Migration Letters

Volume: 21, No: S12 (2024), pp. 130-152

ISSN: 1741-8984 (Print) ISSN: 1741-8992 (Online)

www.migrationletters.com

Liu And Bias-Corrected Liu Estimator For The Multinomial Logit Model

Shaista Shumaila[1], Azeem Ali[2], Muhammad Amin[3], Mohammad Hanif[4]

Abstract:

The multinomial logit Liu estimator and the bias-corrected multinomial logit Liu estimator are proposed as solutions to mitigate the issue of multicollinearity in the multinomial logit model. Furthermore, the superior properties of these estimators in terms of mean squared error are presented when compared to both the maximum likelihood estimator and the ridge estimator. The optimal values of the biasing parameter for the proposed estimators [1]are derived. A simulation is conducted to demonstrate the effectiveness of proposed estimators against ridge and traditional MLE using MSE and bias as performance criteria. The performance of estimators is judged by varying different factors such as the number of values, the number of predictors, levels of the response variable, and the multicollinearity levels. The result of the Monte-Carlo simulation and real data applications reveal that proposed estimators have lower MSE and bias compared to the MLE and ridge estimator.

Keywords: *Multinomial logit model, Multicollinearity, Liu estimator, Almost unbiased estimator.*

1. Introduction

In regression models, as Frisch (1934) indicated, multicollinearity exists among predictors when they have high degree of linear dependence. For the multinomial logit model (MNL), the unbiased and efficient estimates of parameters are acquired by utilizing the traditional maximum likelihood estimator (MLE) under the condition that there is no multicollinearity. However, Abonazel and Farghali (2019) identified that the MLE behave unstable, consequently producing unbiased yet inefficient estimations in presence of multicollinearity. Therefore, the confidence intervals are wider, and due to this theoretically important variables become statistically insignificant in the testing of hypothesis (Qasim et al., 2020a).

In the context of regression models, to alleviate the dire impacts of multicollinearity biased estimators are utilized. In presence of multicollinearity, shrinkage estimators have been proposed to generate efficient estimators (Hoerl and Kennard, 1970; Liu, 1993, Saleh et al., 2019, 2022). Of these, the Liu estimator by Liu (1993) and ridge regression by Hoerl and Kennard (1970), stand out as being broadly adopted techniques. Liu estimator and the ridge estimator are distinguished by the fact that the first-mentioned has a nonlinear while the other one has a linear relation with the shrinkage parameter. Thus, as recognized by Qasim et al. (2018), the Liu estimator has the superiority of being simple when selecting the value of biasing

[1]National College of Business Administration and Economics, Pakistan; [2] Department of Statistics, University of Veterinary and Animal Sciences, Pakistan; [3]Department of Statistics, University of Sargodha, Sargodha, Pakistan; [4] Department of Statistics, National College of Business Administration and Economics, Pakistan.

parameter, which makes it a effective technique in real-world applications. In order to model high-dimensional data, Arashi et al. (2022) worked on the Liu regression following a random forest. The multicollinearity under linear predictors link for modeling longitudinal data was recently covered by Taavoni et al. (2023).

For the Bernoulli response model, numerous researchers have proposed biased estimators to tackle the issue of multicollinearity. Schaefer et al., (1984) introduced the logistic ridge estimator, while Schaefer (1986) introduced the logistic Stein estimator. Månsson and Shukur (2011) proposed the shrinkage estimators for the logistic ridge regression, and subsequently, Mansson et al., (2012) proposed the logistic Liu estimator. Ogoke et al. (2013) presented the modified logistic ridge estimator, and Inan and Erdogan (2013) introduced the logistic Liu-Type estimator. Further contributions include the Almost Unbiased Logistic Liu Estimator by Xinfeng (2015), Asar (2016), Asar and Genc (2016), the almost unbiased ridge logistic estimator (AURLE) by Wu and Asar (2016), the optimal generalized logistic estimator by Varathan and Wijekoon (2018), the modified almost unbiased logistic Liu estimator by Varathan and Wijekoon (2021), Qasim et al., (2021) proposed a beta ridge regression estimator, Mustafa et al. (2022) worked on the beta ridge regression with different link functions, Hadia et al. (2022) discussed the logistic ridge estimator with different link functions, and recently, Lukman et al (2023) proposed the robust biased estimators for the Poisson regression model. Recently, there have been few studies published demonstrating the advantages of almost unbiased estimators in generalized linear models. Some noteworthy examples include the works of Qasim et al. (2020b), Alheety et al. (2021), Amin et al. (2022), Sami et al. (2023) and Algamal et al. (2023), among others. However, the literature on the biased estimation methods for the Multinomial Logistic Regression (MLR) models is limited. Only a limited number of studies have been published that specifically address the issue of multicollinearity, for example, Escabias (2008), Comminatillo and Lucadimo (2008), El-Dash et al. (2011), and Zahid and Tutz (2013). Abonazel and Farghali (2019) introduced a Liu-type multinomial logistic estimator aimed at addressing the challenges posed by multicollinearity.

As far as our knowledge extends, the application of the Liu estimator within the MNL model has not been explored. This paper sets out to introduce both the MNL Liu estimator and a bias-corrected multinomial logit Liu estimator, both adapted to address multicollinearity challenges in the MNL model. The structure of this study is as follows: Section 2 outlines the proposed estimators, while Section 3 presents the shrinkage parameters associated with these estimators. In Section 4, we present Monte Carlo simulation studies to illustrate the theoretical findings. This is followed by Section 5, which showcases the performance of these estimators using real datasets. Finally, Section 6 provides concluding remarks.

2. Methodology:

2.1 Multinomial Logit Model:

The MNL model is the most widely recognized approach for modeling the relationship between the multi-categorical response variable and the set of regressors (So and Kuhfeld, 1995; Zahid and Tutz, 2013). This approach is particularly applicable when dealing with scenarios where there are m different categories of the dependent variable and $m > 2$. The MNL specifies

$$\pi_{ih} = \frac{\exp\,(x_i\beta_h)}{\sum_{h=1}^{m}\exp\,(x_i\beta_h)'}\,, \qquad i = 1,\ldots,n, h = 1,\ldots,m \tag{1}$$

where x_i is the i^{th}row of X which is an $n \times (p+1)$ data matrix with p non-stochastic predictors and $\beta_h$is a $(p+1)\times 1$ vector of regression coefficients. For estimating β_h, the most common method is the MLE, which involves the maximization of the log-likelihood function:

$$l = \sum_{i=1}^{n}\sum_{h=1}^{m} y_{ih}(1-\pi_{ih}). \tag{2}$$

This is carried out by equating the first derivative of Eq. (2) to 0. Thus, the MLE can consequently be acquired by solving the subsequent equation:

$$\frac{\partial l}{\partial \beta_h} = \sum_{i=1}^{n}(y_{ih}-\pi_{ih})x_i = 0. \tag{3}$$

Since Eq.(3) is nonlinear concerning β_h, an iterative weighted least square (IWLS) algorithm is used to solve the above nonlinear function (See Farghali et al., 2023 for more details):

$$\hat{\beta}_h^{(MLE)} = (X'W_hX)^{-1}X'W_hz, \tag{4}$$

where $W_h = \text{diag}(\pi_{ih}(1-\pi_{ih}))$ and z is a vector where the i^{th}element equals $z_{ih} = \log(\pi_{ih}) + \frac{(y_{ih}-\pi_{ih})}{\pi_{ih}(1-\pi_{ih})}$. The asymptotic covariance matrix of the MLE is expressed as follows:

$$\text{Cov}\left(\hat{\beta}_h^{(ML)}\right) = -E\left(\frac{\partial^2 l}{\partial\beta_j\,\partial\beta_k{}'}\right) = (X'W_hX)^{-1}.$$

The asymptotic mean squared error (MSE) equals:

$$\text{MSE}\left(\widehat{\beta_h}^{(ML)}\right) = \text{tr}[(X'W_hX)^{-1}] = \sum_{j=1}^{p+1}\sum_{h=1}^{m}\frac{1}{\lambda_{hj}}, \tag{4a}$$

where λ_{hj} is the j^{th}eigenvalue of the matrix $X'W_hX$. If $X'W_hX$ is ill-conditioned then λ_{hj} tends to zero and the values of $(X'W_hX)^{-1}$ becomes large which gives rise to unstable and inflated MSE of MLE (Chang 2015; Abonazel and Faraghali, 2019, Qasim et al., 2022). As a remedy to this problem, the new estimators are proposed in the next section.

2.2 Proposed Estimators:

2.2.1 Multinomial logit Liu estimator (MLLE)

In the case of multicollinearity, the most common method is the ridge regression estimator by Hoerl and Kennard (1970) and is given by $\hat{\beta}_k = (I + k(X'X)^{-1})^{-1}\hat{\beta}_{MLE}$. The ridge estimator for MNL was proposed by Månsson et al. (2018) which is given by

$$\hat{\beta}^{(RR)} = (X'W_hX + kI)^{-1}(X'W_hX)\hat{\beta}_h^{(MLE)}, \tag{5}$$

where $k = \left(\frac{p+1}{\prod_{j=1}^{p+1}\alpha_{hj}{}^2}\right)^{\frac{1}{p+1}}$. MSE of the ridge estimator for the MNL model is given by

$$\text{MSE}\left(\hat{\beta}^{(RR)}\right) = \sum_{j=1}^{p+1}\sum_{h=1}^{m}\frac{\lambda_{hj}}{(\lambda_{hj}+k)^2} + k^2\sum_{j=1}^{p+1}\sum_{h=1}^{m}\frac{\alpha_{hj}{}^2}{(\lambda_{hj}+k)^2}\,. \tag{6}$$

The Liu estimator for the classical linear regression model was introduced by Liu (1993), and it is presented as follows:

$$\hat{\beta}_d = (X'X + I)^{-1}(X'X + dI)\hat{\beta}_{OLS},$$

where $d(0 \leq d \leq 1)$ is the shrinkage parameter. The Liu estimator has an advantage over the ridge regression estimator that is, it is easier to choose d than to choose k because the Liu estimator is a linear function of d. The Multinomial Logit Liu estimator (MLLE) is suggested as a cure for the issue of inflated variance of the MLE in the case of multicollinearity. As the IWLS algorithm is used to find $\beta_h^{(MLE)}$, it approximately minimizes the weighted sum of squared error (WSSE). Therefore, in the WSSE sense, $\beta_h^{(MLE)}$ can be considered as the optimal estimator. For the MNL model, we propose the following Liu estimator:

$$\hat{\beta}^{(MLLE)} = (X'W_hX + I)^{-1}(X'W_hX + dI)\hat{\beta}_h^{(MLE)}. \tag{7}$$

$$\mathrm{Bias}\left(\hat{\beta}^{(MLLE)}\right) = E\left(\hat{\beta}^{(MLLE)}\right) - \beta_h^{(MLE)}.$$

$$\mathrm{Bias}\left(\hat{\beta}^{(MLLE)}\right) = (X'W_hX + I)^{-1}(X'W_hX + dI)\beta_h^{(MLE)} - \beta_h^{(MLE)}$$

$$= [(X'W_hX + I)^{-1}(X'W_hX + dI) - I]\beta_h^{(MLE)}.$$

$$\mathrm{MSE}\left(\hat{\beta}^{(MLLE)}\right) = \mathrm{var}\left(\hat{\beta}^{(MLLE)}\right) + \mathrm{bias}\left(\hat{\beta}^{(MLLE)}\right)\mathrm{bias}(\hat{\beta}^{(MLLE)})'.$$

$$\mathrm{var}\left(\hat{\beta}^{(MLLE)}\right) = (X'W_hX + I)^{-1}(X'W_hX + dI)(\hat{\beta}^{(MLE)})(X'W_hX + dI)(X'W_hX + I)^{-1}$$

$$= (X'W_hX + I)^{-1}(X'W_hX + dI)(X'W_hX)^{-1}(X'W_hX + dI)(X'W_hX + I)^{-1}.$$

Since the scalar $\mathrm{MSE}\left(\hat{\beta}^{(MLLE)}\right)$ can be obtained by applying the trace operator which is stated as follows:

$$\mathrm{MSE}\left(\hat{\beta}^{(MLLE)}\right) = \sum_{j=1}^{p+1}\sum_{h=1}^{m}\frac{(\lambda_{hj} + d)^2}{\lambda_{hj}(\lambda_{hj} + 1)^2} + (1-d)^2\sum_{j=1}^{p+1}\sum_{h=1}^{m}\frac{\alpha_{hj}{}^2}{\left(\lambda_{hj} + 1\right)^2} \tag{8}$$

$$= \gamma_1(d) + \gamma_2(d),$$

where λ_{hj} are the ordered eigenvalues of the matrix $X'W_hX$ and $\alpha = V'\beta$, where V' is an orthogonal matrix whose columns are eigenvectors of the corresponding eigenvalues λ_{hj}.

2.2.2 Bias-corrected MLLE:

Before proceeding with the development of the bias-corrected Liu estimator for the MNL model, it is essential to establish the definition of the almost unbiased (bias-corrected) estimator:

Definition. (Xu and Yang, 2011 and Amin et al., 2022) Suppose that $\hat{\beta}$ is a biased estimator of β with, $\mathrm{Bias}\left(\hat{\beta}\right) = E\left(\hat{\beta}\right) - \beta = C\beta$, which implies that $E\left(\hat{\beta} - C\beta\right) = \beta$, then the almost unbiased estimator based on a biased estimator $\hat{\beta}$ is defined as $\tilde{\beta} = \hat{\beta} - C\hat{\beta} = (I - C)\hat{\beta}$.

Now, we develop an almost unbiased MLLE (AUMLLE) for the MNL model:

$$\mathrm{Bias}\left(\hat{\beta}^{(MLLE)}\right) = -(1-d)(X'W_hX + I)^{-1}\beta_h^{(MLE)}$$

$$\hat{\beta}^{(AUMLLE)} = \left(I + (X'W_hX + I)^{-1}(1-d)\right)\hat{\beta}^{(MLLE)}$$

$$= [I - (X'W_hX + I)^{-2}(1-d)^2)\hat{\beta}_h^{(MLE)} \tag{9}$$

$$Bias\left(\hat{\beta}^{(AUMLLE)}\right) = -(X'W_hX + I)^{-2}(1-d)^2\hat{\beta}_h^{(MLE)}.$$

$$Var\left(\hat{\beta}^{(AUMLLE)}\right) = [I - (X'W_hX + I)^{-2}(1-d)^2]var(\hat{\beta}_h^{(MLE)})[I - (X'W_hX + I)^{-2}(1-d)^2]$$

$$= [I - (X'W_hX + I)^{-2}(1-d)^2](X'W_hX)^{-1}[I - (X'W_hX + I)^{-2}(1-d)^2].$$

The scalar MSE of the AUMLLE for the MNL model is given by:

$$MSE\left(\hat{\beta}^{(AUMLLE)}\right) = \sum_{j=1}^{p+1}\sum_{h=1}^{m}\frac{(1-\frac{(1-d)^2}{(\lambda_{hj}+1)^2})^2}{\lambda_{hj}} + (1-d)^4\sum_{j=1}^{p+1}\sum_{h=1}^{m}\frac{{\alpha_{hj}}^2}{(\lambda_{hj}+1)^4}$$

$$MSE\left(\hat{\beta}^{(AUMLLE)}\right) = \sum_{j=1}^{p+1}\sum_{h=1}^{m}\frac{\left(\lambda_{hj}+d\right)^2\left(\lambda_{hj}+2-d\right)^2}{\lambda_{hj}\left(\lambda_{hj}+1\right)^4} + (1-d)^4\sum_{j=1}^{p+1}\sum_{j=1}^{m}\frac{{\alpha_{hj}}^2}{\left(\lambda_{hj}+1\right)^4}. \quad (10)$$

where λ_{hj} are the ordered eigenvalues of the matrix $X'W_hX$ and $\alpha = V'\beta$, where V' is an orthogonal matrix whose columns are eigenvectors of the corresponding eigenvalues λ_{hj}.

2.3 Superiority of the new estimators

The performance of the proposed estimators is compared with the MLE. The AUMLLE has a smaller MSE as compared to the MLLE under some conditions. The following lemma is required to prove the theorem.

Lemma 1: Let two linear estimators of α are $\hat{\alpha}_i = A_iX, i = 1,2.$ Suppose that $D = Cov(\hat{\alpha}_1) - Cov(\hat{\alpha}_2)$ is p.d. then $\Delta = MSE(\hat{\alpha}_1) - MSE(\hat{\alpha}_2)$ is n.n.d iff $\acute{a}_2(D + \acute{a}_1a_1)a_2 \leq 1$, where a_j denotes the bias vector of $\hat{\alpha}_j$.

Theorem 1: Under the MNL model, we have $||Bias_{MLLE}||^2 - \left||Bias_{AUMLLE}|\right|^2 > 0$ for $0 < d < 1$.

Proof:

$$||Bias_{MLLE}||^2 - \left||Bias_{AUMLLE}|\right|^2 = (d-1)^2\sum_{j=1}^{p+1}\sum_{h=1}^{m}\frac{{\alpha_{hj}}^2}{(\lambda_{hj}+1)^2} - (1-d)^4\sum_{j=1}^{p+1}\sum_{h=1}^{m}\frac{{\alpha_{hj}}^2}{(\lambda_{hj}+1)^4}$$

$$=(1-d)^2\sum_{j=1}^{p+1}\sum_{h=1}^{m}\left(\frac{{\alpha_{hj}}^2}{(\lambda_{hj}+1)^2} - (1-d)^2\frac{{\alpha_{hj}}^2}{(\lambda_{hj}+1)^4}\right)$$

$$=(1-d)^2\sum_{j=1}^{p+1}\sum_{h=1}^{m}{\alpha_{hj}}^2\left(\frac{(\lambda_{hj}+1)^2-(1-d)^2}{(\lambda_{hj}+1)^4}\right)$$

$$=(1-d)^2\sum_{j=1}^{p+1}\sum_{h=1}^{m}{\alpha_{hj}}^2\left(\frac{(\lambda_{hj}+d)(\lambda_{hj}-d+2)}{(\lambda_{hj}+1)^4}\right).$$

It can be easily seen that the difference of $||Bias_{MLLE}||^2 - \left||Bias_{AUMLLE}|\right|^2$ is positive, when $0 < d < 1$.

2.4 Comparison between MLLE and AUMLLE

Consider

$\Delta = MSE(\widehat{\beta}^{(MLLE)}) - MSE(\widehat{\beta}^{(AUMLLE)}) = D + bias_{MLLE}' bias_{MLLE} - bias_{AUMLLE}' bias_{AUMLLE}$, (11)

where $D = \{(X'W_hX+I)^{-1}(X'W_hX+dI)(X'W_hX)^{-1}(X'W_hX+dI)(X'W_hX+I)^{-1}\} - \{[I-(X'W_hX+I)^{-2}(1-d)^2](X'W_hX)^{-1}[I-(X'W_hX+I)^{-2}(1-d)^2]\}$. The following theorem gives necessary and sufficient conditions for $\widehat{\beta}^{(AUMLLE)}$ to be superior to $\widehat{\beta}^{(MLLE)}$.

Theorem 2: Under the MNL model, if ${\alpha_{hj}}^2 > \frac{\lambda_{hj}+2}{2\lambda_{hj}+3}$, $0 < m < 1$, where $m = 1 - d$, then $\widehat{\beta}^{(AUMLLE)}$ is superior to $\widehat{\beta}^{(MLLE)}$ in the MSE sense namely $MSE(\widehat{\beta}^{(MLLE)}) - MSE(\widehat{\beta}^{(AUMLLE)}) \geq 0$

Proof: From Equations (6) and (8), the difference in scalar MSE is:

$$MSE(\widehat{\beta}^{(MLLE)}) - MSE(\widehat{\beta}^{(AUMLLE)})$$
$$= \sum_{j=1}^{p+1}\sum_{h=1}^{m}\frac{(\lambda_{hj}+d)^2}{\lambda_{hj}(\lambda_{hj}+1)^2} + (1-d)^2\sum_{j=1}^{p+1}\sum_{h=1}^{m}\frac{{\alpha_{hj}}^2}{(\lambda_{hj}+1)^2}$$
$$-\sum_{j=1}^{p+1}\sum_{h=1}^{m}\frac{(\lambda_{hj}+d)^2(\lambda_{hj}+2-d)^2}{\lambda_{hj}(\lambda_{hj}+1)^4} - (1-d)^4\sum_{j=1}^{p+1}\sum_{j=1}^{m}\frac{{\alpha_{hj}}^2}{(\lambda_{hj}+1)^4}.$$

$$= \sum_{j=1}^{p+1}\sum_{h=1}^{m}\frac{\omega_{hj}{\alpha_{hj}}^2(1-\omega_{hj})}{\lambda_{hj}}\{\lambda_{hj}\omega_{hj}(1+\omega_{hj}) - \alpha_{hj}^*(1-\omega_{hj})(2+\omega_{hj})\},$$

where $\omega_{hj} = \frac{1-d}{(\lambda_{hj}+1)}$, and $\alpha_{hj}^* = \frac{1}{{\alpha_{hj}}^2}$.

We can see that $0 < \alpha_{hj}^* = \omega_{hj} < 1$ and $\lambda_{hj} > 0$, then $MSE(\widehat{\beta}^{(MLLE)}) - MSE(\widehat{\beta}^{(AUMLLE)}) \geq 0$ iff $\lambda_{hj}\omega_{hj}(1+\omega_{hj}) - \alpha_{hj}^*(1-\omega_{hj})(2+\omega_{hj}) \geq 0$ or

$(\lambda_{hj}+\alpha_{hj}^*)m^2 + (\lambda_{hj}+\alpha_{hj}^*)(\lambda_{hj}+1)m - 2\alpha_{hj}^*(1-\omega_{hj})(\lambda_{hj}+1)^2 \geq 0$, where $m = 1-d$. $(\lambda_{hj}+\alpha_{hj}^*)m^2 + (\lambda_{hj}+\alpha_{hj}^*)(\lambda_{hj}+1)m - 2\alpha_{hj}^*(1-\omega_{hj})(\lambda_{hj}+1)^2 = 0$

has two roots given by

$$m_1 = \frac{(\lambda_{hj}+1)}{2}\left(-1-\sqrt{\frac{\lambda_{hj}+9\alpha_{hj}^*}{\lambda_{hj}+\alpha_{hj}^*}}\right) \text{ and } m_2 = \frac{(\lambda_{hj}+1)}{2}\left(-1+\sqrt{\frac{\lambda_{hj}+9\alpha_{hj}^*}{\lambda_{hj}+\alpha_{hj}^*}}\right)$$

It is evident that $m_1 < 0$ and $m_1 < m_2$. Since $0 < m < 1$, we acquire ${\alpha_{hj}}^2 > \frac{\lambda_{hj}+2}{2\lambda_{hj}+3}$,

Then for $0 < m < 1$, $MSE(\widehat{\beta}^{(MLLE)}) - MSE(\widehat{\beta}^{(AUMLLE)}) \geq 0$.

3. Selection of shrinkage parameter

We can obtain the optimal value of d by minimizing the MSE of the MLLE and the AUMLLE estimator as follows:

- Taking the first derivative of $MSE\left(\widehat{\beta}^{(MLLE)}\right)$ concerning d and equating it to zero, we get:

$$\sum_{j=1}^{p+1}\sum_{h=1}^{m}\left\{\frac{2}{\lambda_{hj}}\frac{(\lambda_{hj}+d)}{(\lambda_{hj}+1)^2}-\frac{2(1-d)\alpha_{hj}{}^2}{(\lambda_{hj}+1)^4}\right\}=0.$$

Solving the above equation, we get

$$d_{MLLE}=\frac{\sum_{j=1}^{p+1}\sum_{h=1}^{m}\frac{\lambda_{hj}(\alpha_{hj}{}^2-1)}{\lambda_{hj}(\lambda_{hj}+1)^2}}{\sum_{j=1}^{p+1}\sum_{h=1}^{m}\frac{(\lambda_{hj}\alpha_{hj}{}^2+1)}{\lambda_{hj}(\lambda_{hj}+1)^2}}$$

And taking the first derivative of $MSE\left(\beta^{(AUMLLE)}\right)$ concerning d and equating it to 0 we get:

$$\sum_{j=1}^{p+1}\sum_{h=1}^{m}\left\{\frac{4}{\lambda_{hj}}\left(1-\frac{(1-d)^2}{(\lambda_{hj}+1)^2}\right)\frac{(1-d)}{(\lambda_{hj}+1)^2}-4\frac{(1-d)^3\alpha_{hj}{}^2}{(\lambda_{hj}+1)^4}\right\}=0.$$

Solving this equation for d we get:

$$d_{AUMLLE}=1-\sqrt{\frac{\sum_{j=1}^{p+1}\sum_{h=1}^{m}\frac{1}{\lambda_{hj}(\lambda_{hj}+1)^2}}{\sum_{j=1}^{p+1}\sum_{h=1}^{m}\frac{(\lambda_{hj}\alpha_{hj}{}^2+1)}{\lambda_{hj}(\lambda_{hj}+1)^4}}}.$$

respectively. Since d_{MLLE} and d_{AUMLLE} depend on the unknown parameter α_{hj} , so, we replace them with their estimates and the estimated values are

$$\hat{d}_{MLLE}=\frac{\sum_{j=1}^{p+1}\sum_{h=1}^{m}\frac{\lambda_{hj}(\hat{\alpha}_{hj}{}^2-1)}{\lambda_{hj}(\lambda_{hj}+1)^2}}{\sum_{j=1}^{p+1}\sum_{h=1}^{m}\frac{(\lambda_{hi}\hat{\alpha}_{hj}{}^2+1)}{\lambda_{hj}(\lambda_{hj}+1)^2}}$$

And

$$\hat{d}_{AUMLLE}=1-\sqrt{\frac{\sum_{j=1}^{p+1}\sum_{h=1}^{m}\frac{1}{\lambda_{hj}(\lambda_{hj}+1)^2}}{\sum_{j=1}^{p+1}\sum_{h=1}^{m}\frac{(\lambda_{hj}\hat{\alpha}_{hj}{}^2+1)}{\lambda_{hj}(\lambda_{hj}+1)^4}}}$$

4. Monte Carlo Simulation

4.1 Simulation Layout

The simulation study is conducted to assess and compare the performance of the MLLE, and AUMLLE based on d_{MLLE} and d_{AUMLLE} with ML estimator and ridge estimator. In this study, the dependent variable of the MNL model is generated using pseudo-random numbers from the multinomial regression model, where

$$\pi_{ih} = \frac{exp\ (x_i\beta_h)}{\sum_{h=1}^{m} exp\ (x_i\beta_h)}, \qquad i = 1, \ldots, n, h = 1, \ldots, m$$

The parameter values are chosen so that $\beta'\beta = 1$, which is a generally used constraint in this field see Kibria(2003), Asar (2016) and Mannson et al.(2018). Following (Suhail et al., 2020), (Babar & Chand, 2022), and (Wasim et al., 2023), the correlated explanatory variables have been generated as:

$$x_{ij} = (1-\rho^2)^{1/2} z_{ij} + \rho z_{i(j+1)}\ , \qquad i\ = 1,2, \ldots, n, \quad j = 1,2, \ldots, p$$

where z_{li} are the pseudo-random numbers produced from the normal distribution. The factors we choose to vary are the correlation between the predictor variables (ρ), the number of predictor variables (p), the sample size(n), and the levels of the outcome variable (m). The scenarios for the simulation study are summarized in Table 1:

Table 1: Parametric conditions of the simulation layout

n	50	100	150	200	300
ρ	0.80	0.85	0.90	0.95	0.99
p	2	4	8	—	—
m	3	5	—	—	—

To investigate whether the MLLE and AUMLLE estimators are better than the ML estimator, MSE is computed using the following equation:

$$MSE = \frac{\sum_{i=1}^{R}\sum_{j=2}^{m}(\hat{\beta}_j - \beta_j)'(\hat{\beta}_j - \beta_j)}{R},$$

where $\hat{\beta}_j$ is the estimator of β_j obtained from the MLLE, AUMLLE, MLE and ridge estimator. $R = 2000$ indicating the number of replicates used in the Monte Carlo simulation. The values in parentheses in all tables represent the Bias of the estimators. R software was used to run the simulation.

4.2 Results and Discussion:

Tables 2-4 illustrate the simulated MSE values for the proposed estimators and the MLE, considering various factors we deliberately varied. The performance evaluation encompasses AUMLLE, MLLE, ridge, and MLE, with the variations encompassing n (sample size), p (number of explanatory variables), m (levels of the response variable), and ρ (degree of correlation). Specifically, Table 2 showcases MSEs for p =2 with varying n, m, and ρ. Similarly, Table 3 presents MSEs for p=4, and Table 4 depicts MSEs for p=8.

Across all scenarios, the MLE exhibits the highest MSE values in comparison to ridge, MLLE, and AUMLLE, indicative of the superior performance of the proposed estimators. Furthermore, a consistent trend is observed where increased n yields decreased MSE values for all estimators (refer to Figures 1-2), highlighting the significant positive impact of larger sample sizes on estimator performance.

As ρ escalates while keeping m and p constant, the MSE of all estimators increases, signifying a detrimental impact on estimator performance. However, it's noteworthy that the adverse impact of ρ can be mitigated by augmenting the sample size, as evidenced by the simulation outcomes. It's noteworthy that MLE experiences the greatest influence from rising ρ.

Moreover, with increasing levels of the response variable (m), the MSE and Bias values of all estimators rise in tandem, holding ρ and p constant. The escalation in MSE and Bias is particularly pronounced for MLE and ridge estimators when the response variable levels increase. Similarly, augmenting the number of explanatory variables elevates the MSE and Bias across all estimators. The deleterious effect of heightened p is more pronounced when m and ρ take on larger values with a smaller n. Notably, the performance of MLE is compromised as the number of regressors increases. Consequently, for given n, increases in m, ρ, and p inflate the MSE and Bias values for all estimators. Specifically, the highest MSE value is observed when n = 50, p=8, m=5, and ρ = 0.99, while the lowest MSE value occurs when n = 300, p=2, m =3, and ρ =0.80. Overall, the simulation study consistently demonstrates the superior performance of the proposed estimators in comparison to MLE, with MLLE yielding the smallest MSE values across all scenarios.

Table 2: MSE values and biases of MNL estimators when $p = 2$

n	$m = 3$				$m = 5$			
	MLE	Ridge	MLLE	AUMLLE	MLE	Ridge	MLLE	AUMLLE
				$\rho = 0.80$				
50	2.0907	1.1416 (0.6289)	1.0125 (0.5666)	1.0598 (0.5128)	5.2891	2.8027 (0.9881)	2.4184 (0.8689)	2.6624 (0.7395)
100	0.5653	0.3949 (0.2232)	0.3406 (0.2985)	0.3478 (0.2597)	1.8924	1.1292 (0.4383)	0.9963 (0.5121)	1.109 (0.4171)
200	0.4765	0.3181 (0.1702)	0.2735 (0.2561)	0.2738 (0.225)	1.4821	0.8846 (0.3392)	0.7634 (0.4337)	0.8338 (0.354)
250	0.4332	0.2804 (0.1468)	0.2379 (0.2328)	0.2362 (0.2053)	1.0752	0.6773 (0.2494)	0.5727 (0.366)	0.6192 (0.2971)
300	0.4042	0.2252 (0.1152)	0.1969 (0.1956)	0.1924 (0.1735)	1.1327	0.6821 (0.2519)	0.5709 (0.3693)	0.6148 (0.3008)
				$\rho = 0.85$				
50	1.7017	0.9419 (0.4998)	0.8434 (0.4929)	0.877 (0.4314)	6.4725	3.4817 (1.1289)	2.9955 (0.9918)	3.2595 (0.8483)
100	0.9628	0.5725 (0.3228)	0.4997 (0.3703)	0.5154 (0.3256)	3.4445	1.837 (0.6885)	1.6242 (0.6815)	1.7829 (0.5628)
200	0.717	0.419 (0.2306)	0.3598 (0.3061)	0.3634 (0.2713)	2.1759	1.2047 (0.479)	1.055 (0.5351)	1.1602 (0.4424)
250	0.4352	0.27 (0.1397)	0.2307 (0.2251)	0.2284 (0.1983)	1.3229	0.799 (0.3057)	0.6818 (0.4083)	0.7417 (0.3338)

300	0.4188	0.2589 0.2216	0.2130 (0.21)	0.2158 (0.2150)	0.3597	0.2417 (0.1248)	0.2001 (0.2169)	0.1978 (0.1895)
					$\rho = 0.90$			
50	4.8994	2.5521 (1.1471)	2.0503 (0.8676)	2.1133 (0.8135)	8.8407	4.7392 (1.4795)	3.8815 (1.1974)	4.1675 (1.0653)
100	0.5127	0.3402 (0.2014)	0.2847 (0.282)	0.2921 (0.2502)	3.5857	1.9611 (0.7732)	1.7316 (0.738)	1.9274 (0.623)
200	0.5543	0.3589 (0.2164)	0.3006 (0.2926)	0.3095 (0.26)	2.0099	1.1327 (0.4824)	0.9979 (0.5213)	1.1134 (0.4339)
250	0.4774	0.3238 (0.194)	0.271 (0.270)	0.2791 (0.2455)	1.6479	0.9937 (0.3995)	0.8616 (0.4784)	0.9523 (0.3927)
300	0.4108	0.295 (0.1757)	0.2431 (0.2428)	0.2484 (0.2312)	1.7162	0.9966 (0.4123)	0.8684 (0.483)	0.9561 (0.4013)
					$\rho = 0.95$			
50	5.4215	2.8531 (1.2377)	2.2425 (0.9118)	2.313 (0.8622)	26.2441	15.273 (3.1176)	10.8061 (2.1529)	10.9836 (2.0667)
100	2.8559	1.4228 (0.7965)	1.2282 (0.6613)	1.2808 (0.6126)	8.1979	4.3359 (1.4333)	3.5811 (1.1611)	3.8501 (1.0329)
200	1.5001	0.7659 (0.489)	0.6784 (0.4813)	0.7105 (0.4375)	3.3036	1.7726 (0.7368)	1.5667 (0.7114)	1.7476 (0.5997)
250	0.7358	0.4347 (0.2751)	0.3693 (0.336)	0.3837 (0.3001)	3.2763	1.7429 (0.731)	1.5373 (0.7021)	1.7186 (0.5916)
300	0.7651	0.4615 (0.2981)	0.3959 (0.3514)	0.4143 (0.3179)	2.4459	1.3422 (0.5899)	1.1905 (0.6034)	1.3353 (0.5078)
					$\rho = 0.99$			
50	22.0622	13.6848 (3.0794)	8.4651 (1.8991)	8.4749 (1.878)	100.4386	69.639 (7.1928)	(40.6852) (4.3125)	40.6568 (4.2782)
100	12.2971	6.8417 (2.1547)	4.7819 (1.4182)	4.8103 (1.3885)	39.9562	24.2575 (4.1342)	16.7709 (2.7741)	16.7553 (2.7143)
200	5.7754	2.8531 (1.3069)	2.289 (0.9696)	2.3419 (0.9297)	19.4418	10.854 (2.6096)	8.2149 (1.9246)	8.3365 (1.8297)

250	5.1394	2.5407 (1.2163)	2.0629 (0.9151)	2.1109 (0.8725)	15.9589	8.6349 (2.2806)	6.626 (1.7029)	6.8042 (1.5969)
300	3.9958	2.0245 (1.0565)	1.7149 (0.8367)	1.759 (0.7907)	13.4225	7.2117 (2.0416)	5.6077 (1.5516)	5.8361 (1.441)

Note: The bias values are given in parentheses.

Table 3: MSE values and biases of MNL estimators when $p = 4$

n		$m = 3$				$m = 5$		
	MLE	Ridge	MLLE	AUMLE	MLE	Ridge	MLLE	AUMLE
					$\rho = 0.80$			
50	4.7167	2.7177 (1.0284))	2.3574 (0.8806)	2.5372 (0.8081)	17.1145	10.653 2.4548	8.0486 1.6847	8.7291 1.511
100	2.3636	1.4012 (0.585)	1.2513 (0.6184)	1.3788 (0.515)	6.8466	3.9878 (1.1905)	3.5429 (1.0471)	3.9794 (0.8017)
200	0.85	0.6124 (0.2071)	0.5167 (0.3774)	0.551 (0.3186)	4.0409	2.4518 (0.7142)	2.2452 (0.8104)	2.5307 (0.678)
250	0.7354	0.5404 (0.1836)	0.4507 (0.3451)	0.4838 (0.2814)	2.0135	2.0135 (1.3657)	1.1799 (0.5601)	1.3259 (0.4631)
300	0.5785	0.4486 (0.1442)	0.3707 (0.3075)	0.3925 (0.2515)	1.7473	1.2002 (0.3226)	1.0263 (0.5079)	1.1674 (0.4007)
					$\rho = 0.85$			
50	5.842	3.4399 (1.2534)	2.8516 (0.9936)	3.0747 (0.9025)	26.3351	17.2116 (3.336)	11.9801 (2.1181)	12.7941 (1.9598)
100	2.3849	1.4086 (0.5785)	1.2584 (0.6185)	1.3804 (0.5252)	9.5108	5.4412 (1.5027)	4.8248 (1.28)	5.3596 (1.121)
200	1.294	0.8458 (0.331)	0.7309 (0.4604)	0.804 (0.3741)	3.7664	2.3116 (0.6907)	2.1066 (0.7862)	2.3994 (0.6521)

250	0.9075	0.6294 (0.2283)	0.5297 (0.3789)	0.5779 (0.3035)	2.8511	1.7927 (0.5248)	1.6013 (0.6608)	1.8268 (0.5328)
300	0.7238	0.5298 (0.1818)	0.4402 (0.3414)	0.475 (0.2751)	2.4764	1.5761 (0.463)	1.3957 (0.6047)	1.5931 (0.4806)
					$\rho = 0.90$			
50	10.3011	6.3256 (1.9254)	4.6803 (1.3304)	4.9955 (1.225)	36.1543	24.2872 (4.11)	15.9527 (2.5076)	16.8183 (2.3558)
100	3.7503	2.1342 (0.8751)	1.8744 (0.7904)	2.0593 (0.672)	10.9184	6.3662 (1.722)	5.4425 (1.381)	6.0336 (1.1962)
200	1.7691	1.095 (0.432)	0.9656 (0.5565)	1.0667 (0.471)	4.9134	2.8924 (0.8691)	2.6309 (0.8932)	2.9933 (0.7412)
250	1.3694	0.9079 (0.3503)	0.7825 (0.4909)	0.8659 (0.404)	4.535	2.6809 (0.809)	2.4306 (0.8554)	2.7763 (0.7007)
300	1.1379	0.767 (0.2887)	0.6563 (0.4363)	0.7225 (0.3566)	3.6341	2.2224 (0.673)	2.0133 (0.7608)	2.308 (0.613)
					$\rho = 0.95$			
50	21.1369	14.5099 (3.1486)	8.9644 (1.9032)	9.329 (1.8201)	55.5389	39.4116 (5.4669)	24.734 (3.2339)	25.7803 (3.108)
100	8.6715	5.0046 (1.6912)	3.9972 (1.2453)	4.2759 (1.1442)	23.6196	14.721 (3.0749)	11.0199 (2.1017)	11.7709 (1.9308)
200	3.3298	1.8709 (0.769)	1.691 (0.7759)	1.8613 (0.6729)	10.5341	6.1114 (1.6697)	5.2996 (1.3905)	5.9086 (1.2074)
250	2.5832	1.486 (0.6244)	1.3239 (0.6588)	1.4791 (0.5482)	7.9676	4.6037 (1.3334)	4.0252 (1.1657)	4.5455 (0.979)
300	2.2202	1.3124 (0.5477)	1.1693 (0.6162)	1.3065 (0.5081)	7.968	4.5845 (1.3373)	4.0293 (1.1655)	4.559 (0.9796)

				$\rho = 0.99$				
50	105.3262	88.1138 (8.5189)	44.0244 (4.4437)	44.2843 (4.4115)	519.1215	471.8716 (20.0551)	215.4333 (9.9099)	216.0097 (9.8602)
100	43.8742	32.6166 (5.1036)	18.8912 (2.9275)	19.148 (2.861)	138.624	109.3103 (9.6006)	61.4149 (5.3723)	62.0112 (5.2883)
200	18.8052	12.2006 (2.9712)	8.4255 (1.9399)	8.7189 (1.8415)	50.1524	34.0863 (5.1123)	23.2528 (3.2658)	24.1018 (3.1336)
250	14.6668	9.2066 (2.5146)	6.6405 (1.6963)	6.9547 (1.5861)	51.4938	35.2846 (5.2094)	23.8129 (3.3153)	24.58 (3.18)
300	13.2734	8.0937 (2.3255)	5.9343 (1.5874)	6.2385 (1.4739)	39.164	25.9081 (4.3507)	18.0784 (2.85)	18.8447 (2.7011)

Note: The bias values are given in parentheses.

Table 4: MSE values and biases MNL estimators when $p = 8$.

n		$m = 3$				$m = 5$		
	MLE	Ridge	MLLE	AUMLLE	MLE	Ridge	MLLE	AUMLLE
				$\rho = 0.80$				
50	9.1483	6.137 (1.3973)	4.2453 (0.8445)	4.6194 (0.7475)	63.2999	46.6185 (5.8353)	27.7896 (3.0036)	29.8519 (2.7811)
100	5.0298	2.9617 (0.832)	2.7092 (0.8956)	3.0211 (0.7623)	17.1669	10.1919 (2.1072)	8.947 (1.6701)	10.015 (1.4095)
200	1.984	1.3951 (0.3202)	1.1938 (0.5628)	1.3693 (0.4171)	6.5603	4.0825 (0.8784)	3.7618 (0.9901)	4.3972 (0.7371)
250	1.5831	1.1648 (0.2593)	0.9783 (0.4912)	1.1312 (0.3506)	5.4459	3.465 (0.7209)	3.1579 (0.8947)	3.7047 (0.6576)
300	1.3451	1.0014 (0.2234)	0.8342 (0.436)	0.961 (0.3067)	3.9754	2.6763 (0.5219)	2.3771 (0.7461)	2.8088 (0.5294)

				$\rho = 0.85$				
50	26.2735	19.0593 (3.5049)	11.4888 (1.917)	12.3892 (1.7492)	83.7186	65.0172 (6.9608)	35.3503 (3.4256)	37.7222 (3.2076)
100	7.4764	4.4647 (1.2977)	3.884 (1.0851)	4.3878 (0.8869)	26.0738	16.4895 (3.0719)	12.8151 (2.0697)	14.0963 (1.7978)
200	2.858	1.8294 (0.5132)	1.6302 (0.6367)	1.8911 (0.4709)	9.4957	5.7647 (1.2818)	5.2944 (1.2228)	6.1528 (0.9472)
250	2.1532	1.4796 (0.3593)	1.2808 (0.5751)	1.486 (0.4173)	7.292	4.5063 (0.984)	4.1468 (1.0509)	4.8796 (0.7807)
300	1.7228	1.2397 (0.292)	1.054 (0.5056)	1.2259 (0.3605)	5.5689	3.5832 (0.7466)	3.2757 (0.9162)	3.8747 (0.6703)
				$\rho = 0.90$				
50	40.3766	30.8419 (4.5612)	17.0475 (0.4092)	18.2186 (2.2583)	121.9551	98.9983 (8.8485)	50.5742 (4.245)	53.3242 (4.0445)
100	9.86	5.7878 (1.599)	4.9693 (1.2726)	5.5765 (1.0684)	35.0418	22.5983 3.7857	17.2413 2.4975	18.8231 2.2504
200	4.2214	2.575 (0.7496)	2.324 (0.7987)	2.7025 (0.602)	13.9984	8.3062 (1.7855)	7.451 (1.5259)	8.5471 (1.2363)
250	3.3492	2.1119 (0.5709)	1.8936 (0.7288)	2.2079 (0.5444)	10.8504	6.5121 (1.4171)	5.9699 (1.3366)	6.9403 (1.0472)
300	2.6141	1.7378 (0.4566)	1.5242 (0.6258)	1.7931 (0.4514)	9.2554	5.7223 (1.2767)	5.2096 (1.2169)	6.1355 (0.923)
				$\rho = 0.95$				
50	70.5223	57.6306 (6.5785)	28.9351 (3.3197)	30.2354 (3.1691)	203.0149	171.8041 (12.022)	83.1832 (5.5876)	86.8274 (5.4176)
100	22.848	14.9695 (3.1496)	10.7119 (2.0166)	11.6263 (1.8114)	71.2762	50.828 (6.2632)	33.6347 (3.687)	35.7547 (3.4655)

200	8.8546	5.2802 (1.4886)	4.5716 (1.2249)	5.2368 (0.9929)	29.0078	18.0046 (3.2473)	14.6012 (2.3226)	16.1215 (2.0431)
250	6.4524	3.8314 (1.1163)	3.4063 (1.0192)	3.9572 (0.7973)	22.0353	13.6651 (2.6813)	11.3921 (2.0021)	12.7677 (1.702)
300	4.9719	3.0156 (0.8615)	2.7205 (0.9033)	3.1947 (0.6902)	17.1314	10.3937 (2.1749)	9.0599 (1.7325)	10.3189 (1.429)
					$\rho = 0.99$			
50	380.5692	354.4097 (17.2317)	152.0629 (7.8848)	153.8081 (7.836)	1690.285	1629.199 (37.4551)	640.0813 (16.1554)	644.3039 (16.097)
100	120.2018	99.2879 (9.2546)	53.3928 (4.9243)	54.3901 (4.8223)	359.4421	310.074 (16.7569)	162.1685 (8.6225)	164.8815 (8.5141)
200	49.7937	36.5371 (5.3244)	22.3056 (3.1197)	23.3379 (2.9376)	166.0478	131.7042 (10.626)	76.8979 (5.9377)	79.3824 (5.7536)
250	38.798	27.29 (4.5146)	17.6661 (2.751)	18.6869 (2.5538)	127.1571	97.7 (9.0419)	60.0541 (5.2481)	62.3953 (5.0443)
300	30.4368	20.4919 (3.8437)	14.3238 (2.464)	15.3325 (2.2531)	98.892	72.6789 (7.6907)	47.3181 (4.625)	49.62 (4.4092)

Note: The bias values are given in parentheses.

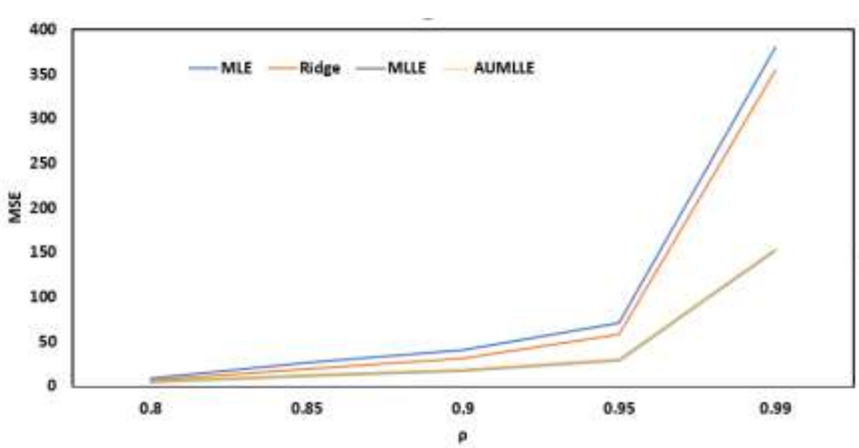

Figure 1. Effect of Multicollinearity on the Performance of Estimators for n=50 and $p = 8$

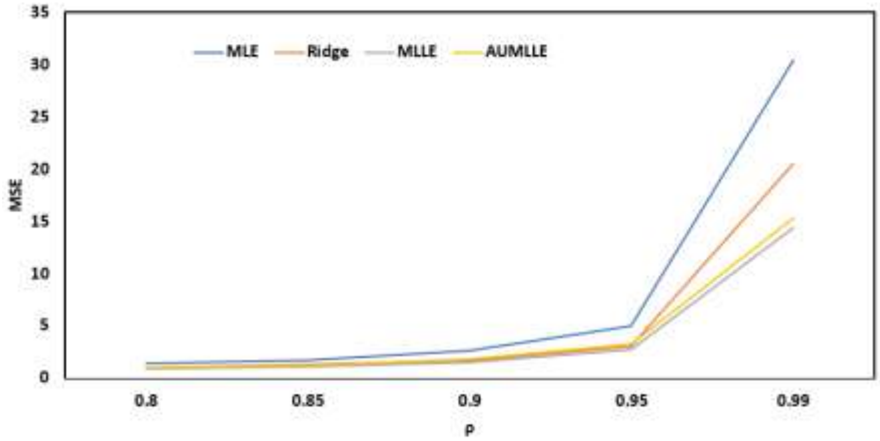

Figure 2. Effect of Multicollinearity on the Performance of Estimators for n=300 and $p = 8$

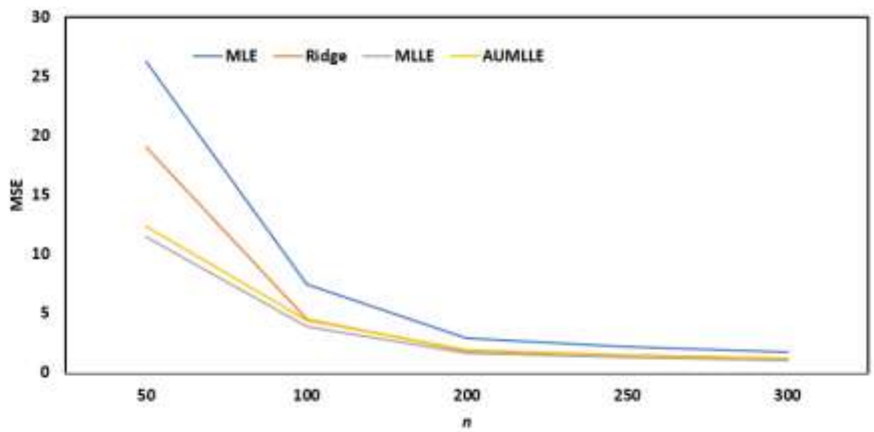

Figure 3. Effect of Sample Size on the Performance of Estimators for $\rho = 0.85$ and $p = 8$

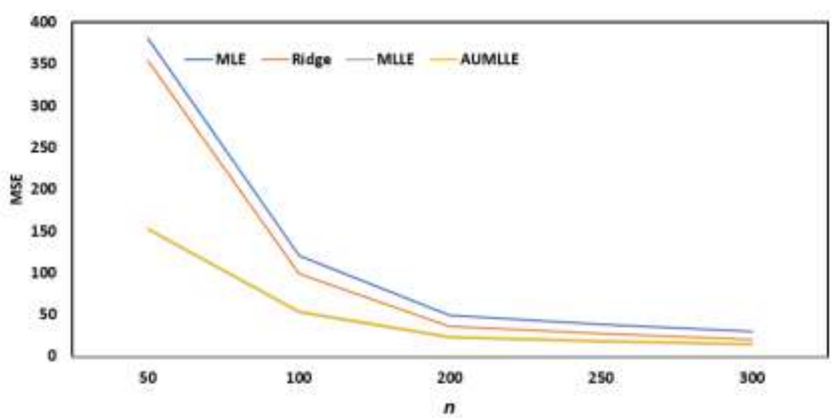

Figure 4. Effect of Sample Size on the Performance of Estimators for $\rho = 0.99$ and $p = 8$

5. Applications

In this section, we check the performance of the considered estimators with the help of three real examples.

5.1 Example 1: High School Data

A real dataset is employed to examine and compare the effectiveness of the proposed estimators against the MLE and ridge estimators. Specifically, the hsbdemo dataset is utilized, previously employed by Abonazel and Faraghali (2019). This dataset encompasses variables that impact the program selection (general, academic, vocational) of high school students. The data set has 11 variables, including the program of study (y), Gender (x_1), Social economic status (x_2) (Ses), Type of school (x_3) (Schtyp), Honors status (x_4), Reading test score (x_5) (Read), Writing test score (x_6) (Write), Math test score (x_7) (Math), Science test score (x_8) (Science), social studies score (x_9) (Socst), Number of awards (x_{10}) (Awards) in total for 200 students. Their program choice can be modeled by taking the study program as the response variable and other variables as predictors. Abonazel and Faraghali (2019) showed that the data have multicollinearity as the generalized variance inflation factor for seven variables is greater than 10 (see Table 9 of Abonazel and Faraghali (2019).

The coefficients of the MLE, ridge estimator, MLLE, AUMLLE and are calculated using Eqs (4), (5), (7) and (9) respectively. The values of MSEs of MLE, ridge estimator, MLLE and AUMLLE are computed using Eqs (4a), (6),(8) and (10) respectively. The estimated coefficients and MSE (in parentheses) are given in Table 5. Among all estimators, the AUMLLE exhibits the smallest MSE, signifying its superior performance. Conversely, the MLE displays a larger MSE in comparison to the other estimators.

Table 5: Estimates and MSEs of the MLE, MLLE and AUMLLE for the high school data.

Variable	Estimates			
	MLE (59.2331)	Ridge (31.1152)	MLLE (13.3425)	AUMLLE (13.1181)
Level: academic				
Intercept	-5.6912	-1.4371	-4.3582	-4.1581
x_1	-0.1547	-0.1669	-0.1581	-0.1531
x_{21}	0.2810	0.0803	0.2145	0.2771
x_{21}	0.9633	0.6201	0.8501	0.9429
x_3	0.5871	0.6224	0.5956	0.5756
x_4	0.0442	0.0376	0.0421	0.0442
x_5	0.0543	-0.0119	0.0335	0.0543
x_6	0.1001	0.0847	0.0953	0.1001
x_7	-0.1039	-0.0984	-0.1021	-0.1039
x_8	0.0248	0.0215	0.0238	0.0248
x_9	0.5965	0.1248	0.4469	0.5621
x_{10}	-0.2610	0.1387	-0.1353	-0.2608
Level: vocational				
Intercept	4.1460	1.2628	3.2338	3.0104
x_1	0.2529	0.1584	0.2221	0.2485
x_{21}	1.5057	0.9910	1.3351	1.4793
x_{21}	0.9643	0.4568	0.7957	0.9275
x_3	-1.3222	-0.8357	-1.1596	-1.2255
x_4	0.0029	0.0178	0.0077	0.0029
x_5	0.0028	0.0245	0.0095	0.0028
x_6	-0.0208	0.0075	-0.0118	-0.0208
x_7	-0.0.0405	-0.0393	-0.0401	-0.0405
x_8	-0.0452	-0.0392	-0.0432	-0.0452
x_9	1.6955	0.7107	1.3768	1.5217
x_{10}	-0.3469	-0.4212	-0.36806	-0.3463

5.2 Example 2: Cancer Data

The evaluation of estimator performance is conducted utilizing the cancer remission dataset previously employed in studies by Lukman et al. (2023), Özkale and Arıcan (2016), and Lesaffre and Marx (1993). The data set consists of one binary response y_i which takes the value 1 if the patient has complete cancer remission otherwise it takes the value 0. There were five explanatory variables including cell index (x_1), smear index (x_2), infil index (x_3), blast index (x_4) and temperature (x_5). There were in total 27 patients, and of those 9 have experienced complete remission. To evaluate the presence of multicollinearity among the explanatory variables, Lukman et al. (2023) utilized the condition index.

The analysis finds a moderate level of multicollinearity, with a CI value of 17.2. A CI value falling within the range of 10 to 30 indicates moderate collinearity, while exceeding 30 is indicative of severe multicollinearity, as elaborated in Gujrati (1993). The estimated coefficients are presented in Table 6. These coefficients for the MLE, ridge estimator, MLLE, and AUMLLE are calculated using Equations (4), (5), (7), and (9) respectively. The MSEs for the MLE, ridge estimator, MLLE, and AUMLLE are computed using Equations (4a), (6), (8), and (10) respectively. It can be seen that the MSE (in the parenthesis of Table 6) of AUMLLE and MLLE is small and MLE has the largest value among all the considered estimators which depicts that the proposed estimator performs better.

Table 6: Estimates and MSEs of the MLE, MLLE and AUMLLE for the cancer data

Variable	Estimators			
	MLE (6572.759)	Ridge (1124.387)	MLLE (965.0741)	AUMLLE (963.0328)
constant	18.3640	-0.5314	2.3803	5.0325
x_1	12.3812	-0.3494	1.6135	3.3948
x_2	-11.7317	-0.2082	-1.8671	-3.3183
x_3	3.66387	0.9277	0.9954	1.3776
x_4	-0.8812	0.4083	0.1199	-0.1583
x_5	-22.2248	-0.7393	-3.6215	-6.5076

5.3 Example 3: Football Data

For further assessing the performance of the proposed estimators, we examine data concerning the performance of Swedish football teams in the top Swedish league for the year 2018. This dataset was initially utilized by Qasim et al. (2020a). In total, there are 242 observations in this dataset. The dependent variable (Y) is full-time results (H: Home win, D: Draw and A: Away win) and there are 9 predictors including the pinnacle home win odds (PH), pinnacle draws odds (PD), pinnacle away win odds (PA), maximum Odds-portal draw win odds (MD), maximum Odds-portal home win odds (MH), maximum Odds-portal away win odds (MA), average Odds-portal home win odds (AvgH), average Odds-portal draw win odds (AvgD), and average Odds-portal away win odds (AvgA). The influence of these regressors on the response variable Y is assessed using the MNL model. It is worth noting that all the VIFs exceed the threshold of 10, indicating a multicollinearity problem. Furthermore, on many occasions, the correlation coefficients between the predictors are greater than 0.85, which also signals the presence of multicollinearity issues. Table 7 demonstrates the values of estimated coefficients and MSEs (in the parenthesis). The coefficients of the MLE, ridge estimator, MLLE, and AUMLLE are calculated using Equations (4), (5), (7) and (9) respectively. The values of MSEs of MLE, ridge estimator, MLLE and AUMLLE are computed using Equations (4a), (6), (8) and (10) respectively. The efficient performance of the proposed estimators is evidenced as the results presented through analysis depict the smaller MSE of proposed estimator as compared to ridge and MLE.

Table 7: Estimates and MSEs of the MLE, MLLE and AUMLLE for the football data

Variable	Estimates			
	MLE (110.9167)	Ridge (93.4301)	MLLE (33.2434)	AUMLLE (31.0915)
Level: D				
Intercept	0.0391	0.1931	0.0988	0.0248
PH	-0.3010	-0.0259	-0.1555	-0.2285
PD	1.5559	0.0299	0.7448	1.1090
PA	-0.7685	-0.0909	-0.4039	-0.6305
MH	1.4817	0.1423	0.7762	1.0960
MD	-7.0140	-0.7007	-3.6972	-4.6615
MA	0.3870	0.2167	0.3056	0.3324
AvgH	-1.1603	-0.0322	-0.5721	-0.8017
AvgD	4.9816	0.0659	2.4258	3.1529
AvgA	1.0673	0.2824	0.6291	0.7815
Level: H				
Intercept	0.7295	0.0341	0.3632	0.4662
PH	0.3027	0.0888	0.1766	0.2159
PD	2.7864	0.4740	1.5414	2.0149
PA	-0.6463	0.0666	-0.2613	-0.5406
MH	1.5831	0.1033	0.7994	1.1240
MD	-2.9121	-0.3741	-1.5720	-1.9748
MA	-0.4690	-0.3822	-0.3999	-0.4092
AvgH	-2.2502	-0.4480	-1.2894	-1.4957
AvgD	-0.5897	-0.4287	-0.4641	-0.3760
AvgA	2.0409	0.8778	1.3807	1.5292

Various empirical applications have been utilized to emphasize the effectiveness of the proposed estimators (MLLE and AUMLLE). The classical MLE has not performed sufficiently good when there is multicollinearity in data. To address this very concern, Månsson et al. (2018) presented the multinomial ridge regression (ridge) estimator. However, both empirical and simulation implementations disclosed that the ridge estimator exhibited smaller MSE when compared with MLE, it is subject to bias due to its non-linear relationship with the shrinkage parameter. In contrast, both MLLE and AUMLLE are characterized by a linear relationship with the shrinkage parameter. Consequently, in simulation and empirical findings, the proposed estimators, namely MLLE and AUMLLE, both consistently outperformed the ridge (Månsson et al., 2018) and MLE in the sense that they have smaller MSE. Hence, MLLE and AUMLLE have performed better than the existing estimators.

6 Conclusion

Liu and bias-corrected Liu estimator are derived to address the multicollinearity problem for the MNL model. The MSE and Bias of the estimators are acquired along with the optimal values of the biasing parameters. Also, we examined the superiority of proposed estimators to the MLE. The optimal values of the biasing parameter d are obtained for the MLLE and AUMLLE. A Monte Carlo simulation study is conducted to illustrate the performance of the MLLE, and AUMLLE against the MLE and ridge estimator by varying factors such as the number of regressors, sample size, levels of the response variable and multicollinearity level. MSE and bias are used as performance criteria to evaluate the performance of the proposed estimator. Based on simulation results, we concluded that the increase in the level of response, the correlation between predictors and the number of regressors harms the performance of estimators. However, it is worth noting that the number of observations exerts a positive influence on estimator performance, even in scenarios characterized by high levels of multicollinearity and an increased number of regressors. Thus, MLLE and AUMLLE are superior to the MLE and ridge estimators in nearly all scenarios. The application of real data examples demonstrated that the proposed estimators outperformed both the MLE and ridge estimators.

Acknowledgments

We are thankful to Dr. Mohamed R. Abonazel for his support in developing software code for the simulation study. Moreover, we are thankful to the reviewers and editor for their constructive comments to improve the quality of the study.

References:

1. Abonazel, M.R., and R.A. Farghali. 2019. Liu-type multinomial logistic estimator. Sankhya B, 81(2):203-225.
2. Aguilera, A., and M. Escabias. 2008. Solving multicollinearity in functional multinomial logit models for nominal and ordinal responses. In Functional and operatorial statistics : 7-13. Physica-Verlag HD.
3. Algamal, Z. Y., Abonazel, M. R., Awwad, F. A., & Eldin, E. T. (2023). Modified jackknife ridge estimator for the Conway-Maxwell-Poisson model. Scientific African, 19, e01543.
4. Alheety, M. I., and B. M. G. Kibria. 2009. On the Liu and almost unbiased Liu estimators in the presence of multicollinearity with heteroscedastic or correlated errors. Surveys in Mathematics and Applications 4:155–67.
5. Alheety, M. I., Qasim, M., Månsson, K., and Kibria, B.M.G. (2021). Modified almost unbiased two-parameter estimator for the Poisson regression model with an application to accident data. SORT-Statistics and Operations Research Transactions, 45(2), 121-142.
6. Amin, M., Qasim, M., Yasin, A., and Amanullah, M. (2022). Almost unbiased ridge estimator in the gamma regression model. Communications in Statistics-Simulation and Computation, 51(7), 3830-3850.
7. Arashi, M., Lukman, A. F. and Algamal, Z. Y. (2022) Liu regression after random forest for prediction and modeling in high-dimension, Journal of Chemometrics, 36(4), e3393.
8. Arumairajan, S., and P. Wijekoon. 2017. Modified almost unbiased Liu estimator in linear regression model. Communications in Mathematics and Statistics, 5(3): 261-276.doi: https://doi.org/10.1007/s40304-017-0111-z
9. Asar, Y. 2016. Liu-type logistic estimators with optimal shrinkage parameter. Journal of Modern Applied Statistical Methods, 15(1):36.
10. Asar, Y., and A. Genç. 2016. New shrinkage parameters for the Liu-type logistic estimators. Communications in Statistics-Simulation and Computation, 45(3):1094-1103.
11. Babar, I., and Chand, S. (2022). Weighted ridge and Liu estimators for linear regression model. Concurrency and Computation: Practice and Experience, 34(27), e7343.
12. Camminatiello, I., and A. Lucadamo. 2008. Estimating multinomial logit model with multicollinear data. MTISD 2008. Methods, models and information technologies for decision support systems, 1(1):51-54.
13. Chang, X.F. (2015). On the almost unbiased ridge and Liu estimator in the logistic regression model. In International Conference on Social Science, Education Management and Sports Education, Shanghai :1-3.
14. El-Dash, A., A. El-Hefnawy, and R.A. Farghali. (2011). Goal programming technique for correcting multicollinearity problem in multinomial logistic regression. In The 46th Annual Conference on Statistics, Computer Sciences, and Operation Research: 26-29.

15. Farghali, R. A., Qasim, M., Kibria, B.M.G., and Abonazel, M. R. (2023). Generalized two-parameter estimators in the multinomial logit regression model: methods, simulation and application. Communications in Statistics-Simulation and Computation, 52(7), 3327-3342.
16. Frisch, R. 1934. Statistical Confluence Analysis by Means of Complete Regression Systems. Oslo: University Institute of Economics.
17. Gujarati, D.N. 1995. Basic Econometrics; McGraw-Hill: New York, NY, USA.
18. Hadia, M., Amin, M., and Akram, M.N. (2022). Comparison of link functions for the estimation of logistic ridge regression: an application to urine data. Communications in Statistics - Simulation and Computation, https://doi.org/10.1080/03610918.2022.2127769.
19. Hoerl, A. E., and R. W. Kennard. 1970. Ridge regression: Biased estimation for nonorthogonal problems. Technometrics, 12 (1):55–67.
20. Inan, D., and B.E. Erdogan. 2013. Liu-type logistic estimator. Communications in Statistics-Simulation and Computation, 42(7): 1578-1586.
21. Kibria, B. M. G. 2003. Performance of some new ridge regression estimators. Communications in Statistics—Theory and Methods 32(2):419–435.
22. Lesaffre, E. and Marx, B.D. (1993). Collinearity in generalized linear regression. Communications in Statistics-Theory and Methods, 22(7), 1933-1952.
23. Liu, K. 1993. A new class of biased estimate in linear regression. Communication in Statistics - Theory Methods 22 (2):393–402.
24. Liu, K. 2003. Using Liu-type estimator to combat collinearity. Communications in Statistics-Theory and Methods, 32(5):1009-1020.
25. Lukman, A.F., Kibria, B.M.G., Nziku, C.K., Amin, M., Adewuyi, E.T. and Farghali, R. (2023). KL estimator: Dealing with multicollinearity in the logistic regression Model. Mathematics, 11(2), 340.
26. Lukman, A. F., Arashi, M. and Prokaj, V. (2023). Robust biased estimators for Poisson regression model: simulation and applications, Concurrency and Computation: Practice and Experience, e7594.
27. Månsson, K., and G. Shukur. 2011. On ridge parameters in logistic regression. Communications in Statistics-Theory and Methods, 40(18): 3366-3381.
28. Månsson, K., B. M. G. Kibria, and G. Shukur. 2012. On Liu estimators for the logit regression model. Economic Modelling 29 (4):1483–88.
29. Månsson, K., G. Shukur, and B. M. G. Kibria. 2018. Performance of some ridge regression estimators for the multinomial logit model. Communications in Statistics-Theory and Methods, 47(12): 2795-2804.
30. Mustafa, S., Amin, M., Akram, M.N., Afzal, N. (2022). On the performance of some link functions in the beta ridge regression model: Simulation and Application. Concurrency and Computation: Practice and Experience, 34(18), e7005.
31. Ogoke, U.P., E.C. Nduka, and M.E. Nja. 2013. The logistic regression model with a modified weight function in survival analysis. Mathematical Theory & Modeling, 3(8):12-17.
32. Özkale, M.R. and Arıcan, E. 2016. A new biased estimator in logistic regression model. Statistics, 50(2), 233-253.
33. Qasim, M., Amin, M., & Amanullah, M. (2018). On the performance of some new Liu parameters for the gamma regression model. Journal of Statistical Computation and Simulation, 88(16), 3065-3080.
34. Qasim, M., Kibria, B. M. G., Månsson, K., & Sjölander, P. (2020). A new Poisson Liu regression estimator: method and application. Journal of Applied Statistics, 47(12), 2258-2271.
35. Qasim, M., Månsson, K., & Kibria, B.M.G. (2021). On some beta ridge regression estimators: method, simulation and application. Journal of Statistical Computation and Simulation, 91(9), 1699-1712.
36. Qasim, M., Månsson, K., Amin, M., Kibria, B.M.G., & Sjölander, P. (2020). Biased adjusted Poisson ridge estimators-method and application. Iranian Journal of Science and Technology, Transactions A: Science, 44, 1775-1789.
37. Qasim, M., Månsson, K., Sjölander, P., & Kibria, B.M.G. (2022). A new class of efficient and debiased two-step shrinkage estimators: method and application. Journal of applied Statistics, 49(16), 4181-4205.
38. Saleh, A. K. Md. Ehsanes, Arashi, M., Saleh, R. A. and Norouzirad, M. (2022). Rank-Based Methods for Shrinkage and Selection: With Application to Machine Learning, Wiley, USA.

39. Saleh, A. K. Md. Ehsanes, Arashi, M., and Kibria, B. M. G. (2019). Theory of Ridge Regression Estimation with Applications, Wiley, USA.
40. Sami, F., Amin, M., Butt, M.M., and Yasin, S. (2023). An almost unbiased ridge estimator for the Conway–Maxwell–Poisson regression model. Iranian Journal of Science, 47, 1209-1219.
41. Schaefer, R. L., L. D. Roi, and R. A. Wolfe. 1984. A ridge logistic estimator. Communications in Statistics - Theory and Methods, 13 (1): 99–113.
42. Schaefer, R.L. 1986. Alternative estimators in logistic regression when the data are collinear. Journal of Statistical Computation and Simulation, 25(1-2): 75-91.
43. So, Y., and W. F. Kuhfeld. 1995. Multinomial logit models. In SUGI 20 conference proceedings : 1227-1234.
44. Suhail, M., Chand, S., & Kibria, B. M. G. (2020). Quantile based estimation of biasing parameters in ridge regression model. Communications in Statistics-Simulation and Computation, 49(10), 2732–2744.
45. Taavoni, M. Arashi, M. and Manda, S. (2023) Multicollinearity and linear predictor link function problems in regression modelling of longitudinal data. Mathematics, 11, 530.
46. Varathan, N. and P. Wijekoon. 2016. Ridge estimator in logistic regression under stochastic linear restrictions. British Journal of Mathematics & Computer Science, 15(3):1.
47. Varathan, N., and P. Wijekoon. 2018. Liu-type logistic estimator under stochastic linear restrictions. Ceylon Journal of Science 47 (1):21–34.
48. Varathan, N., and P. Wijekoon. 2018. Optimal generalized logistic estimator. Communications in Statistics-Theory and Methods, 47(2): 463-474.
49. Varathan, N., and P. Wijekoon. 2021. Modified almost unbiased Liu estimator in logistic regression. Communications in Statistics-Simulation and Computation, 50(11): 3530-3546.
50. Wasim, D., Khan, S. A., & Suhail, M. (2023). Modified robust ridge M-estimators for linear regression models: an application to tobacco data. Journal of Statistical Computation and Simulation, 1–22.
51. Wu, J. B., Asar, Y. (2016). On almost unbiased ridge logistic estimator for the logistic regression model. Hacettepe Journal of Mathematics and Statistics 45(3):989–998.
52. Xu, J., and Yang, H. 2011. More on the bias and variance comparisons of the restricted almost unbiased estimators. Communications in Statistics-Theory and Methods 40(22): 4053-64.
53. Yasin, A.S.A.R., A. Karaibrahimoğlu, and G.E.N.Ç. Aşır. 2013. Modified ridge regression parameters: A comparative Monte Carlo study. Hacettepe Journal of Mathematics and Statistics, 43(5): 827-841.
54. Zahid, F.M., and G. Tutz. 2013. Multinomial logit models with implicit variable selection. Advances in Data Analysis and Classification, 7(4): 393-416.