094	0.95298	0.62997	0.68980	0.65578	**0.61651**
	0.99	4.45693	3.57454	3.53053	3.64819	3.00590	**2.60609**	2.73916	3.21025
75	0.80	0.16877	0.16550	0.16569	0.16555	0.16382	**0.15973**	0.16064	0.16720
	0.85	0.23469	0.21661	0.21613	0.21717	0.18019	0.19546	0.19010	**0.17814**
	0.90	0.29926	0.27466	0.27404	0.27562	**0.23279**	0.24519	0.24101	0.23424
	0.95	0.47427	0.41641	0.41402	0.41941	0.31468	0.34334	0.33184	**0.31403**
	0.99	2.89975	2.29097	2.24387	2.32997	1.54070	1.48173	**1.47316**	1.82727
100	0.80	0.12263	0.11928	0.11928	0.11936	0.11501	**0.11477**	0.11511	0.11909
	0.85	0.19585	0.18274	0.18244	0.18314	**0.15737**	0.16793	0.16436	0.15825
	0.90	0.26347	0.24074	0.24005	0.24162	0.19690	0.21368	0.20800	**0.19653**
	0.95	0.41981	0.38500	0.38557	0.38637	0.36993	**0.35040**	0.35363	0.36821
	0.99	2.03478	1.67365	1.66633	1.70185	1.46601	**1.32533**	1.34182	1.41049
150	0.80	0.06951	0.06813	0.06812	0.06815	**0.06573**	0.06645	0.06628	0.06742
	0.85	0.11408	0.11022	0.11021	0.11030	**0.10433**	0.10558	0.10541	0.10787
	0.90	0.15976	0.15056	0.15042	0.15083	**0.13697**	0.14023	0.13959	0.14190
	0.95	0.24181	0.22419	0.22408	0.22487	0.21177	**0.20766**	0.20857	0.21595
	0.99	1.54384	1.27453	1.26421	1.29578	1.03024	0.98221	**0.97214**	0.98428
200	0.80	0.05868	0.05735	0.05734	0.05736	**0.05527**	0.05588	0.05573	0.05669
	0.85	0.07547	0.07373	0.07373	0.07376	**0.07146**	0.07184	0.07178	0.07386
	0.90	0.10213	0.09866	0.09864	0.09872	**0.09492**	0.09525	0.09524	0.09834
	0.95	0.18979	0.17622	0.17597	0.17671	**0.15864**	0.16176	0.16120	0.16479
	0.99	1.02459	0.86418	0.85993	0.87599	0.73839	0.71111	0.70078	**0.69295**

TABLE 2: MSEs of MLE, GRR, GKL, GOK, and GDK estimators when $p = 3, \phi = 1$

n	ρ	MLE	GRR	GKL	GOK	GDK.1	GDK.2	GDK.3	GDK.4
50	0.80	0.12836	0.12712	0.12711	0.12713	**0.11984**	0.12251	0.12158	0.12088
	0.85	0.22735	0.22178	0.22173	0.22188	**0.19077**	0.19953	0.19730	0.19972
	0.90	0.23055	0.22545	0.22543	0.22553	**0.20510**	0.20839	0.20735	0.21261
	0.95	0.59499	0.57044	0.56990	0.57122	0.40745	0.45804	0.43966	**0.40258**
	0.99	2.35608	2.17676	2.17029	2.18741	1.43646	1.46851	1.41894	**1.40561**
75	0.80	0.09250	0.09219	0.09219	0.09219	**0.08989**	0.09017	0.09010	0.09414
	0.85	0.12767	0.12616	0.12615	0.12617	**0.11471**	0.11967	0.11796	0.11497
	0.90	0.16389	0.16160	0.16158	0.16162	**0.14588**	0.15138	0.14978	0.14930
	0.95	0.25869	0.25309	0.25303	0.25319	**0.21163**	0.22589	0.22162	0.21710
	0.99	1.56165	1.47050	1.46725	1.47447	0.84510	0.98158	0.91253	**0.82926**
100	0.80	0.06637	0.06614	0.06614	0.06614	**0.06436**	0.06479	0.06473	0.06676
	0.85	0.10680	0.10595	0.10595	0.10596	**0.09833**	0.10149	0.10046	0.09988
	0.90	0.14246	0.14092	0.14092	0.14094	**0.12711**	0.13258	0.13089	0.12881
	0.95	0.23129	0.22754	0.22752	0.22759	**0.20491**	0.20873	0.20768	0.21527
	0.99	1.12194	1.07236	1.07129	1.07426	0.78270	0.82506	0.79499	**0.77474**
150	0.80	0.03816	0.03808	0.03808	0.03808	**0.03734**	0.03761	0.03753	0.03793
	0.85	0.06302	0.06277	0.06277	0.06277	**0.06055**	0.06128	0.06112	0.06210
	0.90	0.08859	0.08809	0.08809	0.08809	**0.08339**	0.08487	0.08456	0.08570
	0.95	0.13272	0.13141	0.13140	0.13142	**0.12244**	0.12440	0.12401	0.12804
	0.99	0.84840	0.82146	0.82097	0.82227	**0.59650**	0.64753	0.62435	0.60797
200	0.80	0.03215	0.03209	0.03209	0.03209	**0.03147**	0.03168	0.03163	0.03175
	0.85	0.04217	0.04208	0.04208	0.04208	**0.04120**	0.04148	0.04141	0.04214
	0.90	0.05638	0.05617	0.05617	0.05617	**0.05442**	0.05496	0.05484	0.05580
	0.95	0.10623	0.10546	0.10546	0.10546	**0.09810**	0.10028	0.09989	0.10158
	0.99	0.56808	0.55390	0.55370	0.55426	**0.43541**	0.46131	0.45083	0.45131

TABLE 3: MSEs of MLE, GRR, GKL, GOK, and GDK estimators when $p = 3, \phi = 1.25$

n	ρ	MLE	GRR	GKL	GOK	GDK.1	GDK.2	GDK.3	GDK.4
50	0.80	0.07959	0.07811	0.07810	0.07813	**0.07504**	0.07626	0.07580	0.07671
	0.85	0.14172	0.13485	0.13478	0.13501	**0.12203**	0.12555	0.12481	0.12747
	0.90	0.14403	0.13815	0.13812	0.13826	0.13369	**0.13284**	0.13293	0.13753
	0.95	0.36877	0.33447	0.33332	0.33603	**0.25433**	0.28230	0.27277	0.25811
	0.99	1.46652	1.22552	1.21509	1.24390	0.95298	0.92167	**0.90701**	0.91044
75	0.80	0.05849	0.05826	0.05827	0.05826	0.05780	**0.05743**	0.05753	0.06127
	0.85	0.08036	0.07788	0.07785	0.07791	**0.07210**	0.07516	0.07402	0.07287
	0.90	0.10360	0.10002	0.09998	0.10009	**0.09310**	0.09580	0.09496	0.09580
	0.95	0.16344	0.15389	0.15371	0.15419	**0.13352**	0.14150	0.13909	0.13816
	0.99	0.97767	0.81954	0.80970	0.82967	**0.50804**	0.57619	0.53998	0.51419
100	0.80	0.04162	0.04133	0.04133	0.04134	**0.04073**	0.04074	0.04079	0.04274
	0.85	0.06726	0.06559	0.06558	0.06562	**0.06158**	0.06358	0.06289	0.06297
	0.90	0.08914	0.08605	0.08602	0.08611	**0.07847**	0.08212	0.08089	0.08037
	0.95	0.14641	0.13994	0.13990	0.14009	0.13405	**0.13245**	0.13294	0.14023
	0.99	0.71096	0.62794	0.62616	0.63285	0.53787	0.52905	0.51994	**0.51620**
150	0.80	0.02412	0.02398	0.02398	0.02398	**0.02366**	0.02376	0.02374	0.02420
	0.85	0.03996	0.03939	0.03939	0.03939	**0.03832**	0.03869	0.03861	0.03957
	0.90	0.05636	0.05523	0.05523	0.05525	**0.05300**	0.05373	0.05358	0.05471
	0.95	0.08387	0.08149	0.08147	0.08153	0.07900	**0.07890**	0.07894	0.08286
	0.99	0.53644	0.47813	0.47630	0.48140	**0.38834**	0.39852	0.39085	0.38838
200	0.80	0.02031	0.02015	0.02015	0.02015	**0.01982**	0.01993	0.01990	0.02010
	0.85	0.02698	0.02679	0.02679	0.02679	**0.02641**	0.02652	0.02649	0.02729
	0.90	0.03571	0.03531	0.03531	0.03531	**0.03474**	0.03485	0.03483	0.03589
	0.95	0.06794	0.06609	0.06608	0.06612	**0.06278**	0.06374	0.06358	0.06542
	0.99	0.36093	0.32920	0.32844	0.33078	0.28695	0.28935	**0.28672**	0.29098

TABLE 4: MSEs of MLE, GRR, GKL, GOK, and GDK estimators when $p = 6, \phi = 0.75$

n	ρ	MLE	GRR	GKL	GOK	GDK.1	GDK.2	GDK.3	GDK.4
50	0.80	1.04528	0.87279	0.86843	0.87398	0.60499	0.68287	0.64038	**0.56198**
	0.85	1.83555	1.51131	1.50339	1.51352	0.91310	1.04840	0.97336	**0.83689**
	0.90	1.81604	1.43699	1.42344	1.44017	0.78723	0.92210	0.84696	**0.72584**
	0.95	3.33652	2.56584	2.53953	2.57277	1.03268	1.24371	1.11657	**0.93351**
	0.99	17.09026	12.48343	12.30430	12.53271	6.53834	**5.26834**	5.93525	7.77815
75	0.80	0.56829	0.48797	0.48639	0.48842	0.39725	0.43327	0.41386	**0.38276**
	0.85	0.93971	0.76637	0.76082	0.76796	0.52415	0.59760	0.55838	**0.49578**
	0.90	1.30964	1.06695	1.06281	1.06830	0.68791	0.78680	0.73368	**0.64948**
	0.95	2.49502	1.93581	1.91731	1.94091	0.92448	1.09558	0.99562	**0.86705**
	0.99	13.47306	9.69970	9.58423	9.73043	4.64593	**3.85197**	4.23737	5.31237
100	0.80	0.42429	0.37051	0.36947	0.37082	0.31608	0.34050	0.32699	**0.30642**
	0.85	0.61556	0.51586	0.51360	0.51653	0.40109	0.44480	0.42133	**0.38513**
	0.90	1.93201	1.60427	1.59652	1.60645	0.79688	0.91777	0.84023	**0.76527**
	0.95	1.72440	1.31860	1.30497	1.32256	0.70392	0.82740	0.75975	**0.68782**
	0.99	8.81448	6.13810	6.02985	6.16990	2.56071	**2.19936**	2.34814	2.83732
150	0.80	0.28182	0.25234	0.25200	0.25245	0.22834	0.24204	0.23411	**0.22212**
	0.85	0.34096	0.30211	0.30155	0.30229	0.26744	0.28524	0.27523	**0.26054**
	0.90	0.54272	0.45624	0.45437	0.45678	0.36460	0.40292	0.38208	**0.35285**
	0.95	1.17756	0.93546	0.93019	0.93703	0.60791	0.69159	0.64843	**0.59607**
	0.99	5.23515	3.62223	3.56057	3.63818	1.19361	1.23128	**1.17810**	1.23615
200	0.80	0.21227	0.19490	0.19476	0.19494	0.18229	0.19012	0.18556	**0.17823**
	0.85	0.26830	0.24084	0.24051	0.24095	0.21845	0.23115	0.22391	**0.21299**
	0.90	0.38345	0.33254	0.33178	0.33279	0.28676	0.30862	0.29649	**0.27968**
	0.95	0.76421	0.61376	0.61026	0.61478	0.43948	0.49842	0.46770	**0.42996**
	0.99	3.67588	2.60855	2.57109	2.61856	0.94163	1.07440	0.98407	**0.92497**

TABLE 5: MSEs of MLE, GRR, GKL, GOK, and GDK estimators when $p = 6, \phi = 1$

n	ρ	MLE	GRR	GKL	GOK	GDK.1	GDK.2	GDK.3	GDK.4
50	0.80	0.53610	0.51200	0.51178	0.51211	0.40264	0.43441	0.41726	**0.37614**
	0.85	0.94069	0.88982	0.88924	0.89006	0.63062	0.69413	0.66091	**0.58733**
	0.90	0.94240	0.88235	0.88151	0.88274	0.57241	0.64513	0.60748	**0.53154**
	0.95	1.74715	1.61303	1.61075	1.61408	0.83533	0.98426	0.90532	**0.74612**
	0.99	8.91866	7.85333	7.82646	7.86528	2.30026	**2.17275**	2.18971	2.50589
75	0.80	0.29762	0.29014	0.29010	0.29016	0.25003	0.26298	0.25591	**0.24102**
	0.85	0.49798	0.48014	0.47999	0.48023	0.37252	0.40303	0.38704	**0.35309**
	0.90	0.69882	0.67002	0.66984	0.67013	0.49401	0.53861	0.51566	**0.46827**
	0.95	1.32340	1.24487	1.24379	1.24546	0.73414	0.83855	0.78403	**0.68638**
	0.99	7.14594	6.53427	6.52446	6.53912	**1.90266**	2.01780	1.90882	1.93443
100	0.80	0.22750	0.22341	0.22340	0.22342	0.19832	0.20657	0.20202	**0.19154**
	0.85	0.33171	0.32365	0.32361	0.32367	0.27097	0.28730	0.27844	**0.25925**
	0.90	0.92158	0.84773	0.84652	0.84824	0.50640	0.56332	0.52670	**0.47417**
	0.95	0.92477	0.88017	0.87972	0.88042	0.55435	0.62364	0.58949	**0.52552**
	0.99	4.70562	4.33029	4.32434	4.33362	1.20127	1.40247	1.27047	**1.16036**
150	0.80	0.15385	0.15217	0.15216	0.15217	0.13985	0.14411	0.14169	**0.13550**
	0.85	0.18561	0.18323	0.18323	0.18324	0.16593	0.17170	0.16848	**0.16041**
	0.90	0.29487	0.28920	0.28918	0.28921	0.24441	0.25856	0.25088	**0.23454**
	0.95	0.64203	0.62230	0.62220	0.62237	0.45617	0.49569	0.47609	**0.43539**
	0.99	2.85986	2.69204	2.68999	2.69317	0.91398	1.14048	1.01376	**0.83065**
200	0.80	0.11738	0.11652	0.11652	0.11652	0.10961	0.11198	0.11066	**0.10702**
	0.85	0.14698	0.14559	0.14559	0.14559	0.13413	0.13801	0.13586	**0.13007**
	0.90	0.20966	0.20681	0.20681	0.20682	0.18306	0.19060	0.18646	**0.17638**
	0.95	0.42132	0.41160	0.41156	0.41163	0.32327	0.34800	0.33506	**0.30762**
	0.99	2.01630	1.92052	1.91970	1.92106	0.79157	0.96731	0.87583	**0.73761**

TABLE 6: MSEs of MLE, GRR, GKL, GOK, and GDK estimators when $p = 6, \phi = 1.25$

n	ρ	MLE	GRR	GKL	GOK	GDK.1	GDK.2	GDK.3	GDK.4
50	0.80	0.32708	0.29435	0.29393	0.29449	0.26409	0.27998	0.27083	**0.25374**
	0.85	0.57518	0.50195	0.50077	0.50231	0.41965	0.45333	0.43495	**0.40535**
	0.90	0.57942	0.49351	0.49158	0.49407	0.39310	0.43260	0.41128	**0.37863**
	0.95	1.07972	0.87532	0.87010	0.87684	0.58658	0.67310	0.62746	**0.55620**
	0.99	5.50251	3.91121	3.84768	3.92831	1.36855	1.41841	**1.35831**	1.38572
75	0.80	0.18390	0.17152	0.17142	0.17156	0.16072	0.16716	0.16339	**0.15699**
	0.85	0.30948	0.27855	0.27806	0.27874	0.24543	0.26173	0.25257	**0.23793**
	0.90	0.43607	0.38758	0.38712	0.38776	0.33425	0.35721	0.34457	**0.32554**
	0.95	0.82331	0.68858	0.68544	0.68960	0.50395	0.56333	0.53199	**0.48867**
	0.99	4.44815	3.29874	3.26932	3.30783	1.21882	1.35649	1.25542	**1.19720**
100	0.80	0.14249	0.13491	0.13485	0.13493	0.12818	0.13224	0.12984	**0.12530**
	0.85	0.20822	0.19264	0.19249	0.19270	0.17768	0.18610	0.18118	**0.17302**
	0.90	0.67359	0.60457	0.60389	0.60475	0.34620	0.39720	0.36303	**0.32182**
	0.95	0.57855	0.49229	0.49049	0.49293	0.37968	0.41812	0.39840	**0.37085**
	0.99	2.93285	2.17670	2.15155	2.18483	0.82096	0.98392	0.88614	**0.80045**
150	0.80	0.09704	0.09315	0.09313	0.09315	0.08994	0.09207	0.09078	**0.08810**
	0.85	0.11698	0.11161	0.11158	0.11162	0.10706	0.10994	0.10822	**0.10475**
	0.90	0.18564	0.17276	0.17265	0.17280	0.15995	0.16724	0.16297	**0.15590**
	0.95	0.40522	0.35981	0.35929	0.36000	0.30803	0.32917	0.31769	**0.30126**
	0.99	1.80556	1.39551	1.38420	1.39896	0.64804	0.79202	0.71393	**0.62652**
200	0.80	0.07453	0.07230	0.07230	0.07231	0.07046	0.07166	0.07094	**0.06929**
	0.85	0.09301	0.08949	0.08947	0.08949	0.08653	0.08844	0.08730	**0.08478**
	0.90	0.13234	0.12523	0.12519	0.12525	0.11875	0.12257	0.12031	**0.11597**
	0.95	0.26714	0.24180	0.24153	0.24189	0.21491	0.22842	0.22067	**0.20903**
	0.99	1.27581	1.00699	1.00061	1.00907	0.55564	0.66061	0.60650	**0.54758**

TABLE 7: MSEs of MLE, GRR, GKL, GOK, and GDK estimators when $p = 9, \phi = 0.75$

n	ρ	MLE	GRR	GKL	GOK	GDK.1	GDK.2	GDK.3	GDK.4
50	0.80	2.01084	1.63950	1.62828	1.64115	1.05289	1.17333	1.10140	**0.98470**
	0.85	3.14732	2.58170	2.56907	2.58333	1.45768	1.65254	1.53273	**1.33103**
	0.90	4.55284	3.57835	3.55267	3.58198	1.52167	1.77906	1.61805	**1.38610**
	0.95	7.18276	5.50448	5.45391	5.51169	1.75053	2.01848	1.83901	**1.62137**
	0.99	37.68948	28.55036	28.29416	28.58026	12.72221	**10.60061**	11.87035	15.48784
75	0.80	1.26879	1.02675	1.02094	1.02745	0.77772	0.85102	0.80686	**0.74181**
	0.85	1.55394	1.24536	1.23786	1.24627	0.87442	0.97015	0.91268	**0.82016**
	0.90	2.30522	1.82235	1.81228	1.82364	1.10179	1.24534	1.15955	**1.03901**
	0.95	4.55263	3.42009	3.39283	3.42410	1.21670	1.47446	1.31440	**1.07889**
	0.99	21.05423	15.19501	15.04044	15.21423	5.30898	**4.49125**	4.94365	6.32971
100	0.80	0.86713	0.70845	0.70513	0.70890	0.57624	0.62446	0.59440	**0.55351**
	0.85	8.13789	7.25573	7.24732	7.25687	3.78276	3.68556	3.93842	**3.33549**
	0.90	1.63370	1.27682	1.26926	1.27787	0.86396	0.96654	0.90639	**0.82343**
	0.95	3.30935	2.48081	2.46355	2.48306	1.10766	1.32819	1.19400	**0.99553**
	0.99	17.69790	12.19990	12.01945	12.22115	4.07152	**3.35406**	3.75628	4.89893
150	0.80	0.53353	0.45285	0.45167	0.45303	0.41222	0.43623	0.42085	**0.40114**
	0.85	0.70160	0.58071	0.57873	0.58101	0.50618	0.54184	0.51961	**0.49153**
	0.90	1.03631	0.82585	0.82231	0.82633	0.65433	0.71547	0.67861	**0.63117**
	0.95	2.23317	1.64510	1.63026	1.64696	0.88909	1.04210	0.95103	**0.82013**
	0.99	10.51072	7.21385	7.14283	7.22402	1.90923	1.85941	**1.85729**	1.96468
200	0.80	0.70802	0.63292	0.63199	0.63306	0.49058	0.52780	0.50182	**0.46530**
	0.85	0.49989	0.42716	0.42624	0.42730	0.39558	0.41636	0.40303	**0.38642**
	0.90	0.72556	0.59776	0.59603	0.59800	0.52055	0.55669	0.53441	**0.50775**
	0.95	1.39617	1.06106	1.05395	1.06200	0.73659	0.82938	0.77482	**0.70726**
	0.99	8.72502	6.18236	6.12310	6.18965	**1.68437**	1.76663	1.70359	1.79345

TABLE 8: MSEs of MLE, GRR, GKL, GOK, and GDK estimators when $p = 9, \phi = 1$

n	ρ	MLE	GRR	GKL	GOK	GDK.1	GDK.2	GDK.3	GDK.4
50	0.80	1.00699	0.93428	0.93340	0.93450	0.70245	0.75532	0.72308	**0.65787**
	0.85	1.59030	1.47289	1.47128	1.47325	1.01098	1.10053	1.04691	**0.93365**
	0.90	2.29757	2.08446	2.08122	2.08525	1.15422	1.30018	1.21450	**1.06917**
	0.95	3.58637	3.20386	3.19746	3.20558	1.28574	1.51892	1.37736	**1.15613**
	0.99	19.09173	16.63960	16.58982	16.65077	3.95165	**3.55262**	3.74851	4.66415
75	0.80	0.65426	0.62329	0.62305	0.62336	0.50399	0.53331	0.51531	**0.48112**
	0.85	0.80039	0.76123	0.76091	0.76132	0.58597	0.62655	0.60177	**0.55299**
	0.90	1.19824	1.12708	1.12641	1.12727	0.79315	0.85824	0.81988	**0.74752**
	0.95	2.36958	2.17780	2.17563	2.17839	1.06467	1.22927	1.13285	**0.94918**
	0.99	10.95984	9.86672	9.85084	9.87084	**1.87684**	1.99851	1.89034	1.94815
100	0.80	0.44200	0.42778	0.42772	0.42781	0.36421	0.38141	0.37053	**0.34995**
	0.85	5.62867	5.34882	5.26119	10.92247	2.97445	3.38652	3.08617	**2.57719**
	0.90	0.86330	0.82523	0.82500	0.82530	0.61501	0.65990	0.63306	**0.58278**
	0.95	1.74367	1.64277	1.64204	1.64300	0.93508	1.05349	0.98488	**0.85362**
	0.99	9.31641	8.44352	8.42980	8.44726	**1.51464**	1.64422	1.53594	1.54726
150	0.80	0.28489	0.27926	0.27924	0.27926	0.25071	0.25884	0.25360	**0.24322**
	0.85	0.37830	0.36986	0.36983	0.36987	0.32127	0.33421	0.32602	**0.30968**
	0.90	0.55656	0.54006	0.54000	0.54008	0.43896	0.46353	0.44828	**0.41870**
	0.95	1.19794	1.14377	1.14340	1.14389	0.72604	0.80417	0.75785	**0.66914**
	0.99	5.64920	5.27784	5.27467	5.27891	1.28737	1.56290	1.38805	**1.17071**
200	0.80	0.19858	0.19610	0.19609	0.19610	0.18202	0.18615	0.18347	**0.17758**
	0.85	0.26075	0.25683	0.25683	0.25684	0.23210	0.23905	0.23459	**0.22554**
	0.90	0.39032	0.38238	0.38236	0.38239	0.33018	0.34366	0.33522	**0.31897**
	0.95	0.78995	0.76520	0.76511	0.76523	0.56525	0.60763	0.58237	**0.53622**
	0.99	4.41762	4.12353	4.12093	4.12440	1.33021	1.54700	1.42128	**1.24463**

TABLE 9: MSEs of MLE, GRR, GKL, GOK, and GDK estimators when $p = 9, \phi = 1.25$

n	ρ	MLE	GRR	GKL	GOK	GDK.1	GDK.2	GDK.3	GDK.4
50	0.80	0.60770	0.51246	0.51063	0.51268	0.47419	0.50043	0.48353	**0.45737**
	0.85	0.96464	0.80132	0.79812	0.80169	0.69358	0.74044	0.71159	**0.66631**
	0.90	1.39444	1.10142	1.09541	1.10226	0.84549	0.92447	0.87813	**0.81964**
	0.95	2.15874	1.62083	1.60760	1.62253	1.01203	1.15029	1.06933	**0.96892**
	0.99	11.60600	8.10648	8.00744	8.11996	2.16641	2.24890	**2.16401**	2.21763
75	0.80	0.40107	0.35106	0.35039	0.35116	0.33314	0.34740	0.33813	**0.32536**
	0.85	0.49006	0.42337	0.42240	0.42350	0.39223	0.41241	0.39932	**0.38104**
	0.90	0.73793	0.61948	0.61767	0.61974	0.54661	0.58027	0.55955	**0.53261**
	0.95	1.46088	1.13243	1.12598	1.13349	0.80118	0.89628	0.84010	**0.76301**
	0.99	6.75575	4.76170	4.71132	4.76847	1.24071	1.48601	1.32241	**1.17545**
100	0.80	0.27466	0.24831	0.24805	0.24835	0.23991	0.24793	0.24256	**0.23490**
	0.85	3.47702	3.25475	3.25319	3.25486	1.95277	2.21034	2.02171	**1.68182**
	0.90	0.53605	0.46058	0.45964	0.46072	0.41975	0.44267	0.42808	**0.40977**
	0.95	1.07931	0.87424	0.87128	0.87471	0.68222	0.74666	0.70790	**0.65818**
	0.99	5.77598	4.02337	3.96946	4.03072	1.05223	1.27477	1.12810	**1.00143**
150	0.80	0.17766	0.16524	0.16516	0.16525	0.16200	0.16578	0.16321	**0.15926**
	0.85	0.23749	0.21765	0.21750	0.21767	0.21113	0.21728	0.21316	**0.20713**
	0.90	0.34831	0.31013	0.30981	0.31018	0.29390	0.30586	0.29794	**0.28740**
	0.95	0.74903	0.62173	0.61981	0.62200	0.52084	0.56244	0.53637	**0.50385**
	0.99	3.53718	2.60012	2.58337	2.60277	1.01439	1.22175	1.09767	**0.98457**
200	0.80	0.12456	0.11853	0.11850	0.11853	0.11711	0.11896	0.11769	**0.11549**
	0.85	0.16379	0.15355	0.15350	0.15356	0.15066	0.15386	0.15167	**0.14823**
	0.90	0.24565	0.22490	0.22477	0.22493	0.21749	0.22395	0.21963	**0.21348**
	0.95	0.49750	0.43084	0.43022	0.43095	0.39036	0.41202	0.39811	**0.38132**
	0.99	2.52565	1.87688	1.86334	1.87878	0.89999	1.07031	0.97037	**0.87037**

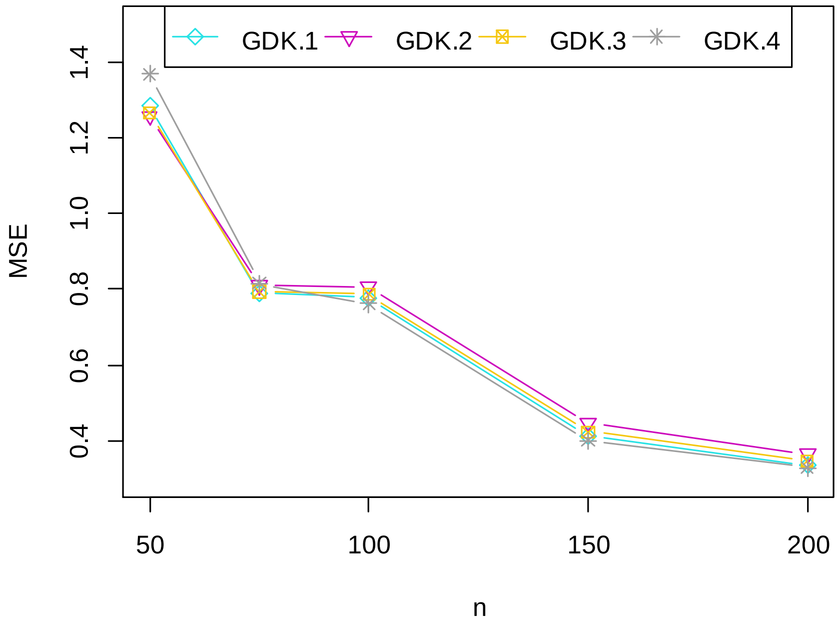

FIGURE 1: Average MSE of the GDK estimator using four different k across the sample size (n)

From Fig. 1, we can note that all four estimators show a consistent decrease in MSE as the sample size increases from 50 to 200, indicating improved accuracy with larger datasets. The rate of MSE reduction is most dramatic between $n = 50$ and $n = 100$, with more gradual improvements for larger sample sizes. The estimators perform very similarly, with only small differences in MSE across the range of sample sizes. GDK.4 appears to have a slightly higher MSE at the smallest sample size ($n = 50$) but performs well when $n \geq 100$. For sample sizes of 100 and higher, GDK.4 shows equivalent or slightly superior performance compared to the other estimators.

Fig. 2 illustrates the performance of four GDK estimators (GDK.1, GDK.2, GDK.3, and GDK.4) across varying degrees of multicollinearity (ρ). As ρ increases, indicating stronger correlations among explanatory variables, the MSE generally rises for all estimators. Notably, GDK.4 demonstrates superior performance across most levels of ρ, particularly as $\rho = 0.99$. This suggests that GDK.4 is more robust to multicollinearity issues, which are common in real-world datasets. The divergence in performance becomes more pronounced at higher levels of multicollinearity, underscoring GDK.4's efficacy in handling highly correlated explanatory variables.

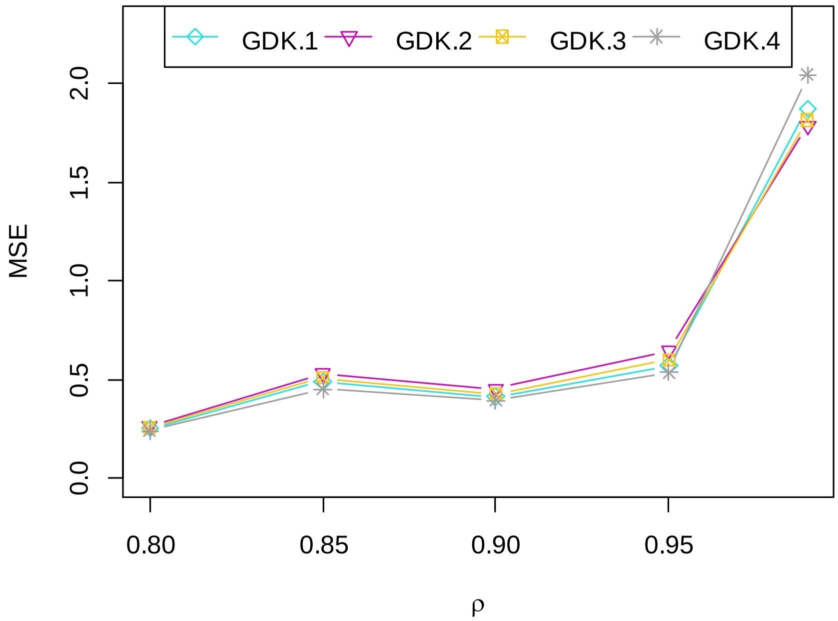

FIGURE 2: Average MSE of the GDK estimator using four different *k* across the degree of multicollinearity (ρ)

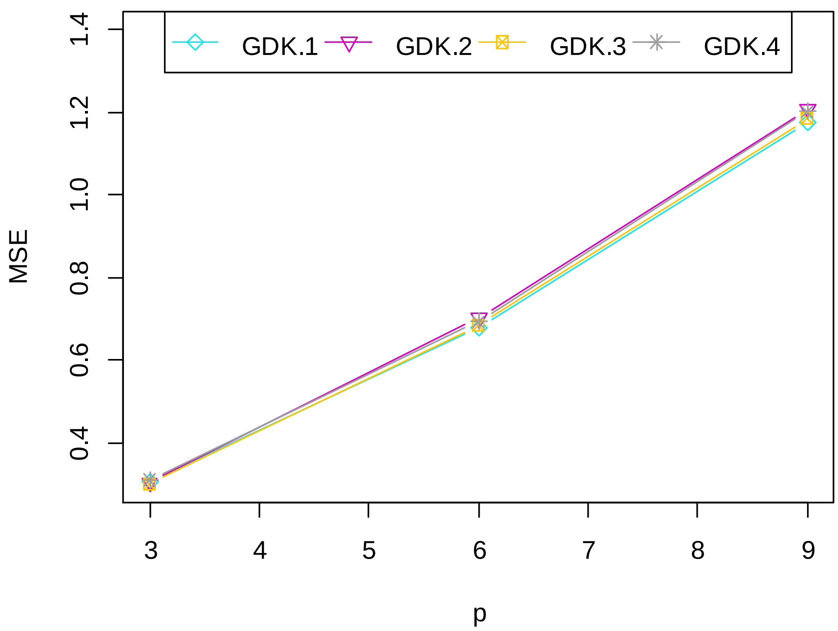

FIGURE 3: Average MSE of the GDK estimator using four different *k* across the number of regression parameters (p)

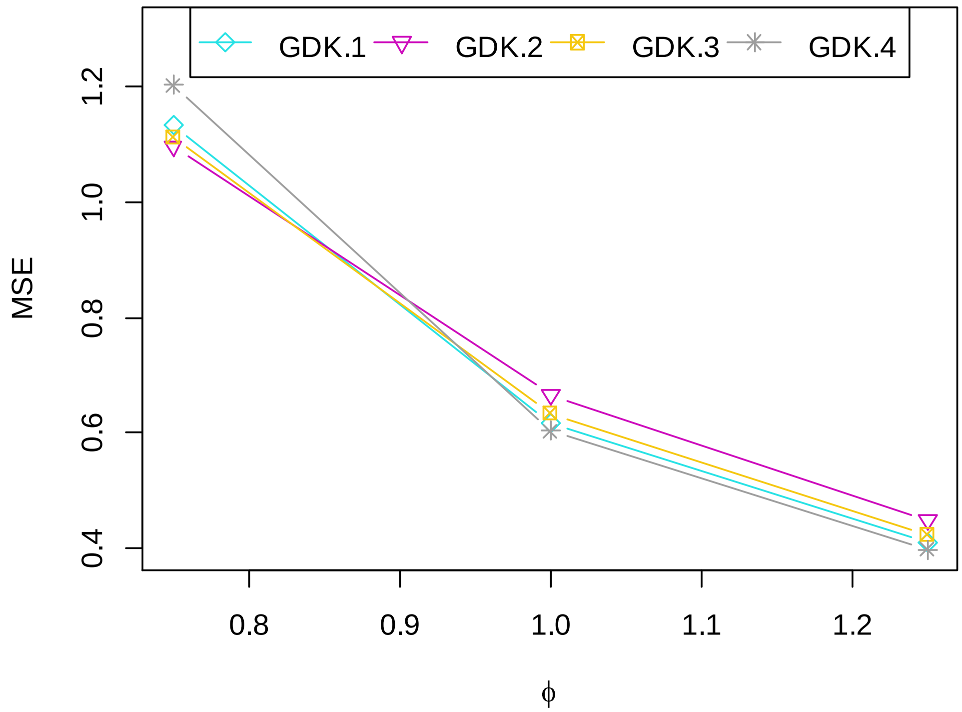

FIGURE 4: Average MSE of the GDK estimator using four different k across the three levels of dispersion parameter (ϕ)

Fig. 3 shows the behavior of GDK estimators as the number of regression parameters (p) increases. All four estimators (GDK.1, GDK.2, GDK.3, and GDK.4) show an upward trend in MSE as p increases. GDK.4 consistently outperforms the other estimators across all values of p, with its advantage becoming more pronounced as the number of regression parameters increases. This suggests that GDK.4 is particularly well-suited for high-dimensional settings, maintaining relatively lower MSE even as the model complexity grows. The stability of GDK.4's performance relative to the other estimators highlights its potential value in analyzing datasets with many explanatory variables.

Fig. 4 examines the impact of the dispersion parameter (ϕ) on the performance of four GDK estimators (GDK.1, GDK.2, GDK.3, and GDK.4). The graph shows three distinct levels of ϕ, representing different degrees of data dispersion. Across all levels, GDK.4 maintains the lowest MSE compared to the other estimators. The performance gap is particularly noticeable at higher levels of dispersion, suggesting that GDK.4 is more robust to variations in data spread. This characteristic makes GDK.4 a potentially valuable tool in scenarios where data overdispersion is a concern, as it maintains its estimation accuracy more effectively than its counterparts.

In all four scenarios examined by the four figures above, GDK.4 consistently emerges as the superior estimator, demonstrating robust performance across varying the sample sizes, degrees of multicollinearity, numbers of explanatory variables, and levels of data dispersion. This consistent outperformance underscores the potential advantages of employing GDK.4 in a wide range of statistical modeling contexts.

4. Application to Body Fat Data

Within this section, we examined a dataset to demonstrate the discoveries outlined in the paper. The dataset referred to as "body fat data" was initially examined in [32] and subsequently utilized in gamma regression modeling by [33], [34]. This dataset consists of 71 observations of healthy female subjects. There are 9 explanatory variables: age in years (age), waist circumference (waistcirc), hip circumference (hipcirc), breadth of the elbow (elbowbreadth), breadth of the knee (kneebreadth), and the sum of logarithms of three anthropometric measurements divided into four groups (anthro3a, anthro3b, anthro3c, and anthro4). Asar and Korkmaz [34] employed the Anderson-Darling (AD) test to examine the distribution of the response variable (DEXfat), which represents body fat. They discovered that the AD test-statistic yielded a value of 0.361, with a corresponding p-value of 0.886. These results indicate that the gamma regression model is appropriate for analyzing this dataset.

Three diagnostic measures (Pearson correlation coefficients, variance inflation factor (VIF), and condition number (CN)) revealed significant multicollinearity among the nine explanatory variables. The correlation matrix (see Fig. 5) shows particularly strong associations (0.88-0.98) between anthropometric measures (anthro3a, anthro3b, anthro3c, anthro4), with waist circumference also demonstrating high correlations (0.73-0.87) with hip circumference and knee breadth. These values, especially those exceeding the 0.8 threshold, confirm severe multicollinearity concerns that warrant remediation in subsequent analyses.

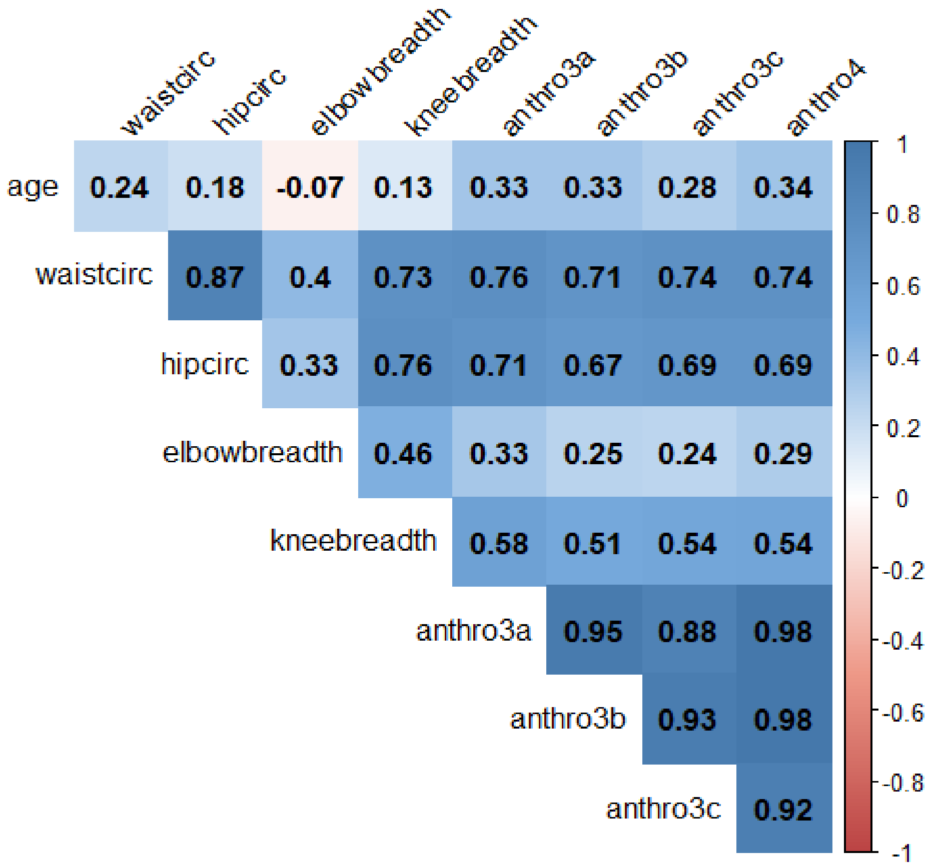

FIGURE 5: Pearson correlation coefficient matrix heatmap

We fit the model including the intercept term and compute the CN based on eigenvalues of $(X'\widehat{W}X)$ matrix: $CN = \sqrt{\gamma_{max}/\gamma_{min}} = 3394.3$, where γ_{max} is the maximum eigenvalue and γ_{min} is the minimum eigenvalue. Since the value of CN (3394.3) is greater than 30, it strongly suggests the presence of multicollinearity. The VIF analysis shows significant multicollinearity issues in the dataset, with variables categorized into three distinct groups based on their VIF values:

1. Severe Multicollinearity (VIF > 10):
   - anthro4 (108.03), anthro3b (46.87), and anthro3a (39.04) exhibit extremely high VIF values, indicating severe multicollinearity that could substantially distort regression results.
2. Moderate to High Multicollinearity (5 < VIF ≤ 10):
   - anthro3c (8.91), hipcirc (6.45), and waistcirc (5.63) demonstrate VIF values above the threshold of 5, suggesting moderate but still problematic multicollinearity.
3. Low Multicollinearity (VIF ≤ 5):

- age (1.21), elbowbreadth (1.45), and kneebreadth (3.18) show VIF values well below 5, indicating minimal multicollinearity concerns.

Given that VIF values above 5 indicate multicollinearity issues, many variables in this dataset exhibit this problem.

TABLE 10: Coefficients and asymptotic MSE (AMSE) values of each estimator

Variable	MLE	GRR	GKL	GOK	GDK.1	GDK.2	GDK.3	GDK.4
Intercept	-0.21947	-0.02451	0.04028	-0.04273	-0.09162	-0.11889	-0.10570	-0.05397
age	0.00161	0.00150	0.00141	0.00150	0.00150	0.00152	0.00151	0.00147
waistcirc	0.00420	0.00491	0.00508	0.00482	0.00464	0.00455	0.00460	0.00477
hipcirc	0.01077	0.01032	0.00987	0.01029	0.01029	0.01039	0.01034	0.01016
elbowbreadth	0.01412	0.00034	-0.00582	0.00130	0.00407	0.00619	0.00516	0.00117
kneebreadth	0.04256	0.03527	0.03511	0.03645	0.03914	0.03989	0.03953	0.03808
anthro3a	-0.12272	0.00410	0.02217	-0.01207	-0.05477	-0.06949	-0.06240	-0.03400
anthro3b	0.14043	0.10533	0.11841	0.11244	0.13603	0.13783	0.13703	0.13266
anthro3c	0.12918	0.13041	0.14245	0.13325	0.13632	0.13482	0.13555	0.13833
anthro4	0.16368	0.09127	0.06183	0.09736	0.11086	0.12164	0.11640	0.09636
AMSE	0.12907	0.06738	0.11078	0.06014	0.05434	0.05892	0.05598	0.05737

Table 10 compares the performance of MLE, GRR, GKL, GOK, and GDK estimators by presenting their coefficients and asymptotic MSE (AMSE) values calculated using equations (8), (11), (14), (17), and (20) respectively. The results demonstrate the clear superiority of the proposed GDK estimators, with GDK.1 achieving the lowest AMSE (0.05434) and GDK.3 closely following (0.05598), while GDK.2 and GDK.4 also show strong performance (0.05737-0.05892). In contrast, traditional estimators perform significantly worse, with GRR and GOK showing moderate results (0.06014-0.06738 AMSE), and GKL (0.11078) and MLE (0.12907) exhibiting the poorest performance, confirming the GDK estimators as more accurate and stable alternatives for this dataset.

5. CONCLUSIONS

This paper introduces the GDK estimator, a novel two-parameter shrinkage estimator proposed to address multicollinearity in the gamma regression model. Through comprehensive Monte Carlo simulations and real-world application, the GDK estimator demonstrates superior performance over MLE, GRR, GKL, and GOK estimators, consistently achieving lower MSE values. The empirical results validate the theoretical advantages of the GDK estimator, particularly in handling highly collinear data. These findings strongly support adopting the GDK estimator as the preferred choice for gamma regression analysis when multicollinearity is present, offering improved estimation accuracy and reliability compared to existing estimators. As future work, we can develop new robust one-parameter and two-parameter GRM estimators to deal with outliers and multicollinearity problems together as suggested by [25], [35], [36], [37], [38], [39], [40], [41], [42] in other regression models.

CONFLICT OF INTERESTS

The author(s) declare that there is no conflict of interests.

REFERENCES

[1] M. Amin, M. Qasim, M. Amanullah, Performance of Asar and Genç and Huang and Yang's Two-Parameter Estimation Methods for the Gamma Regression Model, Iran. J. Sci. Technol. Trans. A: Sci. 43 (2019), 2951–2963. https://doi.org/10.1007/s40995-019-00777-3.

[2] A.M. Al-Abood, D.H. Young, Improved Deviance Goodness of Fit Statistics for a Gamma Regression Model, Commun. Stat. – Theory Methods 15 (1986), 1865–1874. https://doi.org/10.1080/03610928608829223.

[3] M.W. Hattab, A Derivation of Prediction Intervals for Gamma Regression, J. Stat. Comput. Simul. 86 (2016), 3512–3526. https://doi.org/10.1080/00949655.2016.1169421.

[4] Z.Y. Algamal, Developing a Ridge Estimator for the Gamma Regression Model, J. Chemom. 32 (2018), e3054. https://doi.org/10.1002/cem.3054.

[5] E. Dunder, S. Gumustekin, M.A. Cengiz, Variable Selection in Gamma Regression Models via Artificial Bee Colony Algorithm, J. Appl. Stat. 45 (2018), 8–16. https://doi.org/10.1080/02664763.2016.1254730.

[6] A.E. Hoerl, R.W. Kennard, Ridge Regression: Biased Estimation for Nonorthogonal Problems, Technometrics 12 (1970), 55–67. https://doi.org/10.1080/00401706.1970.10488634.

[7] B. Segerstedt, On Ordinary Ridge Regression in Generalized Linear Models, Commun. Stat. – Theory Methods 21 (1992), 2227–2246. https://doi.org/10.1080/03610929208830909.

[8] R.L. Schaefer, L.D. Roi, R.A. Wolfe, A Ridge Logistic Estimator, Commun. Stat. – Theory Methods 13 (1984), 99–113. https://doi.org/10.1080/03610928408828664.

[9] K. Månsson, G. Shukur, A Poisson Ridge Regression Estimator, Econ. Model. 28 (2011), 1475–1481. https://doi.org/10.1016/j.econmod.2011.02.030.

[10] N.K. Rashad, N.M. Hammood, Z.Y. Algamal, Generalized Ridge Estimator in Negative Binomial Regression

Model, J. Phys.: Conf. Ser. 1897 (2021), 012019. https://doi.org/10.1088/1742-6596/1897/1/012019.

[11] R. Tharshan, P. Wijekoon, Ridge Estimator in a Mixed Poisson Regression Model, Commun. Stat. – Simul. Comput. 53 (2024), 3253–3270. https://doi.org/10.1080/03610918.2022.2101064.

[12] K. Månsson, On Ridge Estimators for the Negative Binomial Regression Model, Econ. Model. 29 (2012), 178–184. https://doi.org/10.1016/j.econmod.2011.09.009.

[13] M.R. Abonazel, A.R.R. Alzahrani, A.A. Saber, et al. Developing Ridge Estimators for the Extended Poisson-Tweedie Regression Model: Method, Simulation, and Application, Sci. Afr. 23 (2024), e02006. https://doi.org/10.1016/j.sciaf.2023.e02006.

[14] M.R. Abonazel, I.M. Taha, Beta Ridge Regression Estimators: Simulation and Application, Commun. Stat. – Simul. Comput. 52 (2023), 4280–4292. https://doi.org/10.1080/03610918.2021.1960373.

[15] B.M.G. Kibria, A.F. Lukman, A New Ridge-Type Estimator for the Linear Regression Model: Simulations and Applications, Scientifica 2020 (2020), 9758378. https://doi.org/10.1155/2020/9758378.

[16] A.F. Lukman, I. Dawoud, B.M.G. Kibria, Z.Y. Algamal, B. Aladeitan, A New Ridge-Type Estimator for the Gamma Regression Model, Scientifica 2021 (2021), 5545356. https://doi.org/10.1155/2021/5545356.

[17] I. Dawoud, B.M.G. Kibria, A New Biased Estimator to Combat the Multicollinearity of the Gaussian Linear Regression Model, Stats 3 (2020), 526–541. https://doi.org/10.3390/stats3040033.

[18] M.R. Abonazel, A New Biased Estimation Class to Combat the Multicollinearity in Regression Models: Modified Two--Parameter Liu Estimator, Comput. J. Math. Stat. Sci. 4 (2025), 316–347. https://doi.org/10.21608/cjmss.2025.347818.1096.

[19] M.R. Özkale, S. Kaçiranlar, The Restricted and Unrestricted Two-Parameter Estimators, Commun. Stat. - Theory Methods 36 (2007), 2707–2725. https://doi.org/10.1080/03610920701386877.

[20] J.M. Hardin, J. Hilbe, Generalized Linear Models and Extensions, Stata Press, 2001.

[21] F. Kurtoğlu, M.R. Özkale, Liu Estimation in Generalized Linear Models: Application on Gamma Distributed Response Variable, Stat. Papers 57 (2016), 911–928. https://doi.org/10.1007/s00362-016-0814-3.

[22] Z.Y. Algamal, Y. Asar, Liu-Type Estimator for the Gamma Regression Model, Commun. Stat. – Simul. Comput. 49 (2020), 2035–2048. https://doi.org/10.1080/03610918.2018.1510525.

[23] E. Ertan, A. Erkoç, K.U. Akay, A New Liu-Type Estimator for the Gamma Regression Model, Commun. Stat. – Simul. Comput. (2023). https://doi.org/10.1080/03610918.2023.2220999.

[24] A.E. Hoerl, R.W. Kannard, K.F. Baldwin, Ridge Regression:Some Simulations, Commun. Stat. 4 (1975), 105–123. https://doi.org/10.1080/03610927508827232.

[25] I. Dawoud, M.R. Abonazel, Robust Dawoud–Kibria Estimator for Handling Multicollinearity and Outliers in the Linear Regression Model, J. Stat. Comput. Simul. 91 (2021), 3678–3692. https://doi.org/10.1080/00949655.2021.1945063.

[26] N. Afzal, M. Amanullah, Dawoud–Kibria Estimator for the Logistic Regression Model: Method, Simulation and Application, Iran. J. Sci. Technol. Trans. A: Sci. 46 (2022), 1483–1493. https://doi.org/10.1007/s40995-022-01354-x.

[27] M.R. Abonazel, I. Dawoud, F.A. Awwad, et al. Dawoud–Kibria Estimator for Beta Regression Model: Simulation and Application, Front. Appl. Math. Stat. 8 (2022), 775068. https://doi.org/10.3389/fams.2022.775068.

[28] M.R. Abonazel, I. Dawoud, F.A. Awwad, et al. New Estimators for the Probit Regression Model with Multicollinearity, Sci. Afr. 19 (2023), e01565. https://doi.org/10.1016/j.sciaf.2023.e01565.

[29] A. Kamel, M.R. Abonazel, A Simple Introduction to Regression Modeling Using R, Comput. J. Math. Stat. Sci. 2 (2023), 52–79. https://doi.org/10.21608/cjmss.2023.189834.1002.

[30] M.M. Abdelwahab, M.R. Abonazel, A.T. Hammad, A.M. El-Masry, Modified Two-Parameter Liu Estimator for Addressing Multicollinearity in the Poisson Regression Model, Axioms 13 (2024), 46. https://doi.org/10.3390/axioms13010046.

[31] M.R. Abonazel, A.A. Saber, F.A. Awwad, Kibria–Lukman Estimator for the Conway–Maxwell Poisson Regression Model: Simulation and Applications, Sci. Afr. 19 (2023), e01553. https://doi.org/10.1016/j.sciaf.2023.e01553.

[32] A.L. Garcia, K. Wagner, T. Hothorn, et al. Improved Prediction of Body Fat by Measuring Skinfold Thickness, Circumferences, and Bone Breadths, Obesity Res. 13 (2005), 626–634. https://doi.org/10.1038/oby.2005.67.

[33] O. Reangsephet, S. Lisawadi, S.E. Ahmed, Adaptive Estimation Strategies in Gamma Regression Model, J. Stat. Theory Pract. 14 (2020), 8. https://doi.org/10.1007/s42519-019-0076-1.

[34] Y. Asar, M. Korkmaz, Almost Unbiased Liu-Type Estimators in Gamma Regression Model, J. Comput. Appl. Math. 403 (2022), 113819. https://doi.org/10.1016/j.cam.2021.113819.

[35] N. Yasmin, B.M.G. Kibria, Performance of Some Improved Estimators and Their Robust Versions in Presence of Multicollinearity and Outliers, Sankhya B (2025). https://doi.org/10.1007/s13571-025-00352-4.

[36] I. Dawoud, F.A. Awwad, E.T. Eldin, et al. New Robust Estimators for Handling Multicollinearity and Outliers in the Poisson Model: Methods, Simul. Appl. Axioms 11 (2022), 612. https://doi.org/10.3390/axioms11110612.

[37] A.A. El-Sheikh, M.C. Ali, M.R. Abonazel, Development of Two Methods for Estimating High-Dimensional Data in the Case of Multicollinearity and Outliers, Int. J. Anal. Appl. 22 (2024), 187. https://doi.org/10.28924/2291-8639-22-2024-187.

[38] F.A. Awwad, I. Dawoud, M.R. Abonazel, Development of Robust Özkale–Kaçiranlar and Yang–Chang Estimators for Regression Models in the Presence of Multicollinearity and Outliers, Concurr. Comput. Pract. Exp. 34 (2022), e6779. https://doi.org/10.1002/cpe.6779.

[39] M. Işılar, Y.M. Bulut, Robust Liu-type Estimator Based on GM Estimator, Stat. Neerlandica 78 (2024), 167–190. https://doi.org/10.1111/stan.12310.

[40] W.B. Altukhaes, M. Roozbeh, N.A. Mohamed, Robust Liu Estimator Used to Combat Some Challenges in Partially Linear Regression Model by Improving LTS Algorithm Using Semidefinite Programming, Mathematics 12 (2024), 2787. https://doi.org/10.3390/math12172787.

[41] S. Ahmad, A. Majid, M. Aslam, On Some Robust Liu Estimators for the Linear Regression Model with Outliers: Theory, Simulation and Application, J. Stat. Theory Pract. 18 (2024), 49. https://doi.org/10.1007/s42519-024-00404-4.

[42] W.B. Altukhaes, M. Roozbeh, N.A. Mohamed, Feasible Robust Liu Estimator to Combat Outliers and Multicollinearity Effects in Restricted Semiparametric Regression Model, AIMS Math. 9 (2024), 31581–31606. https://doi.org/10.3934/math.20241519.