On a Mixed Poisson Liu Regression Estimator for Overdispersed and Multicollinear Count Data

Ramajeyam Tharshan and Pushpakanthie Wijekoon

1Postgraduate Institute of Science, University of Peradeniya, Peradeniya, Sri Lanka
2Department of Mathematics and Statistics, University of Jaffna, Jaffna, Sri Lanka
3Department of Statistics and Computer Science, University of Peradeniya, Peradeniya, Sri Lanka

Correspondence should be addressed to Ramajeyam Tharshan; tharshan10684@gmail.com

Received 10 March 2022; Revised 13 June 2022; Accepted 24 June 2022; Published 20 July 2022

Copyright © 2022 Ramajeyam Tharshan and Pushpakanthie Wijekoon. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The mixed Poisson regression models are commonly employed to analyze the overdispersed count data. However, multicollinearity is a common issue when estimating the regression coefficients by using the maximum likelihood estimator (MLE) in such regression models. To deal with the multicollinearity, a Liu estimator was proposed by Liu (1993). The Poisson-Modification of the Quasi Lindley (PMQL) regression model is a mixed Poisson regression model introduced recently. The primary interest of this paper is to introduce the Liu estimator for the PMQL regression model to mitigate the multicollinearity issue. To estimate the Liu parameter, some exiting methods are used, and the superiority conditions of the new estimator over the MLE and PMQL ridge regression estimator are obtained based on the mean square error (MSE) criterion. A Monte Carlo simulation study and applications are used to assess the performance of the new estimator in the scalar mean square error (SMSE) sense. Based on the simulation study and the results of the applications, it is shown that the PMQL Liu estimator performs better than the MLE and some other existing biased estimators in the presence of multicollinearity.

1. Introduction

The Poisson regression model is a commonly used statistical method for analyzing the count response variable [1]. One disadvantage of this model is that it is an overdispersion issue that is common in the real-world applications of actuarial, engineering, biomedical, and economic sciences. The overdispersion occurs when the conditional variance of the count response variable exceeds the conditional mean of the count response variable. In this context, the index of dispersion (variance-to-mean ratio) is greater than one. To tackle this issue in the Poisson regression model, researchers have proposed several mixed Poisson regression models. The standard mixed Poisson distribution is obviously the negative binomial (NB)/Poisson-gamma regression model introduced by Greenwood and Yule [2]. However, the NB distribution fails to fit well for a count data with a higher value of the index of dispersion and long right-tail behavior. Then, the regression model based on NB is not a good choice for such a count response variable. As an alternative to the NB regression model, several mixed Poisson regression models are in the literature. However, most of the probability mass functions (pmfs) of these mixed Poisson distributions are not in an explicit form. Some notable examples of such regression models are the Poisson-Inverse Gaussian regression model [3] and the Poisson-Inverse gamma regression model [4]. This algebraic intractability in such distributions leads to computational complexity, and their regression models are limited in practice.

In the last decade, several researchers have highlighted mixed Poisson distributions obtained by mixing the Poisson and Lindley family of distributions due to their explicit form of the pmf and work efficiency. The Lindley family of distributions are two-component mixtures. Some notable such
mixed Poisson distributions are Poisson–Lindley distribution, Generalized Poisson–Lindley distribution, Poisson–generalized Lindley distribution, Poisson-Quasi Lindley distribution, and Poisson-weighted Lindley distribution, proposed by Sankaran [5], Mahmoudi and Zakerzadeh [6], Wongrin and Bodhisuwon [7], Grine and Zeghdoudi [8], and Atikankul et al. [9], respectively. They may have the flexibility to capture various ranges of horizontal symmetries, right-tail behaviors, and variance-to-mean ratios based on their mixing distributions [10–12]. However, the literature on their regression models is rather limited. Some of the most relevant works cited are the Generalized Poisson distributions with the mixing proportion, \( \alpha \) and \( \beta \) as scale parameters, \( \theta \) a shape parameter, and \( y \) the respective random variable bounded to \((0, \infty)\). Equation (1) presents the mixture of exponential (\( \theta \)) and gamma (\( \delta, \theta \)) distributions with the mixing proportion, \( p = (\alpha^3/\alpha^3 + 1) \). By mixing the Poisson and the MQL, Tharshan and Wijekoon [16] derived the Poisson-Modification of Quasi Lindley (MQL) distribution. Its explicit form of the pmf, and some other important statistics are given in Section 2. Authors have shown that the PMQL distribution is an overdispersed distribution, and it has the flexibility to capture the various ranges of horizontal symmetry, right-tail heaviness, and variance-to-mean ratio. Then, by using a reparameterization technique, the same authors [17] derived its regression model to predict the overdispersed count responses with a set of linear independent covariates based on the generalized linear model (GLM) approach. Further, in their paper, it is shown that the PMQL regression model performs better than the NB, GPL, and PQL regression models. More details of this regression model are given in Section 2. The traditional estimator to estimate the unknown regression coefficients of the PMQL regression model is the maximum likelihood estimator (MLE\(_{\text{PMQL}}\)), where the solutions of the nonlinear equations with respect to the regression coefficients are found by applying an iterative weighted least square (IWLS) algorithm. However, the MLE\(_{\text{PMQL}}\) is unstable, and its variance is inflated when the covariates are linearly correlated since it is a GLM. It leads to difficulty in having a valid statistical inference. This problem is commonly known as multicollinearity by Frisch [18]. To overcome the multicollinearity problem in the PMQL regression model, Tharshan and Wijekoon [19] adopted the ridge regression estimator in the PMQL regression model \( \text{RE}_{\text{PMQL}} \). The authors have shown that the \( \text{RE}_{\text{PMQL}} \) performs better than the MLE\(_{\text{PMQL}}\) when multicollinearity exists. Further, they recommended some ridge parameter estimation methods for the \( \text{RE}_{\text{PMQL}} \). The ridge regression estimator was suggested by Hoerl and Kennard [20] for the ordinary linear regression model, and it was extended to GLM by Segerstedt [21]. The ordinary linear regression model is defined as

\[
y = X\beta + \epsilon,
\]

where \( y \) is an \( n \times 1 \) vector of observations on a response variable \( Y \), \( \beta \) is a \( (p + 1) \times 1 \) vector of unknown regression coefficients, \( X \) is the known design matrix of order \( n \times (p + 1) \) with \( p \)-dimensional covariates, and \( \epsilon \) is an \( n \times 1 \) vector of errors with \( E(\epsilon) = 0 \) and \( \text{Var}(\epsilon) = \sigma^2 I_n \). Further, its unknown regression coefficients are estimated by the ordinary least square estimator, which is defined as

\[
\hat{\beta}_{\text{OLSE}} = (X'X)^{-1}X'y.
\]
estimators for the zero-inflated Poisson regression model. We may note that the Liu estimator for the regression model of a mixed Poisson distribution is rather limited in the literature.

This paper adopts the Liu estimator in the PMQL regression model to combat the multicollinearity. Further, we adhere to some possible estimation methods to estimate the Liu parameter \( d \) for the PMQL Liu regression estimator \((\text{LE}_{\text{PMQL}})\) based on the works carried out by Hoerl and Kennard [20], Kibria [35], and Khalaf and Shukur [36]. Then, the performance of the MLE\(_{\text{PMQL}}\), RE\(_{\text{PMQL}}\), and LE\(_{\text{PMQL}}\) will be compared in terms of the scalar mean square error (SMSE) criterion by using an extensive Monte Carlo simulation study. Finally, a simulated data set and a real-world example will be considered to illustrate the benefits of the Liu estimator for the PMQL regression model in handling the overdispersion and multicollinearity issues.

The rest of the paper is organized as follows: Section 2 discusses the PMQL regression model and its regression coefficients estimator. We present the LE\(_{\text{PMQL}}\), mean square error (MSE) properties of the LE\(_{\text{PMQL}}\), conditions that the LE\(_{\text{PMQL}}\) is superior to the MLE\(_{\text{PMQL}}\) and the RE\(_{\text{PMQL}}\), and possible Liu parameter estimators for the LE\(_{\text{PMQL}}\) in Section 3. Section 4 designs a Monte Carlo simulation study and discusses the results of the simulation study. Section 5 gives a simulated data set and real data applications in order to illustrate the applicability of the PMQL Liu regression model. Finally, the conclusion of the paper is given in Section 6.

2. PMQL Regression Model

In this section, we present the PMQL regression model and its regression coefficients estimation.

The PMQL distribution [16] is a resultant distribution or unconditional distribution by assuming that the Poisson parameter follows the MQL distribution. The pdf of the MQL distribution is given in equation (1). The probability mass function of the PMQL distribution is given as

\[
f(y) = \frac{\theta^y}{y!(\lambda^3 + 1)(1 + \theta)^{\lambda^3} \Gamma(y + 1)} \left( \Gamma(\delta) \Gamma(y + 1) \lambda^3 (1 + \theta)^{\delta - 1} + \theta^{\delta - 1} \Gamma(y + \delta) \right), \quad y = 0, 1, 2, \ldots, \theta > 0, \delta > 0, \lambda^3 > -1, \quad (5)
\]

where \( y \) is the respective random variable and represents the total counts of an experiment. Its mean and variance are given

\[
E(Y) = \frac{\lambda^3 + \delta}{(\lambda^3 + 1)} \theta = \mu, \quad (6)
\]

\[
\text{Var}(Y) = \mu + \mu^2 \left( \frac{\lambda^3 (\lambda^3 + 2 + \delta (\delta - 1)) + \delta}{(\lambda^3 + \delta)^2} \right), \quad (7)
\]

respectively. Equation (5) represents a two-component mixture of geometric \((\theta/1 + \theta)\) and negative binomial \((\delta, 1/1 + \theta)\) with the mixing proportion, \( \rho = (\lambda^3)/\alpha^3 + 1 \). Further, it possesses to be unimodal and bimodal distributions and overdispersed. The authors have shown that it has the potentiality to accommodate various horizontal symmetry, right-tail behaviors, and index of dispersion for overdispersed count data.

Let \( y_1, y_2, \ldots, y_n \) be the random sample of \( n \) observations from the PMQL distribution. The link between \( p \)-dimensional covariates and the mean responses \( y \) was taken as

\[
\eta_i = g(\mu_i) = \log(\mu_i) = \sum_{j=0}^{p} \beta_j x_{ij} = \mathbf{x}_i^{\prime} \mathbf{\beta}, \quad i = 1, 2, \ldots, n, \quad (8)
\]

where \( \mathbf{x}_i = (1, x_{i1}, x_{i2}, \ldots, x_{ip}) \) is the vector of \( i^{th} \) row of the known design matrix \( X \) supplemented with a 1 in front for the intercept, \( \mathbf{\beta} = (\beta_0, \beta_1, \ldots, \beta_p) \) is a vector of unknown regression coefficients of order \((p + 1) \times 1\) with intercept, and \( \eta \) and \( \delta \) are overdispersion parameters. To approach the GLM, the PMQL distribution was reparametrized based on the relationship between \( \mu \) and \( \theta \) given in equation (6) for a given set of \( \delta \) and values and the link between \( \mu \) and \( p \)-dimensional covariates given in equation (8).

That is, by substituting \( \theta_i = (\lambda^3 + \delta)/(\lambda^3 + 1) \exp(\mathbf{x}_i^{\prime} \mathbf{\beta}) \), \( i = 1, 2, \ldots, n \) in equation (5), the pmf of the \( y_i \) for a given set of covariates \( x_i \) was obtained as

\[
f(y_i|x_i) = \frac{((\lambda^3 + 1) \exp(\mathbf{x}_i^{\prime} \mathbf{\beta}))^{y_i} (\lambda^3 + \delta) \Gamma(\delta) \Gamma(y_i + 1) \lambda^3 A_i^{\delta - 1} + (\lambda^3 + \delta)^{\delta - 1} \Gamma(y_i + \delta))}{y_i!(\lambda^3 + 1)A_i^{y_i + \delta} \Gamma(\delta)}, \quad (9)
\]

where \( A_i = ((\lambda^3 + 1) \exp(\mathbf{x}_i^{\prime} \mathbf{\beta}) + (\lambda^3 + \delta)), i = 1, 2, \ldots, n \). The conditional mean and variance of the regression model are given:
\[ E(Y|x) = \exp(x\beta), \]
\[ \text{Var}(Y|x) = \exp(x\beta)(\exp(x\beta))^2 \times \left( \frac{\alpha^3(\alpha^3 + 2 + \delta(\delta - 1)) + \delta}{(\alpha^3 + \delta)^2} \right), \]  
(10) respectively. Figure 1 depicts the surface plots of the variance function of the PMQL regression model at \( \mu = 2.5 \) for different values of \( \alpha \) and \( \delta \). From Figure 1, it can be observed that the variance as a function of \( \alpha \) or \( \delta \) is not a monotonic function, and it is high for small values of \( \alpha \) and \( \delta \).

The estimation of the unknown regression coefficients is commonly estimated by maximizing the following log-likelihood function of its pmf given in equation (9):

\[
\ell(\beta, \alpha, \delta | y, x) = \sum_{i=1}^{n} y_i \log(\alpha^3 + 1) \exp(x_i \beta) + n \log(\alpha^3 + \delta) - \sum_{i=1}^{n} \log(y_i) - n \log(\alpha^3 + 1) - n \log(\Gamma(\delta))
\]  
(11)

The score function of the vector of regression coefficients \( \beta \) is given as

\[
S(\beta) = \frac{\partial \ell(\beta, \alpha, \delta | y, x)}{\partial \beta} = \sum_{i=1}^{n} y_i \exp(x_i \beta) x_i - \sum_{i=1}^{n} (y_i + \delta)(\alpha^3 + 1) \exp(x_i \beta) x_i / A_i
\]  
\[ + \sum_{i=1}^{n} \Gamma(\delta) \Gamma(y_i + 1) \alpha^3 A_i^{d-1} (\alpha^3 + 1) \exp(x_i \beta) x_i / \Gamma(\delta) \Gamma(y_i + 1) A_i^{d-1} (\alpha^3 + 1) \Gamma(y_i + \delta) \]  
(12)

where \( I(\beta^{(s-1)}) \) is a \((p + 1) \times (p + 1)\) Fisher information matrix and the \( S(\beta^{(s-1)}) \) is the score function of the regression coefficients calculated at \( \beta^{(s-1)} \). In the final step of the IWLS algorithm, \( \hat{\beta}_{MLE,PMQL}^{(s)} \) is obtained as

\[
\hat{\beta}_{MLE,PMQL}^{(s)} = (X'\hat{W}X)^{-1} X'\hat{W}z,
\]  
(14)

where \( \hat{W} = \text{diag}\left(\frac{1}{g'(\mu_i)}\right) = \text{diag}\left(\frac{\mu_i(\alpha^3 + \delta)^2}{(\alpha^3 + \delta)^2 + \mu_i(\alpha^3 + 2 + \delta(\delta - 1) + \delta)}\right) \)

(15)

\( \hat{z} \) is a vector, and its \( i \)th element is given as \( g'(\mu_i) + (y_i - \bar{\mu}_i)g'(\mu_i) = \log(\bar{\mu}_i) + (y_i - \bar{\mu}_i)/\bar{\mu}_i \).

The asymptotic covariance matrix of this estimator is given as

\[
\text{Cov}(\hat{\beta}_{MLE,PMQL}) = (X'\hat{W}X)^{-1}, \]
(16)

and the asymptotic MSE and SMSE of this estimator are given as [16]

\[
\text{MSE}(\hat{\beta}_{MLE,PMQL}) = (X'\hat{W}X)^{-1}, \]
(17)
\[
\text{SMSE}(\hat{\beta}_{MLE,PMQL}) = \sum_{j=1}^{p} \frac{1}{\lambda_j}, \]
(18)

respectively, where \( \lambda_j \) is the \( j \)th eigenvalue of the matrix \( X'\hat{W}X \).
When the covariates are highly correlated, the weighted matrix of cross-product $X'WX$ is ill-conditioned, and this matrix will have some smaller eigenvalues. We can observe that the SMSE ($\hat{\beta}_{MLE/(PMQL)}$) given in equation (18) can easily be inflated for smaller eigenvalues. In this situation, it is very hard to have a valid inference of whether the estimated regression coefficients are significant or not.

### 3. The PMQL Liu Regression Estimator

Note that the PMQL regression model is a GLM. Then, following the Liu estimator for the GLMs, which was proposed by Kurtoloğlu and Özkalet [28] for the GLMs based on the IWLS algorithm, we define the Liu estimator for the PMQL regression model to mitigate the multicollinearity issue as

$$
\tilde{\beta}^{(s)}_{LE/(PMQL)} = \left(X'\hat{W}(\beta^{(s-1)} + 1)\right)^{-1} \left(X'\hat{W}(\beta^{(s-1)} + 1)\right)^{-1} \tilde{\beta}_{MLE/(PMQL)}^{(s)}
$$

where $\tilde{\beta}^{(s)}_{LE/(PMQL)}$ is the estimated value of the $\beta$ by the Liu estimator with the $s$ iterations, $\hat{W}(\beta^{(s-1)})$ is the weighted matrix evaluated at $\beta^{(s-1)}$, and $\tilde{\beta}_{MLE/(PMQL)}^{(s)}$ is the estimated value of the $\beta$ by the maximum likelihood method with the $s$ iterations. In the final step of the IWLS algorithm, $\tilde{\beta}_{LE/(PMQL)}^{(s)}$ can be obtained as

$$
\tilde{\beta}_{LE/(PMQL)}^{(s)} = \left(X'\hat{W}X + dI\right)^{-1}\left(X'\hat{W}X + dI\right)^{-1} \tilde{\beta}_{MLE/(PMQL)}^{(s)}
$$

where $d(0 < d < 1)$ is the Liu parameter, $I$ is a $(p + 1) \times (p + 1)$ identity matrix, and $L_d = (X'\hat{W}X + dI)$. Note that, if $d = 1$, then $\tilde{\beta}_{MLE/(PMQL)} = \tilde{\beta}_{LE/(PMQL)}$.

Asymptotic properties of the PMQL Liu regression estimator are as follows:

$$
E(\tilde{\beta}_{LE/(PMQL)}) = L_d\beta,
$$

$$
\text{Cov}(\tilde{\beta}_{LE/(PMQL)}) = L_d\text{Cov}(\tilde{\beta}_{MLE/(PMQL)})L_d',
$$

$$
= \left(X'\hat{W}X + 1\right)^{-1}(X'\hat{W}X + dI)(X'\hat{W}X)^{-1}(X'\hat{W}X + dI)(X'\hat{W}X + 1)^{-1}
$$

$$
= \left(X'\hat{W}X + 1\right)^{-1}(X'\hat{W}X + dI)(I + d(X'\hat{W}X)^{-1})(X'\hat{W}X + 1)^{-1}
$$

$$
= L_d\left(I + d(X'\hat{W}X)^{-1}\right)(X'\hat{W}X + 1)^{-1},
$$

and then the asymptotic bias and MSE are given as

$$
\text{Bias}(\tilde{\beta}_{LE/(PMQL)}) = E(\tilde{\beta}_{LE/(PMQL)}) - \beta
$$

$$
= (L_d - I)\beta
$$

$$
= \left((X'\hat{W}X + 1)^{-1}(X'\hat{W}X + dI) - I\right)\beta
$$

$$
= (d - 1)(X'\hat{W}X + 1)^{-1}\beta,
$$

$$
\text{MSE}(\tilde{\beta}_{LE/(PMQL)}) = E\left((\tilde{\beta}_{LE/(PMQL)} - \beta)(\tilde{\beta}_{LE/(PMQL)} - \beta)'ight)
$$

$$
= \text{Cov}(\tilde{\beta}_{LE/(PMQL)}) + \text{Bias}(\tilde{\beta}_{LE/(PMQL)})^2
$$

$$
= L_d\left(I + d(X'\hat{W}X)^{-1}\right)(X'\hat{W}X + 1)^{-1} + (d - 1)^2(X'\hat{W}X + 1)^{-1}\beta\beta'\left(X'\hat{W}X + 1\right)^{-1},
$$

where $\beta$ is the true parameter.
respectively. Now, we derive the asymptotic SMSE of the estimator as

\[
\text{SMSE} \left( \hat{\beta}_{\text{LE}(\text{PMQL})} \right) = \text{trace} \left( \text{MSE} \left( \hat{\beta}_{\text{LE}(\text{PMQL})} \right) \right) \\
= \text{trace} \left( \text{Cov} \left( \hat{\beta}_{\text{LE}(\text{PMQL})} \right) + \text{Bias}^2 \left( \hat{\beta}_{\text{LE}(\text{PMQL})} \right) \right) \\
= \text{trace} \left( (X'\hat{W}X)^{-1}L_dL_d' + (d - 1)^2\beta' \left( X'\hat{W}X + I \right)^{-2}\beta \right) \\
= \text{trace} \left( (X'\hat{W}X)^{-1}(X'\hat{W}X + dI)'(X'\hat{W}X + I)^{-2}(X'\hat{W}X + dI) \right) + (d - 1)^2\beta' \left( X'\hat{W}X + I \right)^{-2}\beta. \quad (23)
\]

Let us define an orthogonal matrix \( \Gamma \) whose columns are the normalized eigenvectors of the matrix \( X'\hat{W}X \), a vector \( \alpha = \Gamma'\beta, \) and a diagonal matrix \( \Lambda = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_{p+1}) = \Gamma'X\hat{W}XT \). Then, the asymptotic SMSE can be written by using the spectral decomposition as

\[
\text{SMSE} \left( \hat{\beta}_{\text{LE}(\text{PMQL})} \right) = \text{trace} \left( \Lambda'^{-1}(\Lambda + dI)'(\Lambda + dI)^{-2}\Lambda' \right) + (d - 1)^2\beta' \left( (\Lambda + dI)^{-2}\Lambda' \right) \\
= \sum_{j=1}^{p+1} \frac{\lambda_j + d}{\lambda_j} \left( \sum_{j=1}^{p+1} \frac{\lambda_j + d}{\lambda_j} \right)^2 + (d - 1)^2 \sum_{j=1}^{p+1} \frac{\lambda_j^2}{\lambda_j + 1} \\
= \text{term I} + \text{term II}, \text{ say}, \quad (24)
\]

respectively, where \( \alpha_j (j = 1, 2, \ldots, p + 1) \) is the \( j \)th element of \( \Gamma'\beta \) and term I and term II are the total variance of regression coefficient estimates and squared bias, respectively.

### 3.1. MSE Properties of the PMQL Liu Regression Estimator

In this subsection, we discuss the MSE properties of the Liu estimator for the PMQL regression model. Further, we make a comparison of \( \hat{\beta}_{\text{LE}(\text{PMQL})} \) with the existing estimators such as \( \hat{\beta}_{\text{MLE}(\text{PMQL})} \) and \( \hat{\beta}_{\text{LE}(\text{PMQL})} \) to show the superiority of \( \hat{\beta}_{\text{LE}(\text{PMQL})} \) under different conditions in the MSE sense.

Let us define \( \Delta \) which is the SMSE differences of \( \hat{\beta}_{\text{MLE}(\text{PMQL})} \) and \( \hat{\beta}_{\text{LE}(\text{PMQL})} \). By using equations (18) and (24), we get

\[
\Delta = \text{SMSE} \left( \hat{\beta}_{\text{MLE}(\text{PMQL})} \right) - \text{SMSE} \left( \hat{\beta}_{\text{LE}(\text{PMQL})} \right) \\
= \sum_{j=1}^{p+1} \frac{1}{\lambda_j} \left( \sum_{j=1}^{p+1} \frac{(\lambda_j + d)^2}{\lambda_j (\lambda_j + 1)} \right) + (d - 1)^2 \sum_{j=1}^{p+1} \frac{\lambda_j^2}{(\lambda_j + 1)^2}. \quad (25)
\]

It is clear that \( \Delta = 0 \) when \( d \) equals one, and then \( \text{SMSE} \left( \hat{\beta}_{\text{MLE}(\text{PMQL})} \right) = \text{SMSE} \left( \hat{\beta}_{\text{LE}(\text{PMQL})} \right) \). Further, the estimator \( \hat{\beta}_{\text{LE}(\text{PMQL})} \) is said to be superior to the estimator \( \hat{\beta}_{\text{MLE}(\text{PMQL})} \) in the form of the SMSE criterion if and only if \( \Delta > 0 \). Then, if we can find a \( d (0 < d < 1) \) such that \( \Delta > 0 \), we can say that the estimator \( \hat{\beta}_{\text{LE}(\text{PMQL})} \) is superior to \( \hat{\beta}_{\text{MLE}(\text{PMQL})} \) in the PMQL regression model.

Kejian [22] showed that there exists a \( d (0 < d < 1) \) such that the Liu regression estimator has a lower SMSE than the ordinary least square estimator. Further, Kurtoglu and Ozkale [28] have proven that this property holds for the GLMs. The following two propositions show that this property holds in the PMQL Liu regression model.

**Proposition 1.** The total variance of the regression coefficient estimates of \( \hat{\beta}_{\text{LE}(\text{PMQL})} \) (term I) and squared bias of \( \hat{\beta}_{\text{LE}(\text{PMQL})} \) (term II) are continuous monotonically increasing and decreasing functions of \( d \), respectively.

**Proof.** The first derivative of term I in terms of \( d \) is

\[
\frac{\partial \text{term I}}{\partial d} = 2 \sum_{j=1}^{p+1} \frac{(\lambda_j + d)}{(\lambda_j + 1)^2}, \quad (26)
\]

\[
\lim_{d \to 0^+} \frac{\partial \text{term I}}{\partial d} = 2 \sum_{j=1}^{p+1} \frac{1}{(\lambda_j + 1)^2}, \quad (27)
\]

\[
\lim_{d \to 1^-} \frac{\partial \text{term I}}{\partial d} = 2 \sum_{j=1}^{p+1} \frac{1}{\lambda_j (\lambda_j + 1)} \quad (28)
\]
Since $\lambda_j > 0$ for all $j$, equation (26) is always positive for all $d (0 < d < 1)$. Further, the derivative of the term I in the neighborhood of zero given in equation (27) and the derivative of the term I in the neighborhood of one given in equation (28) are positive.

The first derivative of term II in terms of $d$ is

$$\frac{\partial \text{term II}}{\partial d} = 2 (d - 1) \sum_{j=1}^{p+1} \frac{a_j}{(\lambda_j + 1)^2}.$$  \hspace{1cm} (29)

$$\lim_{d \to 0} \frac{\partial \text{term II}}{\partial d} = -2 \sum_{j=1}^{p+1} \frac{a_j}{(\lambda_j + 1)^2}.$$  \hspace{1cm} (30)

$$\lim_{d \to 1} \frac{\partial \text{term II}}{\partial d} = 0.$$  \hspace{1cm} (31)

Since $\lambda_j > 0$ and $a_j^2 > 0$ for all $j$, equation (29) is always negative for all $d (0 < d < 1)$. Further, the derivative of the term II in the neighborhood of zero given in equation (30) is negative, and the derivative of the term II in the neighborhood of one given in equation (31) is zero.

Then, it is shown that term II is continuous monotonically increasing and decreasing functions of $d$, respectively.

\textbf{Proposition 2.} The SMSE $\tilde{\text{SMSE}}(\tilde{\beta}_{\text{LE(pmq1)}})$ given in equation (24) is a continuous monotonically increasing function of $d$ when $(a_{\max}^2 - 1)/(1/\lambda_{\max}) + a_{\max}^2 < d < 1$, where $a_{\max}^2$ is the maximum element of $a_j^2$ ($j = 1, 2, \ldots, p + 1$) and $\lambda_{\max}$ is the maximum eigenvalue of the matrix $X'WX$.

\textbf{Proof.} The first derivative of equation (24) is

$$\frac{d \text{SMSE}(\tilde{\beta}_{\text{LE(pmq1)}})}{dd} = \frac{p+1}{\lambda_j + 1} \sum_{j=1}^{p+1} \frac{a_j^2}{\lambda_j + 1} + d \sum_{j=1}^{p+1} \left( \frac{a_j^2}{\lambda_j + 1} \right)^2.$$  \hspace{1cm} (32)

One can note that if the individual Liu parameter $d_j > (a_j^2 - 1)/(1/\lambda_j) + a_j^2$ for all $j = 1, 2, \ldots, p + 1$, equation (32) is positive. Then, it is clear that when $(a_{\max}^2 - 1)/(1/\lambda_{\max}) + a_{\max}^2 < d < 1$, equation (32) is always positive. Then, it is shown that the SMSE $\tilde{\text{SMSE}}(\tilde{\beta}_{\text{LE(pmq1)}})$ is a continuous monotonically increasing function of $d$ when $(a_{\max}^2 - 1)/(1/\lambda_{\max}) + a_{\max}^2 < d < 1$

Now, we can conclude that there is a possibility of finding a value $d (0 < d < 1)$ based on the results of Proposition 1. Further, the results of $\Delta = 0$ given in equation (25) at $d$ equal one, and Proposition 2 reveals that the $\Delta > 0$ when $(a_{\max}^2 - 1)/(1/\lambda_{\max}) + a_{\max}^2 < d < 1$

Then, it is shown that there exists a $d (0 < d < 1)$ such that

$$\text{SMSE} (\tilde{\beta}_{\text{LE(pmq1)}}) < \text{SMSE} (\tilde{\beta}_{\text{MLE(pmq1)}}).$$

The following theorem discusses the condition that the LE is superior to the MLE in the PMQL regression model.

\textbf{Theorem 1.} Let $\lambda_j (\lambda_j + 1)^2 - (\lambda_j + d)^2 > 0$, ($j = 1, 2, \ldots, p + 1$) and $b_d = \text{Bias} (\tilde{\beta}_{\text{LE(pmq1)}})$, where $\text{bias} (\tilde{\beta}_{\text{LE(pmq1)}})$, $\text{MSE} (\tilde{\beta}_{\text{LE(pmq1)}}) > 0$ if $b_d (\Gamma (\lambda + I)^{-1} (\lambda + d I) - \lambda^{-1} (\lambda + 1)^{-1} I')^{-1} b_d < 1$

\textbf{Proof.} The difference between the MSE of the MLE and LE is derived as

$$\text{MSE} (\tilde{\beta}_{\text{MLE(pmq1)}}) - \text{MSE} (\tilde{\beta}_{\text{LE(pmq1)}}) = X'WX + dI - b_d b_d' - b_d b_d'$$  \hspace{1cm} (33)

Now, we apply the spectral decomposition for the above matrix. Then, the difference can be written as

$$\text{MSE} (\tilde{\beta}_{\text{MLE(pmq1)}}) - \text{MSE} (\tilde{\beta}_{\text{LE(pmq1)}}) = \frac{1}{\lambda_j} \left( \frac{\lambda_j + d}{\lambda_j (\lambda_j + 1)} \right)^2 \Gamma' - b_d b_d'.$$  \hspace{1cm} (34)

$$\Gamma = \text{diag} \left( \frac{\lambda_j + d}{\lambda_j (\lambda_j + 1)} \right).$$
The diagonal matrix \(A^{-1} - (\Lambda + I)^{-1}(\Lambda + dI)A^{-1} (\Lambda + dI) (\Lambda + I)^{-1}\) is pd if \(\lambda_j (\lambda_j + 1)^2 - (\lambda_j + d)^2 (\lambda_j + k)^2 > 0\), \(j = 1, 2, \ldots, p + 1\). Then, by Lemma A1 (Appendix 1), MSE \((\hat{\beta}_{\text{MLE}(PMQL)}) - \text{MSE}(\hat{\beta}_{\text{LE}(PMQL)}) > 0\) if \(b_d^J (\Lambda^{-1} - (\Lambda + I)^{-1}(\Lambda + dI)\Lambda^{-1} (\Lambda + dI) (\Lambda + I)^{-1})^{-1} b_d < 1\). It completes the proof.

The following theorem discusses the condition that the LE is superior to the RE in the PMQL regression model. □

Theorem 2. Let \(\lambda_j^2 (\lambda_j + 1)^2 - (\lambda_j + d)^2 (\lambda_j + k)^2 > 0\), \(j = 1, 2, \ldots, p + 1\) and \(b_k = \text{Bias}(\hat{\beta}_{\text{RE}(PMQL)}) = -k(\lambda^ TWX + kI)^{-1}\). Then, MSE \((\hat{\beta}_{\text{RE}(PMQL)}) - \text{MSE}(\hat{\beta}_{\text{LE}(PMQL)}) > 0\) iff \(b_d^J (\Lambda^{-1} (\Lambda + kI)^{-1} (\Lambda + dI) (\Lambda + dI) (\Lambda + I)^{-1})^{-1} b_d < 1\).

Proof: The difference between the MSE of the RE and LE is derived as

\[
\text{MSE}(\hat{\beta}_{\text{RE}(PMQL)}) - \text{MSE}(\hat{\beta}_{\text{LE}(PMQL)}) = (X' \tilde{W} X + kI)^{-1} (X' \tilde{W} X) (X' \tilde{W} X + kI)^{-1} - (X' \tilde{W} X + I)^{-1} (X' \tilde{W} X + dI) (X' \tilde{W} X + I)^{-1} + b_d^J - b_d \]

By applying the spectral decomposition for the above matrix, the difference can be written as

\[
\text{MSE}(\hat{\beta}_{\text{RE}(PMQL)}) - \text{MSE}(\hat{\beta}_{\text{LE}(PMQL)}) = \Gamma' \left( (\Lambda + kI)^{-1} \Lambda (\Lambda + kI)^{-1} - (\Lambda + I)^{-1} (\Lambda + dI) \Lambda^{-1} (\Lambda + dI) (\Lambda + I)^{-1} \right) \Gamma' + b_d^J - b_d \]

It is clear that \(b_d^J - b_d\) is a pd matrix, and the diagonal matrix \((\Lambda + kI)^{-1} \Lambda (\Lambda + kI)^{-1} - (\Lambda + I)^{-1} (\Lambda + dI) \Lambda^{-1} (\Lambda + dI) (\Lambda + I)^{-1}\) is pd if \(\lambda_j^2 (\lambda_j + 1)^2 - (\lambda_j + d)^2 (\lambda_j + k)^2 > 0\), \(j = 1, 2, \ldots, p + 1\).

Then, by Lemma A1 (Appendix 1), MSE \((\hat{\beta}_{\text{RE}(PMQL)}) - \text{MSE}(\hat{\beta}_{\text{LE}(PMQL)}) > 0\) if \(b_d^J (\Lambda^{-1} (\Lambda + kI)^{-1} (\Lambda + dI) (\Lambda + I)^{-1})^{-1} b_d < 1\). It completes the proof. □

3.2. Estimation of the Liu Parameter \(d\). Based on the MSE properties of the PMQL Liu regression estimator discussed in Section 3.1, it is clear that the performance of \(\hat{\beta}_{\text{LE}(PMQL)}\) depends on the optimum value of the Liu parameter \(d\). The optimal value of any individual Liu parameter \(d_j\) can be found by setting equation (32) to zero and solving for \(d_j\). Then, it is obtained as

\[
d_j = \frac{\alpha_j^2 - 1}{(1/\lambda_j) + \alpha_j} \quad j = 1, 2, \ldots, p + 1. \quad (37)
\]

From equation (37), we can note that the optimum value is negative when \(\alpha_j^2 < 1\) \((j = 1, 2, \ldots, p + 1)\) and otherwise positive. Since the value of the \(d\) is limited between 0 and 1, we should use the “max” operator to ensure the estimated value of the \(d\) is nonnegative.
In this subsection, we adopt some notable existing Liu parameter estimators in order to estimate the Liu parameter in \( \hat{\beta}_{LE(PMQL)} \). They are summarized in Table 1 (Appendix 2). We define the Liu parameter estimators \( D_1, D_2 - D_3 \), and \( D_4 - D_5 \) for \( \hat{\beta}_{LE(PMQL)} \) based on the theoretical works of Hoerl and Kennard [20], Kibria [35], and Khalaf and Shukur [36], respectively.

To estimate the ridge parameter in the PMQL ridge regression estimator, Tharshan and Wijekoon [19] discussed 12 various estimation methods. Among all, they recommended three ridge parameter estimators, \( \tilde{k}_q, \tilde{k}_r, \) and \( \tilde{k}_{12} \), based on the works of Hoerl and Kennard [20], Nomura [37], and Muniz and Kibria [38]. These ridge parameter estimators are summarized in Table 1 (Appendix 2). Therefore, \( \tilde{k}_q, \tilde{k}_r, \) and \( \tilde{k}_{12} \) will be utilized to estimate \( k \) in \( \hat{\beta}_{RE(PMQL)} \) in the simulation study.

4. The Monte Carlo Simulation Study

In this section, a simulation study is carried out to evaluate the performance of the MLE(PMQL), PMQL ridge regression estimators, PMQL Liu regression estimators based on the various ridge, and Liu parameter estimators, respectively, as discussed in Section 3.2. We compare the performance of different estimators in the SMSE sense. A brief discussion about the simulation study is given in the following.

4.1. The Design of the Simulation Study. Since the degrees of the correlation (\( \rho \)) between the covariates greatly depends on the performance of the various estimators, we generate the covariates with several degrees of multicollinearity by following the same formula as used by McDonald and Galarneau [39]. The formula is given as follows:

\[
x_{ij} = \sqrt{(1 - \rho^2)m_{ij} + pm_{i,p+1}},
\]

\[
i = 1, 2, \ldots, n, \quad j = 1, 2, \ldots, p,
\]

where \( m_{ij} \)'s are independent standard normal pseudorandom numbers. The response variable \( y \) of the PMQL regression model is generated from the PMQL \((\mu_i, \alpha, \delta)\) by using the inverse transform method, where \( u_i = \exp(x'_i\beta) \), \( i = 1, 2, \ldots, n \). The starting values of the slope parameters are selected such that \( \sum_j \hat{\beta}_j^2 = 1 \) and \( \hat{\beta}_1 = \hat{\beta}_2 = \cdots = \hat{\beta}_p \).

Table 2 summarizes the factors and their levels that are considered in this design. Since either higher increments or decrements of variation of \( y \) may lead to a negative impact on the performance of estimators [19, 24, 40, 41], we vary \( \beta_0, \alpha, \) and \( \delta \). When we decrease the value of \( \beta_0 \), the average values of the \( y_i(\mu_i), \) \( i = 1, 2, \ldots, n \) will decrease. This phenomenon leads to having more zeros of \( y \), which makes very less variation in the sample. Further, from Figure 1, we can observe that changing the value of the overdispersion parameters \( \alpha, \) or \( \delta \) affects the variation of \( y \).

The simulation is repeated 1000 times. To judge the performance of the different estimators, we obtain the SMSE values of different estimators by using the following equation:

\[
SMSE(\hat{\beta}) = \frac{\sum_{i=1}^{1000}(\hat{\beta}_r - \beta)^2}{1000},
\]

where \( \hat{\beta}_r \) is an estimator of \( \beta \) at the \( r \)th replication.

4.2. Results and Discussion of the Simulation Study. The estimated SMSE values of the Monte Carlo Simulations are summarized in Tables 3–10 (Appendix 2) for the selected situations shown in Table 2. The minimum SMSE in each combination of different factor levels is presented in bold. In general, we can note that the LE(PMQL) is more efficient than the MLE(PMQL) and RE(PMQL) in all cases reviewed in this study. Further, the performances of all regression coefficient estimators are affected by the factors of degrees of the correlation among the covariates, the sample size, the value of the intercept, the number of covariates, and the values of the overdispersion parameters.

It can be noticed that as the degrees of correlation increase, the estimated SMSE of the MLE(PMQL) increases, and the PMQL Liu regression estimators having the estimated \( d \) values \( D_1, D_2, D_3, \) and \( D_4 \) are also affected negatively in some cases. However, the PMQL Liu regression estimator based on the \( D_5 \) estimator does not affect and yields consistently a smaller SMSE in all cases. Further, its estimated SMSE decreases with \( \rho \).

When the sample size increases, SMSE of the MLE(PMQL) and the PMQL ridge and Liu regression estimators decrease. This phenomenon reveals that the asymptotic property holds for all given estimators. Further, in a given sample size, the PMQL Liu regression estimator based on the \( D_5 \) estimator performs better than the MLE(PMQL) and PMQL ridge regression estimators for all given situations.

It is clearly observed that as the \( \beta_0 \) decreases from 1 to \(-1\), the SMSE of the MLE(PMQL) and PMQL Liu regression estimators based on the \( D_1, D_3, \) and \( D_4 \) estimators are increased with a higher amount and based on \( D_2 \) estimator are also affected negatively for some cases. However, this change does not affect the performance of the PMQL Liu regression estimator based on the \( D_5 \) estimator basically.

The increasing number of covariates affects the performance of the MLE(PMQL) negatively in all given situations, and the PMQL ridge regression estimators also affect negatively in some given situations. However, the PMQL Liu regression estimator based on \( D_5 \) estimator does not affect when increasing the number of covariates in all given situations. Further, when \( p \) increases for a given combination of \( \rho, n, \beta_0, \alpha, \delta \), the SMSEs produced by PMQL Liu regression estimator based on \( D_5 \) estimator decrease.

We can clearly note that the increment of either the overdispersion parameter \( \alpha (0.03 \text{ to } 0.25) \) or \( \delta (0.02 \text{ to } 0.04) \) shows a positive impact on the performance of different estimators. From Figure 1, we can note that, in both situations, the variance of \( y \) decreases. These results are in line with the simulated results of [17, 41].

Then, based on the simulation study results, we may say that the Liu parameter estimator \( D_6 \) is the best option to estimate the Liu parameter \( d \) in the \( \hat{\beta}_{LE(PMQL)} \).
Table 1: The Liu and ridge parameter estimators for the PMQL Liu and ridge regression models.

| Author(s)                  | Liu and ridge parameter estimators                                      |
|----------------------------|--------------------------------------------------------------------------|
| Hoen and Kennard [20]      | $\hat{D}_1 = \max(0, \tilde{a}^2 + 1/(1/\lambda_{\text{max}} + \tilde{a}^2_{\text{max}}))$ |
| Kibria [35]                | $\hat{D}_2 = \max(0, \text{median}(\tilde{a}^2 + 1/(1/\lambda_{\text{j}} + \tilde{a}^2_{\text{j}})))$ |
| Khalaf and Shukur [36]     | $\hat{D}_3 = \max(0, \sum_{j=1}^{p} (\tilde{a}^2 - 1/(1/\lambda_{\text{j}} + \tilde{a}^2_{\text{j}}))(p + 1))$ |
| Nomura [37]                | $\hat{D}_4 = \max(0, \max(\tilde{a}^2 - 1/(1/\lambda_{\text{j}} + \tilde{a}^2_{\text{j}})))$ |
| Muniz and Kibria [38]      | $\hat{k}_5 = p + 1/(\sum_{j=1}^{p} (\tilde{a}^2_{\text{j}})/(1 + (1 + \lambda_{\text{j}})(\tilde{a}^2_{\text{j}})))$. |
|                            | $\hat{k}_{12} = \text{median}(\sqrt{1/\tilde{a}^2_{\text{j}}})$, $j = 1, 2, \ldots, p + 1$, |
|                            | To examine the performance of different estimators when increasing $\delta$. |

Table 2: Factors and their levels that are considered in the design of the simulation study.

| Factors                  | Levels | Purpose                                                      |
|--------------------------|--------|--------------------------------------------------------------|
| Degrees of correlation ($\rho$) | 0.99, 0.999, and 0.9999 | To examine the performance of different estimators when increasing $\rho$. |
| Sample size ($n$)        | 10, 20, 30, 40, and 100 | To examine the asymptotic property and the performance of different estimators with $n$. |
| Number of covariates ($k$) | 2 and 4          | To examine the performance of the various estimators when increasing $p$. |
| Intercept ($\beta_0$)     | 1 and 2         | (i) By changing $\beta_0$.                                  |
| Overdispersion parameter 1 ($\alpha$) | 0.03 and 0.25 | (ii) By changing $\alpha$.                                  |
| Overdispersion parameter 2 ($\delta$) | 0.25 and 0.04  | (iii) By changing $\delta$.                                |

Table 3: Estimated SMSE values for different $n, p, k, d,$ and $\beta_0$ when $\delta = 0.02$, $\alpha = 0.03$, and $p = 2$.

| Estimator | $\beta_0 = -1$ | $\beta_0 = 1$ |
|-----------|----------------|---------------|
| $\rho = 0.99$ |                  |               |
| $n = 10$  | 452.489         | 174.282       |
| $n = 20$  | 108.926         | 78.421        |
| $n = 30$  | 29.614          | 414.628       |
| $n = 40$  | 159.988         | 71.908        |
| $n = 50$  | 27.170          |               |
| $n = 100$ | 1.711           | 1.474         |
| $n = 200$ | 1.297           | 1.166         |
| $n = 300$ | 1.116           | 0.986         |
| $n = 400$ | 0.605           | 0.574         |
| $n = 500$ | 1.564           | 1.567         |
| $n = 600$ | 1.289           | 1.295         |
| $n = 700$ | 1.104           | 1.108         |
| $n = 800$ | 0.972           | 0.974         |
| $n = 900$ | 0.574           | 0.582         |

| $\rho = 0.999$ |                  |               |
| $n = 10$  | 4396.674         | 1692.037      |
| $n = 20$  | 1056.591         | 760.296       |
| $n = 30$  | 4154.024         | 2860.172      |
| $n = 40$  | 1554.024         | 10177.070     |
| $n = 50$  | 6971.687         | 41077.070     |
| $n = 60$  | 2625.934         | 15496.860     |
| $n = 70$  | 1.565           | 1.289         |
| $n = 80$  | 1.104           | 0.972         |
| $n = 90$  | 0.574           | 0.574         |

| $\rho = 0.9999$ |                  |               |
| $n = 10$  | 43836.060        | 16872.260     |
| $n = 20$  | 10534.020        | 7579.503      |
| $n = 30$  | 30177.070        | 2860.172      |
| $n = 40$  | 15496.860        | 10177.070     |
| $n = 50$  | 6954.122         | 41077.070     |
| $n = 60$  | 2625.934         | 15496.860     |
| $n = 70$  | 1.565           | 1.289         |
| $n = 80$  | 1.104           | 0.972         |
| $n = 90$  | 0.574           | 0.574         |
5. Applications

In this section, we use a simulated data set to show the applicability of the PMQL Liu regression model for the count data set with a higher index of dispersion and long right tail over the existing Liu regression models for count data, namely, NB Liu regression and Poisson Liu regression models. Further, we illustrate the applicability of the PMQL Liu regression estimator over the MLE, by using a real data set also. We have observed that the PMQL Liu regression estimator based on $D_e$ estimator performs well in the simulation study. Further, the same Liu parameter

### Table 3: Continued.

| Estimator | $\beta_0 = -1$ | $\beta_0 = 1$ |
|-----------|----------------|----------------|
|           | $n = 10$       | $n = 20$       | $n = 30$       | $n = 40$       | $n = 100$      | $n = 20$       | $n = 30$       | $n = 40$       | $n = 100$      |
| LE ($D_1$) | 8201.712       | 4973.109       | 3898.417       | 3163.635       | 1503.131       | 1.561          | 1.285          | 1.094          | 0.956          | 0.545          |
| LE ($D_2$) | 208.113        | 165.643        | 132.270        | 101.672        | 71.001         | 1.561          | 1.285          | 1.094          | 0.956          | 0.545          |
| LE ($D_3$) | 336.983        | 208.402        | 154.438        | 118.589        | 36.359         | 1.561          | 1.285          | 1.094          | 0.956          | 0.545          |
| LE ($D_4$) | 4596.616       | 3511.205       | 2954.507       | 2525.475       | 1356.645       | 1.561          | 1.285          | 1.094          | 0.956          | 0.545          |
| LE ($D_e$) | 1.559          | 1.288          | 1.103          | 0.970          | 0.571          | 1.561          | 1.285          | 1.094          | 0.956          | 0.545          |

Note. RE (a) indicates the ridge estimator based on the ridge parameter estimator $a$, LE (b) indicates the Liu estimator based on the Liu parameter estimator $b$, and bold values indicate the minimum SMSEs.

### Table 4: Estimated SMSE values for different $n, \rho, k, d,$ and $\beta_0$ when $\delta = 0.04, \alpha = 0.03$, and $p = 2$.

| Estimator | $\beta_0 = -1$ | $\beta_0 = 1$ |
|-----------|----------------|----------------|
|           | $n = 10$       | $n = 20$       | $n = 30$       | $n = 40$       | $n = 100$      | $n = 20$       | $n = 30$       | $n = 40$       | $n = 100$      |
| MLE       | 245.039        | 93.937         | 58.668         | 42.234         | 15.927         | 210.562        | 81.256         | 50.776         | 36.539         | 13.806         |
| RE ($k_1$)| 1.362          | 1.084          | 0.927          | 0.831          | 0.595          | 1.760          | 1.631          | 1.335          | 1.368          |
| RE ($k_2$)| 1.433          | 1.109          | 0.902          | 0.763          | 0.450          | 1.321          | 0.992          | 0.802          | 0.680          | 0.393          |
| RE ($k_3$)| 1.529          | 1.254          | 1.074          | 0.944          | 0.529          | 1.577          | 1.380          | 1.267          | 1.199          | 0.928          |
| LE ($D_1$)| 70.111         | 37.569         | 27.385         | 21.437         | 9.560          | 1.315          | 0.988          | 0.799          | 0.677          | 0.391          |
| LE ($D_2$)| 3.014          | 1.915          | 1.413          | 1.064          | 0.455          | 1.315          | 0.988          | 0.799          | 0.677          | 0.391          |
| LE ($D_3$)| 3.980          | 2.187          | 1.530          | 1.135          | 0.480          | 1.315          | 0.988          | 0.799          | 0.677          | 0.391          |
| LE ($D_4$)| 48.183         | 31.266         | 24.062         | 19.472         | 9.242          | 1.315          | 0.988          | 0.799          | 0.677          | 0.391          |
| LE ($D_e$)| 1.318          | 1.006          | 0.826          | 0.710          | 0.423          | 1.315          | 0.988          | 0.799          | 0.677          | 0.391          |

Note. RE (a) indicates the ridge estimator based on the ridge parameter estimator $a$, LE (b) indicates the Liu estimator based on the Liu parameter estimator $b$, and bold values indicate the minimum SMSEs.
### Table 5: Estimated SMSE values for different $n, \rho, k, d,$ and $\beta_0$ when $\delta = 0.02, \alpha = 0.25,$ and $p = 2$.

| Estimator | $\beta_0 = -1$ | $\beta_0 = 1$ |
|-----------|----------------|----------------|
|           | $n = 10$ | $n = 20$ | $n = 30$ | $n = 40$ | $n = 100$ | $n = 20$ | $n = 30$ | $n = 40$ | $n = 100$ |
| MLE       | 372.624 | 143.374 | 89.597 | 64.506 | 24.352 | 335.750 | 129.559 | 80.947 | 58.240 | 22.006 |
| RE ($k_1$) | 1.524 | 1.268 | 1.106 | 0.999 | 0.709 | 1.830 | 1.721 | 1.646 | 1.593 | 1.440 |
| RE ($k_1$) | 1.596 | 1.323 | 1.127 | 0.984 | 0.556 | 1.500 | 1.205 | 1.012 | 0.879 | 0.513 |
| RE ($k_{12}$) | 1.657 | 1.408 | 1.225 | 1.071 | 0.611 | 1.712 | 1.562 | 1.468 | 1.410 | 1.171 |
| LE ($D_1$) | 81.482 | 48.077 | 36.358 | 29.354 | 13.841 | 1.495 | 1.202 | 1.010 | 0.877 | 0.512 |
| LE ($D_1$) | 3.476 | 2.801 | 2.138 | 1.738 | 0.655 | 1.495 | 1.202 | 1.010 | 0.877 | 0.512 |
| LE ($D_2$) | 4.793 | 3.092 | 2.332 | 1.860 | 0.783 | 1.495 | 1.202 | 1.010 | 0.877 | 0.512 |
| LE ($D_2$) | 49.919 | 36.118 | 29.627 | 24.850 | 12.838 | 1.495 | 1.202 | 1.010 | 0.877 | 0.512 |
| LE ($D_3$) | 1.494 | 1.207 | 1.023 | 0.895 | 0.539 | 1.495 | 1.202 | 1.010 | 0.877 | 0.512 |

Note: RE ($a$) indicates the ridge estimator based on the ridge parameter estimator $a$, LE ($b$) indicates the Liu estimator based on the Liu parameter estimator $b$, and bold values indicate the minimum SMSEs.

### Table 6: Estimated SMSE values for different $n, \rho, k, d,$ and $\beta_0$ when $\delta = 0.04, \alpha = 0.25,$ and $p = 2$.

| Estimator | $\beta_0 = -1$ | $\beta_0 = 1$ |
|-----------|----------------|----------------|
|           | $n = 10$ | $n = 20$ | $n = 30$ | $n = 40$ | $n = 100$ | $n = 20$ | $n = 30$ | $n = 40$ | $n = 100$ |
| MLE       | 3609.870 | 1391.796 | 1056.591 | 625.309 | 235.911 | 3262.905 | 1258.426 | 785.386 | 564.931 | 213.346 |
| RE ($k_1$) | 1.520 | 1.263 | 1.184 | 0.993 | 0.704 | 1.829 | 1.720 | 1.645 | 1.593 | 1.441 |
| RE ($k_1$) | 1.594 | 1.319 | 1.222 | 0.978 | 0.546 | 1.493 | 1.493 | 1.563 | 1.471 | 1.317 |
| RE ($k_{12}$) | 1.661 | 1.406 | 1.291 | 1.061 | 0.589 | 1.712 | 1.712 | 1.816 | 1.638 | 1.417 |
| LE ($D_1$) | 787.324 | 458.141 | 348.280 | 276.544 | 128.660 | 1.489 | 1.191 | 0.995 | 0.858 | 0.472 |
| LE ($D_1$) | 21.560 | 15.514 | 14.279 | 8.557 | 1.508 | 1.491 | 1.491 | 1.510 | 1.491 | 0.872 |
| LE ($D_1$) | 33.215 | 18.962 | 16.545 | 9.874 | 2.662 | 1.489 | 1.191 | 0.995 | 0.858 | 0.472 |
| LE ($D_1$) | 468.722 | 339.052 | 297.506 | 231.827 | 118.588 | 1.489 | 1.191 | 0.995 | 0.858 | 0.472 |
| LE ($D_1$) | 1.488 | 1.197 | 1.104 | 0.878 | 0.502 | 1.489 | 1.191 | 0.995 | 0.858 | 0.472 |

Note: RE ($a$) indicates the ridge estimator based on the ridge parameter estimator $a$, LE ($b$) indicates the Liu estimator based on the Liu parameter estimator $b$, and bold values indicate the minimum SMSEs.

### Table 7: Estimated SMSE values for different $n, \rho, k, d,$ and $\beta_0$ when $\delta = 0.02, \alpha = 0.25,$ and $p = 2$.

| Estimator | $\beta_0 = -1$ | $\beta_0 = 1$ |
|-----------|----------------|----------------|
|           | $n = 10$ | $n = 20$ | $n = 30$ | $n = 40$ | $n = 100$ | $n = 20$ | $n = 30$ | $n = 40$ | $n = 100$ |
| MLE       | 227.529 | 87.144 | 54.416 | 39.171 | 14.768 | 193.512 | 74.676 | 46.666 | 33.583 | 12.689 |
| RE ($k_1$) | 1.330 | 1.050 | 0.896 | 0.802 | 0.577 | 1.746 | 1.613 | 1.533 | 1.483 | 1.355 |
| RE ($k_1$) | 1.400 | 1.068 | 0.862 | 0.725 | 0.416 | 1.285 | 0.954 | 0.766 | 0.648 | 0.376 |
| RE ($k_{12}$) | 1.503 | 1.228 | 1.052 | 0.924 | 0.521 | 1.548 | 1.343 | 1.226 | 1.157 | 0.882 |
| LE ($D_1$) | 67.852 | 36.245 | 25.990 | 20.327 | 8.977 | 1.279 | 0.949 | 0.763 | 0.665 | 0.374 |
| LE ($D_1$) | 2.899 | 1.772 | 1.305 | 0.976 | 0.430 | 1.279 | 0.949 | 0.763 | 0.665 | 0.374 |
| LE ($D_1$) | 3.811 | 2.033 | 1.413 | 1.052 | 0.448 | 1.279 | 0.949 | 0.763 | 0.665 | 0.374 |
| LE ($D_1$) | 47.518 | 30.250 | 23.033 | 18.542 | 8.692 | 1.279 | 0.949 | 0.763 | 0.665 | 0.374 |
| LE ($D_1$) | 1.284 | 0.970 | 0.793 | 0.680 | 0.407 | 1.279 | 0.949 | 0.763 | 0.665 | 0.374 |

Note: RE ($a$) indicates the ridge estimator based on the ridge parameter estimator $a$, LE ($b$) indicates the Liu estimator based on the Liu parameter estimator $b$, and bold values indicate the minimum SMSEs.
### Table 6: Estimated SMSE values for different $n, \rho, k, d$, and $\beta_0$ when $\delta = 0.02$, $\alpha = 0.03$, and $p = 4$.

| Estimator | $n = 10$ | $n = 20$ | $n = 40$ | $n = 100$ | $\beta_0 = -1$ | $\beta_0 = 1$ |
|-----------|----------|----------|----------|-----------|---------------|---------------|
| MLE       | 22036.400 | 8431.337 | 5258.827 | 3783.276  | 1425.136      | 18750.330     |
| RE ($k_1$) | 1.325    | 1.044    | 0.890    | 0.796     | 0.573        | 1.745         |
| RE ($k_2$) | 1.396    | 1.062    | 0.854    | 0.716     | 0.361        | 1.274         |
| RE ($k_3$) | 1.514    | 1.230    | 1.052    | 0.915     | 0.503        | 1.548         |
| LE ($D_1$) | 6435.778 | 3365.452 | 2444.748 | 1882.106  | 822.170       | 1.269         |
| LE ($D_2$) | 161.526  | 74.435   | 44.844   | 29.430    | 2.132        | 1.269         |
| LE ($D_3$) | 241.493  | 101.296  | 58.745   | 35.786    | 3.880        | 1.269         |
| LE ($D_4$) | 4456.390 | 2815.842 | 2135.848 | 1710.447  | 788.719      | 1.269         |
| LE ($D_5$) | 1.275    | 0.954    | 0.771    | 0.652     | 0.344        | 1.269         |

Note. RE (a) indicates the ridge estimator based on the ridge parameter estimator $a$, RE (b) indicates the Liu estimator based on the Liu parameter estimator $b$, and bold values indicate the minimum SMSEs.

### Table 7: Estimated SMSE values for different $n, \rho, k, d$, and $\beta_0$ when $\delta = 0.02$, $\alpha = 0.03$, and $p = 4$.

| Estimator | $n = 10$ | $n = 20$ | $n = 40$ | $n = 100$ | $\beta_0 = -1$ | $\beta_0 = 1$ |
|-----------|----------|----------|----------|-----------|---------------|---------------|
| MLE       | 2300.301 | 639.879  | 369.034  | 259.458   | 93.662        | 1987.509      |
| RE ($k_1$) | 1.512    | 1.257    | 1.100    | 0.994     | 0.720        | 1.850         |
| RE ($k_2$) | 1.537    | 1.252    | 1.060    | 0.925     | 0.589        | 1.441         |
| RE ($k_3$) | 1.809    | 1.736    | 1.687    | 1.645     | 1.484        | 1.944         |
| LE ($D_1$) | 640.236  | 259.627  | 172.729  | 130.386   | 53.893        | 1.439         |
| LE ($D_2$) | 1.437    | 1.155    | 0.984    | 0.874     | 0.587        | 1.439         |
| LE ($D_3$) | 6.649    | 3.124    | 2.185    | 1.653     | 0.733        | 1.439         |
| LE ($D_4$) | 220.533  | 129.204  | 102.887  | 87.006    | 45.331       | 1.439         |
| LE ($D_5$) | 1.437    | 1.155    | 0.984    | 0.872     | 0.587        | 1.439         |

Note. RE (a) indicates the ridge estimator based on the ridge parameter estimator $a$, RE (b) indicates the Liu estimator based on the Liu parameter estimator $b$, and bold values indicate the minimum SMSEs.

Estimator was recommended by Månsson [25] for the NB Liu regression model and by Månsson [23] for the Poisson Liu regression model. Then, we will use the $D_3$ estimator to estimate the Liu parameter in all considered different Liu regression estimators.

5.1. Simulated Data Application. A data set with $\rho = 4$, $\delta = 0.9999$, and $n = 400$ was simulated by using the method discussed in Section 4.1 in order to show the applicability of the proposed Liu regression model over the NB Liu and Poisson Liu regression models. The skewness, excess...
### Table 8: Estimated SMSE values for different $n, \rho, k, d,$ and $\beta_0$ when $\delta = 0.04$, $\alpha = 0.03$, and $p = 4$.

| Estimator | $n = 10$ | $n = 20$ | $n = 30$ | $n = 40$ | $n = 100$ | $n = 20$ | $n = 30$ | $n = 40$ | $n = 100$ |
|-----------|----------|----------|----------|----------|-----------|----------|----------|----------|-----------|
| MLE       | 1290.178 | 351.061  | 201.273  | 141.149  | 50.739    | 1023.268 | 289.402  | 167.742  | 118.192   |
| RE ($k_1$)| 1.278     | 1.001    | 0.854    | 0.764    | 0.559     | 1.770    | 1.675    | 1.623    | 1.592     |
| RE ($k_2$)| 1.286     | 0.952    | 0.762    | 0.694    | 0.517     | 1.185    | 0.881    | 0.725    | 0.634     |
| LE ($D_1$)| 1.667     | 1.544    | 1.468    | 1.405    | 1.184     | 1.893    | 1.854    | 1.827    | 1.802     |
| LE ($D_2$)| 4.821     | 172.171  | 108.377  | 79.652   | 31.243    | 1.183    | 0.879    | 0.723    | 0.632     |
| LE ($D_3$)| 1.185     | 0.895    | 0.747    | 0.660    | 0.488     | 1.183    | 0.879    | 0.723    | 0.632     |
| LE ($D_4$)| 4.905     | 1.815    | 1.145    | 0.843    | 0.492     | 1.183    | 0.879    | 0.723    | 0.632     |
| LE ($D_5$)| 227.801   | 109.915  | 78.862   | 62.425   | 28.354    | 1.183    | 0.879    | 0.723    | 0.632     |
| LE ($D_6$)| 1.185     | 0.895    | 0.747    | 0.660    | 0.488     | 1.183    | 0.879    | 0.723    | 0.632     |

Note: RE (a) indicates the ridge estimator based on the ridge parameter estimator $a$, LE (b) indicates the Liu estimator based on the Liu parameter estimator $b$, and bold values indicate the minimum SMSEs.

### Table 9: Estimated SMSE values for different $n, \rho, k, d,$ and $\beta_0$ when $\delta = 0.02$, $\alpha = 0.25$, and $p = 4$.

| Estimator | $n = 10$ | $n = 20$ | $n = 30$ | $n = 40$ | $n = 100$ | $n = 20$ | $n = 30$ | $n = 40$ | $n = 100$ |
|-----------|----------|----------|----------|----------|-----------|----------|----------|----------|-----------|
| MLE       | 1913.478 | 529.367  | 304.832  | 214.179  | 127.854   | 1615.252 | 457.738  | 265.474  | 187.103   |
| RE ($k_1$)| 1.270     | 0.993    | 0.845    | 0.756    | 0.553     | 1.769    | 1.674    | 1.623    | 1.592     |
| RE ($k_2$)| 1.279     | 0.941    | 0.749    | 0.628    | 0.352     | 1.162    | 0.841    | 0.667    | 0.558     |
| RE ($k_3$)| 1.865     | 1.816    | 1.783    | 1.753    | 1.633     | 1.963    | 1.949    | 1.939    | 1.930     |
| LE ($D_1$)| 482.020   | 1696.644 | 1052.622 | 776.154  | 299.176   | 1.161    | 0.840    | 0.666    | 0.557     |
| LE ($D_2$)| 1.166     | 0.859    | 0.692    | 0.596    | 0.340     | 1.161    | 0.840    | 0.666    | 0.557     |
| LE ($D_3$)| 38.622    | 9.836    | 4.545    | 2.340    | 0.401     | 1.161    | 0.840    | 0.666    | 0.557     |
| LE ($D_4$)| 2240.761  | 1069.435 | 763.603  | 602.223  | 268.874   | 1.161    | 0.840    | 0.666    | 0.557     |
| LE ($D_5$)| 1.166     | 0.861    | 0.698    | 0.596    | 0.340     | 1.161    | 0.840    | 0.666    | 0.557     |
| LE ($D_6$)| 1.166     | 0.857    | 0.693    | 0.589    | 0.323     | 1.159    | 0.836    | 0.660    | 0.549     |

Note: RE (a) indicates the ridge estimator based on the ridge parameter estimator $a$, LE (b) indicates the Liu estimator based on the Liu parameter estimator $b$, and bold values indicate the minimum SMSEs.
kurtosis, and index of dispersion of the response variable \( y \) are 4.628, 20.113, and 35.217, respectively. Then, the distribution of \( y \) has higher positive skewness, a long right tail, and higher overdispersion (index of dispersion > 1). Figure 2 also illustrates the distribution of the response variable \( y \). Table 11 displays the estimated regression coefficients, their standard errors (SEs) (in parentheses), SMSE values, and Akaike information criterion (AIC) values for the given regression models. We can clearly observe that the PMQL Liu regression model produces smaller SEs for the

| Table 9: Continued. |
|----------------------|
| \( \rho = 0.99 \) |
| Estimator          | \( n = 10 \) | \( n = 20 \) | \( n = 30 \) | \( n = 40 \) | \( n = 100 \) | \( \beta_0 = -1 \) | \( n = 20 \) | \( n = 30 \) | \( n = 40 \) | \( n = 100 \) |
|---------|-------------|-------------|-------------|-------------|-------------|----------------|-------------|-------------|-------------|-------------|
| MLE     | 189540.800  | 52332.180   | 30131.600   | 21166.800   | 7627.807    | 159604.500    | 45226.060   | 26292.020   | 18484.280   | 6699.903    |
| RE \( k_1 \) | 1.436       | 1.169       | 1.012       | 0.910       | 0.658       | 1.826          | 1.738       | 1.684       | 1.650       | 1.564       |
| RE \( k_2 \) | 1.459       | 1.151       | 0.953       | 0.818       | 0.455       | 1.347          | 1.037       | 0.852       | 0.727       | 0.397       |
| RE \( k_3 \) | 1.970       | 1.960       | 1.952       | 1.945       | 1.914       | 1.992          | 1.989       | 1.987       | 1.975       | 1.977       |
| LE \( D_L \) | 59285.760   | 22644.310   | 47231.380   | 10972.310   | 4374.427    | 1.346          | 1.036       | 0.851       | 0.726       | 0.396       |
| LE \( D_L \) | 1.346       | 1.044       | 0.867       | 0.749       | 0.432       | 1.346          | 1.036       | 0.851       | 0.726       | 0.396       |
| LE \( D_L \) | 485.473     | 157.635     | 87.075      | 51.146      | 6.811       | 1.346          | 1.036       | 0.851       | 0.726       | 0.396       |
| LE \( D_L \) | 22132.520   | 12130.250   | 9296.824    | 7661.909    | 3746.868    | 1.346          | 1.036       | 0.851       | 0.726       | 0.396       |
| LE \( D_L \) | 1.346       | 1.044       | 0.867       | 0.749       | 0.432       | 1.346          | 1.036       | 0.851       | 0.726       | 0.396       |

Note: RE \( a \) indicates the ridge estimator based on the ridge parameter estimator \( a \), LE \( b \) indicates the Liu estimator based on the Liu parameter estimator \( b \), and bold values indicate the minimum SMSEs.

| Table 10: Estimated SMSE values for different \( n, \rho, k, d, \) and \( \beta_0 \) when \( \delta = 0.04, \alpha = 0.25, \) and \( p = 4 \). |
|----------------------|
| \( \rho = 0.99 \) |
| Estimator          | \( n = 10 \) | \( n = 20 \) | \( n = 30 \) | \( n = 40 \) | \( n = 100 \) | \( \beta_0 = -1 \) | \( n = 20 \) | \( n = 30 \) | \( n = 40 \) | \( n = 100 \) |
|---------|-------------|-------------|-------------|-------------|-------------|----------------|-------------|-------------|-------------|-------------|
| MLE     | 1203.887    | 326.347     | 186.925     | 131.032     | 47.070      | 942.456        | 266.410     | 154.391     | 108.775     | 39.413      |
| RE \( k_1 \) | 1.247       | 0.970       | 0.825       | 0.738       | 0.541       | 1.759          | 1.663       | 1.612       | 1.582       | 1.512       |
| RE \( k_2 \) | 1.251       | 0.915       | 0.728       | 0.644       | 0.490       | 1.152          | 0.849       | 0.698       | 0.611       | 0.467       |
| RE \( k_3 \) | 1.645       | 1.514       | 1.434       | 1.369       | 1.142       | 1.884          | 1.842       | 1.813       | 1.786       | 1.684       |
| LE \( D_L \) | 465.044     | 162.790     | 102.135     | 74.654      | 29.225      | 1.149          | 0.847       | 0.696       | 0.609       | 0.465       |
| LE \( D_L \) | 1.153       | 0.865       | 0.722       | 0.640       | 0.482       | 1.149          | 0.847       | 0.696       | 0.609       | 0.465       |
| LE \( D_L \) | 4.658       | 1.685       | 1.056       | 0.783       | 0.484       | 1.149          | 0.847       | 0.696       | 0.609       | 0.465       |
| LE \( D_L \) | 226.525     | 106.636     | 75.619      | 59.457      | 26.668      | 1.149          | 0.847       | 0.696       | 0.609       | 0.465       |
| LE \( D_L \) | 1.153       | 0.865       | 0.722       | 0.640       | 0.482       | 1.149          | 0.847       | 0.696       | 0.609       | 0.465       |

Note: RE \( a \) indicates the ridge estimator based on the ridge parameter estimator \( a \), LE \( b \) indicates the Liu estimator based on the Liu parameter estimator \( b \), and bold values indicate the minimum SMSEs.
coefficients of covariates, SMSE value, and AIC value than the other regression models. Then, we may say that the PMQL Liu regression model gives a better performance than the other regression models based on the Liu estimator.

5.2. Real Data Application. The applicability of the PMQL Liu estimator over the MLE is illustrated by using the Swedish football data set, which consists of the Swedish football teams’ performance in the top Swedish league (Allsvenskan) for the year 2012. Qasim et al. [32] used a similar data set during the year 2018 to fit a Poisson regression model. It contains 242 observations and represents the number of full-time home team goals (y), the pinnacle home win odds (x1), the pinnacle away win odds (x2), the maximum odds portal home win (x3), the maximum odds portal away win (x4), the average odds portal home win (x5), and the average odds portal away win (x6).

The conditional number, which is the ratio of the maximum to minimum eigenvalues, is 33460.350, which is clearly much larger than 1000. The index of dispersion of y is 1.201, which is greater than one. These results indicate that there exists severe multicollinearity among the covariates, and y is overdispersed. Further, to examine whether the PMQL distribution is suitable for y, the Chi-square (χ2) goodness of fit test is employed. χ2 value is computed as 3.159 with a p-value equal to 0.531. Then, this test confirms that the PMQL distribution fits well for this response variable y.

Table 12 lists the estimated regression coefficients, their standard errors (SEs) (in parentheses), and SMSE values for the MLE, Liu estimator, and PMQL Liu estimator. It can be clearly noted that the PMQL Liu estimator produces smaller SEs and SMSE than the MLE. Then, we can conclude that the PMQL Liu estimator performs better than MLE for this real data set that has severe multicollinearity issues.

Now, we justify Theorem 1 by using the real-world application. The necessary conditions in Theorem 1 are holding asmin(λj(λj + 1)/2 − (λj + d)2)p+1 = 89.004 > 0, the minimum eigenvalue of the difference matrix

\[
\text{MSE} \left( \hat{\beta}_{\text{MLE}} \right) - \text{MSE} \left( \hat{\beta}_{\text{PMQL}} \right) \quad \text{equals} \quad 7.134 \times 10^{-13} > 0
\]

\[
b_d I \left( (\lambda - 1 - (\lambda + I))^{-1} (\lambda + dI) \lambda^{-1} (\lambda + dI) (\lambda + I)^{-1} \right)^{-1} b_d = 0.501 < 1,
\]

and hence Theorem 1 is justified.

6. Conclusion

This paper introduced the Liu estimator for the Poisson-Modification of Quasi Lindley (PMQL) regression model instead of the maximum likelihood estimator (MLE) in order to mitigate the multicollinearity and overdispersion issues. A comprehensive Monte Carlo simulation study was conducted to compare the performance of the MLE, the PMQL Liu regression estimator, and the existing biased estimator, namely, the PMQL ridge regression estimator.
The scalar mean square error (SMSE) was considered as the evaluation criterion. The simulation study results revealed that the performance of the different estimators is affected by the different levels of factors such as correlations between the covariates, sample size, number of covariates, intercept, and the overdispersion parameters of the PMQL regression model. Further, the PMQL Liu regression estimator based on the Liu parameter estimator $D_3$ shows a better performance than the other estimators in all situations reviewed in the simulation study. The results of a simulated data set show that the PMQL Liu regression model performs better than the existing count Liu regression models, namely, the negative binomial Liu regression and the Poison Liu regression models, by mitigating higher overdispersion and multicollinearity. Further, a real-world application also shows that the PMQL Liu regression estimator based on the Liu parameter estimator $D_3$ has a better performance than the MLE. Therefore, based on the simulation study and applications, the PMQL Liu regression estimator based on the Liu parameter estimator $D_3$ is recommended for analyzing the overdispersed count responses with intercorrelated covariates.

Appendix

A. Lemma

Lemma A.1. Let $M$ be a positive definite (pd) matrix and $c$ be a vector of nonzero constants. Then, $M - cc' > 0$ iff $c' M^{-1} c < 1$ [42].

Data Availability

The data set is available for the public at https://www.football-data.co.uk/sweden.php. The data are also available from the author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

[1] C. Cameron and P. K. Trivedi, Regression Analysis of Count Data, Cambridge University Press, Cambridge, UK, 1998.
[2] M. Greenwood and G. U. Yule, “An inquiry into the nature of frequency distributions representative of multiple happenings with particular reference to the occurrence of multiple attacks of disease or of repeated accidents,” Journal of the Royal Statistical Society, vol. 83, no. 2, p. 255, 1920.
[3] M. M. Shoukri, M. H. Asyali, R. VanDorp, and D. Kelton, “The Poisson inverse Gaussian regression model in the analysis of clustered counts data,” Journal of Data Science, vol. 2, no. 1, pp. 17–32, 2021.
[4] G. Tzougas, “EM estimation for the Poisson-inverse gamma regression model with varying dispersion: an application to insurance ratemaking,” Risks, vol. 8, no. 3, p. 97, 2020.
[5] M. Sankaran, “The discrete Poisson-Lindley distribution,” Biometrics, vol. 26, no. 1, p. 145, 1970.
[6] E. Mahmoudi and H. Zakerzadeh, “Generalized Poisson—Lindley distribution,” Communications in Statistics—Theory and Methods, vol. 39, no. 10, pp. 1785–1798, 2010.
[7] W. Wongrin and W. Bodhisuwan, “The Poisson generalised Lindley distribution and its applications,” Songklanakarin Journal of Science and Technology, vol. 38, pp. 645–656, 2016.
[8] R. Grine and H. Zeghdoudi, “On Poisson quasi-Lindley distribution and its applications,” Journal of Modern Applied Statistical Methods, vol. 16, no. 2, pp. 403–417, 2017.
[9] Y. Atikankul, A. Thongteeraparp, and W. Bodhisuwan, “The new Poisson mixed weighted Lindley distribution with applications to insurance claims data,” Songklanakarin Journal of Science and Technology, vol. 42, pp. 1–11, 2020.
[10] J. Lynch, “Mixtures, generalized convexity and balayages,” Scandinavian Journal of Statistics, vol. 15, pp. 203–210, 1988.
[11] R. Tharshan and P. Wijekoon, “Location based generalized Akash distribution: properties and applications,” Open Journal of Statistics, vol. 10, no. 02, pp. 163–187, 2020.
[12] R. Tharshan and P. Wijekoon, “A comparison study on a new five-parameter generalized Lindley distribution with its sub-models,” Statistics in Transition New Series, vol. 21, no. 2, pp. 89–117, 2020.
[13] W. Wongrin and W. Bodhisuwan, “Generalized Poisson-Lindley linear model for count data,” Journal of Applied Statistics, vol. 44, no. 15, pp. 2659–2671, 2016.
[14] E. Altun, “A new model for over-dispersed count data: Poisson quasi-Lindley regression model,” The Mathematical Scientist, vol. 13, no. 3, pp. 241–247, 2019.
[15] R. Tharshan and P. Wijekoon, “A modification of the Quasi Lindley distribution,” Open Journal of Statistics, vol. 11, no. 03, pp. 369–392, 2021.
[16] R. Tharshan and P. Wijekoon, “A new mixed Poisson distribution for over-dispersed count data: theory and applications,” Reliability: Theory and Applications, vol. 17, pp. 33–51, 2022.
[17] R. Tharshan and P. Wijekoon, “Poission-modification of Quasi Lindley regression model for over-dispersed count responses.” Communications in Statistics - Simulation and Computation, pp. 1–16, 2022.
[18] R. Frisch, Statistical Confluence Analysis by Means of Complete Regression Systems, University Institute of Economics, Oslo, Norway, 1934.
[19] R. Tharshan and P. Wijekoon, “Ridge estimator in a mixed Poisson regression model,” Communication in Statistics: Simulation and Computation, pp. 1–18, 2022.
[20] A. E. Hoerl and R. W. Kennard, “Ridge regression: biased estimation for nonorthogonal problems,” Technometrics, vol. 12, no. 1, pp. 55–67, 1970.
[21] B. Segerstedt, “On ordinary ridge in generalized linear models,” Communications in Statistics—Theory and Methods, vol. 21, no. 8, pp. 2227–2246, 1992.
[22] L. Kejian, “A new class of biased estimate in linear regression,” Communications in Statistics—Theory and Methods, vol. 22, no. 2, pp. 393–402, 1993.
[23] K. Månsson, B. M. G. Kibria, P. Sjölander, and G. Shukur, “Improved Liu estimators for the Poisson regression model,” International Journal of Statistics and Probability, vol. 1, p. 2, 2012.
[24] K. Månsson, B. G. Kibria, and G. Shukur, “On Liu estimators for the logit regression model,” Economic Modelling, vol. 29, no. 4, pp. 1483–1488, 2012.
[25] K. Månsson, “Developing a Liu estimator for the negative binomial regression model: method and application,” Journal
of Statistical Computation and Simulation, vol. 83, no. 9, pp. 1773–1780, 2013.

[26] G. Ü. Şiray, S. Toker, and S. Kaçiranlar, “On the restricted Liu estimator in the logistic regression model,” Communications in Statistics—Simulation and Computation, vol. 44, no. 1, pp. 217–232, 2015.

[27] J. Wu, “Modified restricted Liu estimator in logistic regression model,” Computational Statistics, vol. 31, no. 4, pp. 1557–1567, 2016.

[28] F. Kurtoğlu and M. R. Özkale, “Liu estimation in generalized linear models: application on gamma distributed response variable,” Statistical Papers, vol. 57, no. 4, pp. 911–928, 2016.

[29] S. Türkcan and G. Özel, “A Jackknifed estimators for the negative binomial regression model,” Communications in Statistics—Simulation and Computation, vol. 47, no. 6, pp. 1845–1865, 2018.

[30] J. Wu, Y. Asar, and M. Arashi, “On the restricted almost unbiased Liu estimator in the logistic regression model,” Communications in Statistics—Theory and Methods, vol. 47, no. 18, pp. 4389–4401, 2018.

[31] N. Varathan and P. Wijekoon, "Logistic Liu Estimator under stochastic linear restrictions," Statistical Papers, vol. 60, no. 3, pp. 945–962, 2019.

[32] M. Qasim, B. M. G. Kibria, K. Månsson, and P. Sjölander, “A new Poisson Liu Regression Estimator: method and application,” Journal of Applied Statistics, vol. 47, no. 12, pp. 2258–2271, 2020.

[33] Y. Li, Y. Asar, and J. Wu, “On the stochastic restricted Liu estimator in logistic regression model,” Journal of Statistical Computation and Simulation, vol. 90, no. 15, pp. 2766–2788, 2020.

[34] T. Omer, P. Sjölander, K. Månsson, and B. M. G. Kibria, “Improved estimators for the zero-inflated Poisson regression model in the presence of multicollinearity: simulation and application of maternal death data,” Communications in Statistics: Case Studies, Data Analysis and Applications, vol. 7, no. 3, pp. 394–412, 2021.

[35] B. M. G. Kibria, “Performance of some new ridge regression estimators,” Communications in Statistics—Simulation and Computation, vol. 32, no. 2, pp. 419–435, 2003.

[36] G. Khalaf and G. Shukur, “Choosing ridge parameter for regression problems,” Communications in Statistics—Theory and Methods, vol. 34, no. 5, pp. 1177–1182, 2005.

[37] M. Nomura, “On the almost unbiased ridge regression estimator,” Communications in Statistics—Simulation and Computation, vol. 17, no. 3, pp. 729–743, 1988.

[38] G. Muniz and B. M. G. Kibria, “On some ridge regression estimators: an empirical comparisons,” Communications in Statistics—Simulation and Computation, vol. 38, no. 3, pp. 621–630, 2009.

[39] G. C. McDonald and D. I. Galarneau, “A Monte Carlo evaluation of some ridge-type estimators,” Journal of the American Statistical Association, vol. 70, no. 350, pp. 407–416, 1975.

[40] K. Månsson and G. Shukur, “A Poisson ridge regression estimator,” Economic Modelling, vol. 28, no. 4, pp. 1475–1481, 2011.

[41] K. Månsson, “On ridge estimators for the negative binomial regression model,” Economic Modelling, vol. 29, no. 2, pp. 178–184, 2012.

[42] R. W. Farebrother, “Further results on the mean square error of ridge regression,” Journal of the Royal Statistical Society: Series B, vol. 38, no. 3, pp. 248–250, 1976.