JOURNAL OF APPLIED STATISTICS
2020, VOL. 47, NO. 12, 2258–2271
https://doi.org/10.1080/02664763.2019.1707485

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS Check for updates

A new Poisson Liu Regression Estimator: method and application

Muhammad Qasim[a], B. M. G. Kibria [b], Kristofer Månsson[a] and Pär Sjölander[a]

[a]Department of Economics, Finance and Statistics, Jönköping University, Jönköping, Sweden; [b]Department of Mathematics and Statistics, Florida International University, Miami, FL, USA

ABSTRACT
This paper considers the estimation of parameters for the Poisson regression model in the presence of high, but imperfect multicollinearity. To mitigate this problem, we suggest using the Poisson Liu Regression Estimator (PLRE) and propose some new approaches to estimate this shrinkage parameter. The small sample statistical properties of these estimators are systematically scrutinized using Monte Carlo simulations. To evaluate the performance of these estimators, we assess the Mean Square Errors (MSE) and the Mean Absolute Percentage Errors (MAPE). The simulation results clearly illustrate the benefit of the methods of estimating these types of shrinkage parameters in finite samples. Finally, we illustrate the empirical relevance of our newly proposed methods using an empirically relevant application. Thus, in summary, via simulations of empirically relevant parameter values, and by a standard empirical application, it is clearly demonstrated that our technique exhibits more precise estimators, compared to traditional techniques – at least when multicollinearity exist among the regressors.

ARTICLE HISTORY
Received 8 July 2019
Accepted 14 December 2019

KEYWORDS
MLE; MSE; Poisson regression; Liu estimator; shrinkage estimators; simulation study

1. Introduction

The Poisson Regression (PR) model is an appropriate model for studying count variables using appropriate covariates. For instance, the number of patients, bank failures, the number of road accidents, traffic flow and ideal gap distances, number of typing errors on a page, failure of a machine in one month, the occurrences of virus disease, takeover bids and criminal careers can be modeled with the Poisson distribution etc. The common Maximum Likelihood Estimator (MLE) is used to estimate unknown regression coefficients in the PR model. The MLE can be found by applying an Iterative Weighted Least Square (IWLS) algorithm. One problem with MLE occurs when there are linear dependencies among the explanatory variables. This problem is called multicollinearity by Frisch [6]. For example, when counting the number of injuries that occur in the upper seam of mines in the coal fields, then the inside burden thickness, lower seam height, and extraction of the lower earlier mined seam in percentage are the important factors. In such situation, the explanatory variables would be strongly correlated. High (imperfect) multicollinearity

CONTACT Kristofer Månsson kristofer.mansson@ju.se

causes the MLE to overestimate the standard errors, while the standard errors are consistent. It leads difficulty to isolate marginal effects of individual regressors since marginal interpretation implies holding the other independent variables constant.

This problem of multicollinearity has significant impact on the performance of MLE for the estimation of unknown regression coefficients in the PR model. Furthermore, it leads to instability and a high variance of the parameters estimated by MLE. Another consequence of multicollinearity is the wider confidence interval, decreased statistical power which result in increased probability of type II error in hypothesis testing in terms of the parameters. In addition, the uncertainty of the estimated coefficients is higher because of an increased coefficient variance due to multicollinearity. By minimizing the standard errors of the coefficients, we demonstrate that our new Liu estimator is a beneficial and a recommended remedy for the problem of multicollinearity.

In recent research, it is a stylized fact that the shrinkage estimators are considered as an efficient remedial measure to combat multicollinearity problem [11,22]. Many researchers propose different type of shrinkage estimators to overcome multicollinearity for different models. Månsson and Shukur [20] proposed a Poisson ridge regression estimator which was a generalization of the ordinary ridge regression. In 1993, Liu introduced a new estimator, subsequently known as the Liu estimator. It is based on a linear function of d instead of a non-linear function as in the ordinary ridge regression. This leads to a more stable shrinkage of the vector of estimated coefficients. Therefore, due to the linear function of d, researchers have used the more robust Liu estimator instead of the traditional ridge regression. Regarding the vast literature on the Liu estimator in the linear regression model, we refer our readers primarily to Liu [12], Kaciranlar [10], Alheety and Kibria [1], Kibria [14], Qasim *et al.* [22], among others. Furthermore, Arashi *et al.* [4,5] deliberated the improved preliminary test and Stein-rule Liu estimators, and Liu type estimator. Recently, Karbalaee *et al.* [11] introduced a Preliminary test generalized Liu estimator with series of stochastic restrictions. However, the literature on the Liu estimator of a generalized linear model is rather limited. For instance, Månsson *et al.* [19] suggested some shrinkage parameters for the Poisson Liu Regression Estimator (PLRE), Månsson *et al.* [18] introduced a Liu estimator for the logit regression model, Månsson [17] recommended some Liu parameters for the negative binomial regression model, Inan and Erdogan [9] developed a Liu-type logistic estimator, Şiray *et al.* [23] proposed a restricted Liu estimator in the logistic regression model, Amin *et al.* [3] recommend some shrinkage parameters for the gamma regression, Qasim *et al.* [21] developed and adopted some new shrinkage parameters for the Liu estimator for the gamma regression model, Wu *et al.* [25] developed the restricted almost unbiased Liu estimator for the logistic regression model, and, finally, recently, Amin *et al.* [2] proposed Liu type estimators for the gamma regression model.

The main contribution of this paper is to propose some new methods of estimating the shrinkage parameter, d for the PR model. The original methods that inspired our new estimation methods were developed by Hoerl and Kennard [8], Kibria [13], Månsson *et al.* [18]. The Mean Squared Error (MSE) and Mean Absolute Percentage Error (MAPE) are considered as performance criteria for evaluation of the proposed estimators in the Monte Carlo experiment. The intuitive, and empirical relevance, of the Liu estimator is demonstrated by applying proposed estimation methods and traditional MLE on real-world data where we systematically analyze which estimator that can to the highest degree remedy the effects of multicollinearity. In this empirical application, we model the number of goals scored at home and away as a function of the quality of the teams (measured by bookmaker odds). By this approach is easily demonstrated that the standard errors and the estimated MSEs decrease substantially. Hence, the precision of the estimated parameters is increased, which of course is one of the main objectives in an empirical situation

This study is structured as follows: We discuss the model of interest and propose different shrinkage parameters in section 2. The Monte Carlo experiment and the simulated results are addressed in Section 3. An empirical application is outlined in section 4. Finally, the concluding remarks are provided in section 5.

2. A new Poisson Liu Regression Estimator (PLRE)

The Poisson Regression (PR) model is only applicable when the dependent variable deals with count data. Suppose, y_i is the dependent variable and follows a Poisson distribution with parameter (θ_i) and is denoted as P(θ_i) with the following probability mass function

$$f(y_i) = \frac{e^{\theta_i}\theta_i^{y_i}}{y_i!}; y_i = 0, 1, 2, 3, \ldots; i = 1, 2, 3, \ldots, n. \tag{1}$$

The PR model is commonly developed by using the canonical link function, such that $\theta_i = \exp(x_i^t\beta)$, where x_i is the ith row of X which is an $n \times p$ data matrix with p non-stochastic explanatory variables, β is a $p \times 1$ vector of the unknown regression coefficients. The log-likelihood function of the PR model can be defined as

$$l(\theta; y) = \sum_{i=1}^{n}\left\{y_i \ln(\theta_i) - \theta_i - \ln\left(\prod_{i=1}^{n} y_i!\right)\right\}$$

$$l(\theta; y) = \sum_{i=1}^{n}\left\{y_i \ln(\exp(x_i^t\beta)) - \exp(x_i^t\beta) - \ln\left(\prod_{i=1}^{n} y_i!\right)\right\}. \tag{2}$$

The MLE is used to estimate the parameters of the model. The following IWLS algorithm is applying to maximize the log-likelihood function.

$$\hat{\beta}_{MLE} = (X^t\hat{V}X)^{-1}X^t\hat{V}z^*, \tag{3}$$

where $\hat{V} = \text{diag}\{\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_n\}$; $z^* = x_i^t\hat{\beta}_{MLE} + \frac{y_i - \theta_i}{\hat{\theta}_i}$ is the adjusted response variable. Both $\hat{V}$and $z^*$are evaluated by Fisher's scoring iterative procedure. The MSEs of the estimators are obtained by considering $\alpha = \Upsilon^t\beta$ and $\Lambda = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_p) = \Upsilon(X^t\hat{V}X)\Upsilon^t$, where Υ is the orthogonal matrix whose columns are the eigenvectors of $X^t\hat{V}X$; and $\lambda_1 \geq \lambda_2 \geq, \ldots, \geq \lambda_p > 0$ are the eigenvalues of the matrix $X^t\hat{V}X$ and $\alpha_j (j = 1, 2, \ldots, p)$ of the jth element of $\Upsilon^t\beta$. Furthermore, the matrix MSE (MMSE) of the $\hat{\beta}_{MLE}$ is defined as

$$\text{MMSE}(\hat{\beta}_{MLE}) = (\Upsilon\Lambda^{-1}\Upsilon^t). \tag{4}$$

Moreover, the scalar MSE of the $\hat{\beta}_{MLE}$ is defined as

$$\text{MSE}(\hat{\beta}_{MLE}) = E(\hat{\beta}_{MLE} - \beta)^t(\hat{\beta}_{MLE} - \beta) = tr\{\Upsilon\Lambda^{-1}\Upsilon^t\} = \sum_{j=1}^{p}\frac{1}{\lambda_j}, \tag{5}$$

where λ_j is the jth eigenvalue of the $X^t\hat{V}X$ matrix. When the explanatory variables are linearly correlated, some of the eigenvalues will be small and $X^t\hat{V}X$ matrix will be ill-conditioned which inflated the variance of MLE. To overcome this problem of multicollinearity, we define a PLRE which is generalization of Liu [12].

$$\hat{\beta}_{PLRE} = (X^t\hat{V}X + I)^{-1}(X^t\hat{V}X + dI)\hat{\beta}_{MLE} \tag{6}$$

where d ($0 \le d \le 1$) is the shrinkage parameter. If $d = 1$ then $\hat{\beta}_{MLE} = \hat{\beta}_{PLRE}$ and in case $d < 1$ which implies that the absolute norm vector of PLRE is less than or equal to the absolute norm vector of MLE, i.e. $\hat{\beta}_{PLRE} \le \hat{\beta}_{MLE}$. The MMSE and MSE of the PLRE can be defined as

$$Bias(\hat{\beta}_{PLRE}) = \Upsilon {\Lambda_I}^{-1}\alpha(d-1) \tag{7}$$

$$VAR(\hat{\beta}_{PLRE}) = \Upsilon {\Lambda_I}^{-1}\Lambda_d\Lambda{\Lambda_I}^{-1}\Lambda_d\Upsilon^t \tag{8}$$

$$\text{MMSE}(\hat{\beta}_{PLRE}) = \Upsilon {\Lambda_I}^{-1}\Lambda_d\Lambda{\Lambda_I}^{-1}\Lambda_d\Upsilon^t + (d-1)^2\Upsilon{\Lambda_I}^{-1}\alpha\alpha^t{\Lambda_I}^{-1}\Upsilon^t \tag{9}$$

where $\Lambda_I = \text{diag}(\lambda_1 + I, \lambda_2 + I, \ldots, \lambda_p + I)$, $\Lambda_d = \text{diag}(\lambda_1 + d, \lambda_2 + d, \ldots, \lambda_p + d)$ and $\Lambda = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_p) = \Upsilon(X^t\hat{V}X)\Upsilon^t$, where Υ is the orthogonal matrix whose columns are the eigenvectors of $X^t\hat{V}X$. Finally, the scalar MSE of the PLRE is obtained by applying tr (.) operator on Equation (9), which can be defined as

$$\text{MSE}(\hat{\beta}_{PLRE}) = \sum_{j=1}^{p}\left(\frac{\lambda_j + d}{(\lambda_j+1)^2\lambda_j}\right) + (d-1)^2\sum_{j=1}^{p}\left(\frac{\alpha_j^2}{(\lambda_j+1)^2}\right). \tag{10}$$

Liu [12] provided a proof that the Liu estimator is better than the ordinary least squares estimator for the linear regression model. We extend this method for PR model and show that the PLRE perform better than the MLE. In order to do so, we follow Liu [12] and differentiate Equation (10) with respect to *d:*

$$g'(d) = \frac{\partial[MSE(\hat{\beta}_{PLRE})]}{\partial d} = 2\sum_{j=1}^{p}\frac{\lambda_j + d}{\lambda_j(\lambda_j+1)^2} + 2(d-1)\sum_{j=1}^{p}\frac{\alpha_j^2}{(\lambda_j+1)^2}. \tag{11}$$

Thus, by inserting the value 1 (the situation when PLRE and MLE are equal) we can see that:

$$g'(1) = 2\sum_{j=1}^{p}\frac{\lambda_j + d}{\lambda_j(\lambda_j+1)^2} > 0 \tag{12}$$

Therefore there exists $0 < d < 1$ such that $g(d) < g(1)$ or, equivalently, $\text{MSE}(\hat{\beta}_{\text{PLRE}}) < \text{MSE}(\hat{\beta}_{\text{MLE}})$.

2.1. Proposed shrinkage estimators

The PLRE is a robust measure and it performs better than the usual MLE when the explanatory variables are linearly correlated. Furthermore, the performance of the PLRE depends on the optimal value of the shrinkage parameter, d. Therefore, we propose some new estimators in order to obtain the value of d based on the work of Månsson *et al.* [18]. In order to obtain the optimal value of the PLRE, we take the first derivative of Equation (10) and then solve it for d by equating to zero. Consequently, we obtain

$$d_j = \frac{\alpha_j^2 - 1}{\frac{1}{\lambda_j} + \alpha_j^2} \tag{13}$$

The range of d depends on α_j^2. However, as is specified in [12], the value of d is limited between 0 and 1. Therefore, we use max operator with the proposed estimators to ensure the value of d_j lie between 0 to 1. The ideas of the proposed estimators are based on the theoretical work of Hoerl and Kennard [8], Kibria [13], and Månsson *et al.* [18]. For estimation of an optimal value of d, we define D_1, D_2-D_3, and D_4-D_5 estimators based on the work of Hoerl and Kennard [8], Kibria [13], and Månsson *et al.* [18], respectively.

$$D_1 = \max\left(0, \frac{\hat{\alpha}_{max}^2 - 1}{\frac{1}{\hat{\lambda}_{max}} + \hat{\alpha}_{max}^2}\right) \quad D_2 = \max\left(0, \text{median}\left(\frac{\hat{\alpha}_j^2 - 1}{\frac{1}{\hat{\lambda}_j} + \hat{\alpha}_j^2}\right)\right)$$

$$D_3 = \max\left(0, \sum_{j=1}^{p}\left(\frac{\hat{\alpha}_j^2 - 1}{\frac{1}{\hat{\lambda}_j} + \hat{\alpha}_j^2}\right)/p\right) \quad D_4 = \max\left(0, \max\left(\frac{\hat{\alpha}_j^2 - 1}{\frac{1}{\hat{\lambda}_j} + \hat{\alpha}_j^2}\right)\right)$$

$$D_5 = \max\left(0, \min\left(\frac{\hat{\alpha}_j^2 - 1}{\frac{1}{\hat{\lambda}_j} + \hat{\alpha}_j^2}\right)\right)$$

where $\hat{\alpha}_{max}^2$ and $\hat{\lambda}_{max}$ which defined in D_1 estimator are the maximum element of the $\hat{\alpha}_j^2$ and $\Upsilon(X^t\hat{V}X)\Upsilon^t$, respectively.

A contribution of this paper is the following PLRE estimators.

$$D_{kp1} = \max(0, \text{median}(m_j)), \quad D_{kp2} = \max\left(0, \sum_{j=1}^{p}(m_j)/p\right)$$

$$D_{kp3} = \max(0, \max(m_j)), \quad D_{kp4} = \max(0, \min(m_j))$$

$$D_{q1} = \max(0, \text{median}(h_j)), \quad D_{q2} = \max\left(0, \sum_{j=1}^{p}(h_j)/p\right)$$

$$D_{q3} = \max(0, \max(h_j)/p), \quad D_{q4} = \max(0, \min(h_j))$$

where $m_j = \frac{\hat{\alpha}_j^2 - 1}{\max\left(\frac{1}{\hat{\lambda}_j}\right) + \hat{\alpha}_j^2}$ and $h_j = \frac{\hat{\alpha}_j^2 - 1}{\max\left(\frac{1}{\hat{\lambda}_j}\right) + \hat{\alpha}_{\max}^2}$. It should be noted that the PLRE perform much better when the optimal value of d is close to zero. Hence our proposed estimator's value is always close to zero.

2.2. Performance evaluation criteria of the estimators

The main objective of this article is to compare the performance of the proposed estimators with the existing estimators in order to improve the performance of the PLRE in the presence of high, but imperfect, multicollinearity. So, there is a need to define the performance criteria for selecting the best estimator. In this study, we consider the standard measure of MSE and MAPE as the performance criteria. These are defined as follows:

$$MSE = \frac{\sum_{r=1}^{R} (\hat{\beta}_{(r)} - \beta)^{t} (\hat{\beta}_{(r)} - \beta)}{R} \quad (14)$$

$$\text{MAPE} = \left(\frac{1}{R} \left(\sum_{r=1}^{R} \left(\frac{\sum_{j=1}^{P} \left| \frac{(\hat{\beta}-\beta)_j}{(\beta)_j} \right|}{p} \right) \right) \right) 100\%, \quad (15)$$

where $\hat{\beta}$ is the estimator of β at the ith repetition out of $R = 2000$ replicates.

3. The simulation study

In this section, we conduct a Monte Carlo simulation study to evaluate the performance of our newly proposed estimators and the MLE. Below we provide a brief discussion about the simulation results.

3.1. The design of an experiment

The dependent variable of the PR model is generated from the Poisson distribution $P_o(\theta_i)$, where

$$\theta_i = \exp(\beta_o + \beta_1 x_{i1} +, \ldots, +\beta_p x_{ip}); j = 1, 2, \ldots, p; i = 1, 2, \ldots, n. \quad (16)$$

where Equation (16) is the mean function and it is generated for $p = 4, 8$ regressors, respectively. Furthermore, the intercept value is set to be –1 or 1, and the slope parameter values of Equation (16) are chosen to be $\sum_{j=1}^{p} \beta_j^2 = 1$ and $\beta_1 =, \ldots, = \beta_p$ by considering different sample sizes. Table 1 shows the different factors that affect on the performance of the estimators. We use the R 3.2.2 software to conduct the simulation study.

Since the performance of the estimators greatly depends on the strength of the correlation, the following formula is used to generate the correlated explanatory variables

Table 1. Values of the factors that are used in the design of the experiment.

Factors	Notations	Values
Intercept	β_o	−1, 1
Number of explanatory variables	p	4, 8
Degree of correlation	ρ^2	0.90, 0.98, 0.99, 0.999
Sample size	n	25, 50, 75, 100, 125
Replicates	R	2000

[7,13]:

$$x_{ij} = \sqrt{1-\rho^2} z_{ij} + \rho z_{ip} + 1; j = 1, 2, \ldots, p+1; i = 1, 2, \ldots, n. \quad (17)$$

where z_{ij} are the pseudo-random numbers which are generated from the standard normal distribution. We consider four different values of $\rho^2 = 0.90, 0.98, 0.99, 0.999$.

3.2. Simulation results and discussion

The simulated results for MSE and MAPE are presented in Tables 2–5 for $p = 4$and8, respectively. All estimators performed better than the traditional MLE in every case. It is observed that the factors influencing the performance of different estimators are the value

Table 2. Estimated MSE values when $p = 4$.

n	ML	D_1	D_2	D_3	D_4	D_5	D_{kp1}	D_{kp2}	D_{kp3}	D_{kp4}	D_{q1}	D_{q2}	D_{q3}	D_{q4}
							$\beta_o = -1$							
$\rho^2 = 0.90$														
25	2.051	1.509	0.737	0.677	1.429	0.596	0.652	0.658	1.150	0.596	0.624	0.672	0.674	0.595
50	0.960	0.748	0.520	0.508	0.729	0.506	0.517	0.508	0.712	0.506	0.513	0.520	0.551	0.506
75	0.655	0.537	0.436	0.432	0.532	0.432	0.436	0.432	0.528	0.432	0.434	0.436	0.454	0.432
100	0.514	0.437	0.377	0.375	0.435	0.375	0.377	0.375	0.434	0.375	0.376	0.376	0.389	0.375
$\rho^2 = 0.98$														
25	8.961	7.296	2.796	2.428	6.390	0.895	1.430	1.645	3.782	0.741	1.062	1.445	1.098	0.724
50	3.502	2.650	0.943	0.837	2.384	0.750	0.869	0.858	2.038	0.747	0.809	0.947	0.985	0.745
75	2.082	1.617	0.846	0.809	1.541	0.798	0.835	0.812	1.446	0.798	0.818	0.869	0.936	0.797
100	1.457	1.139	0.727	0.702	1.104	0.701	0.722	0.703	1.057	0.701	0.715	0.729	0.780	0.701
$\rho^2 = 0.99$														
25	81.850	78.381	32.987	35.871	70.411	5.304	8.743	11.915	27.129	0.689	4.158	7.958	2.433	0.464
50	35.344	32.479	10.353	10.789	28.638	1.229	3.741	5.640	14.772	0.592	1.847	3.859	1.794	0.445
75	17.582	15.226	4.339	4.479	13.122	0.731	2.482	3.170	7.839	0.605	1.417	2.464	1.411	0.538
100	12.554	10.557	3.199	2.755	9.122	0.788	2.028	2.336	5.813	0.724	1.343	2.008	1.380	0.692
$\rho^2 = 0.999$														
25	819.622	815.484	404.997	522.065	794.182	74.259	75.347	114.554	256.690	4.425	34.536	74.044	16.707	1.273
50	314.751	310.796	116.421	155.670	301.592	12.357	25.246	45.196	110.561	2.332	9.485	27.791	7.588	0.731
75	176.454	172.637	58.213	80.207	164.807	6.463	18.359	29.765	67.385	2.110	7.168	17.903	4.999	0.655
100	120.202	116.572	43.572	55.769	109.314	4.587	12.567	19.374	45.418	1.263	5.087	11.796	3.676	0.482
							$\beta_o = 1$							
$\rho^2 = 0.90$														
25	0.230	0.209	0.204	0.204	0.209	0.204	0.204	0.204	0.208	0.204	0.204	0.204	0.205	0.204
50	0.188	0.185	0.185	0.185	0.185	0.185	0.185	0.185	0.185	0.185	0.185	0.185	0.185	0.185
75	0.216	0.216	0.216	0.216	0.216	0.216	0.216	0.216	0.216	0.216	0.216	0.216	0.216	0.216
100	0.210	0.211	0.211	0.211	0.211	0.211	0.211	0.211	0.211	0.211	0.211	0.211	0.211	0.211
$\rho^2 = 0.98$														
25	0.931	0.671	0.539	0.534	0.657	0.534	0.538	0.534	0.641	0.534	0.537	0.537	0.559	0.534
50	0.521	0.433	0.415	0.415	0.431	0.415	0.415	0.415	0.430	0.415	0.415	0.415	0.418	0.415
75	0.406	0.375	0.372	0.372	0.375	0.372	0.372	0.372	0.375	0.372	0.372	0.372	0.373	0.372
100	0.356	0.338	0.337	0.337	0.338	0.337	0.337	0.337	0.338	0.337	0.337	0.337	0.338	0.337
$\rho^2 = 0.99$														
25	8.108	6.174	2.027	1.368	4.854	0.692	1.286	1.418	3.285	0.678	1.001	1.338	1.068	0.666
50	4.113	2.838	1.120	0.902	2.363	0.800	1.004	0.966	2.007	0.797	0.916	1.034	1.027	0.796
75	2.474	1.621	0.951	0.889	1.510	0.882	0.933	0.896	1.394	0.882	0.913	0.942	0.993	0.882
100	1.817	1.208	0.820	0.805	1.138	0.801	0.815	0.805	1.098	0.801	0.810	0.823	0.867	0.801
$\rho^2 = 0.999$														
25	84.81	81.13	32.01	37.34	74.43	3.378	8.964	12.91	28.92	0.894	4.046	8.321	2.544	0.410
50	40.47	37.12	11.91	13.89	32.80	1.792	4.297	6.500	14.71	0.737	2.042	4.382	1.711	0.443
75	23.14	20.11	5.67	6.11	17.69	0.838	2.893	3.945	9.284	0.627	1.606	2.970	1.491	0.525
100	16.39	13.77	4.54	3.95	11.15	0.801	2.109	2.808	6.861	0.723	1.293	2.270	1.418	0.677

Table 3. Estimated MSE values when $p = 8$.

n	ML	D_1	D_2	D_3	D_4	D_5	D_{kp1}	D_{kp2}	D_{kp3}	D_{kp4}	D_{q1}	D_{q2}	D_{q3}	D_{q4}
							$\beta_o = -1$							
$\rho^2 = 0.90$														
25	4.055	2.827	1.142	1.075	2.625	1.071	1.102	1.102	2.110	1.071	1.086	1.134	1.161	1.071
50	1.795	1.308	0.981	0.980	1.278	0.980	0.981	0.980	1.241	0.980	0.981	0.982	1.010	0.980
75	1.287	0.998	0.861	0.861	0.989	0.861	0.861	0.861	0.980	0.861	0.861	0.861	0.875	0.861
100	0.836	0.674	0.635	0.635	0.672	0.635	0.635	0.635	0.669	0.635	0.635	0.635	0.639	0.635
$\rho^2 = 0.98$														
25	17.323	14.353	3.590	2.331	12.589	1.383	1.865	2.181	6.587	1.341	1.561	1.993	1.648	1.337
50	7.473	5.689	1.874	1.700	5.088	1.614	1.757	1.746	4.193	1.611	1.671	1.842	1.842	1.610
75	4.804	3.566	1.618	1.586	3.309	1.586	1.609	1.595	2.994	1.586	1.596	1.652	1.729	1.586
100	3.167	2.287	1.370	1.366	2.165	1.366	1.369	1.366	2.049	1.366	1.368	1.376	1.440	1.366
$\rho^2 = 0.99$														
25	176.3	171.3	60.17	53.153	156.7	2.041	7.039	11.362	53.19	0.650	2.812	6.894	1.892	0.580
50	66.48	62.03	12.89	12.017	54.07	0.997	4.748	6.619	26.94	0.842	2.233	4.429	1.819	0.798
75	43.00	38.84	8.096	7.011	33.25	1.139	3.912	5.107	18.88	1.074	2.172	3.738	1.909	1.053
100	29.82	26.06	5.372	4.072	21.76	1.240	3.045	3.790	13.46	1.229	1.947	3.082	1.922	1.219
$\rho^2 = 0.999$														
25	1673.9	1668.4	668.1	670.1	1629.2	20.62	53.61	89.87	443.58	0.85	18.95	52.11	7.648	0.328
50	680.8	675.4	168.0	195.4	650.7	4.16	38.22	58.81	258.64	1.05	10.94	30.68	5.057	0.391
75	421.9	416.5	103.3	116.6	397.9	2.78	28.78	41.31	167.34	0.85	8.85	21.69	3.732	0.420
100	304.2	299.0	73.9	79.7	279.2	2.50	20.31	29.69	122.81	0.76	6.06	15.63	3.056	0.395
							$\beta_o = 1$							
$\rho^2 = 0.90$														
25	0.530	0.454	0.451	0.451	0.453	0.451	0.451	0.451	0.453	0.451	0.451	0.451	0.451	0.451
50	0.361	0.351	0.351	0.351	0.351	0.351	0.351	0.351	0.351	0.351	0.351	0.351	0.351	0.351
75	0.318	0.315	0.315	0.315	0.315	0.315	0.315	0.315	0.315	0.315	0.315	0.315	0.315	0.315
100	0.184	0.182	0.182	0.182	0.182	0.182	0.182	0.182	0.182	0.182	0.182	0.182	0.182	0.182
$\rho^2 = 0.98$														
25	2.094	1.370	1.026	1.024	1.333	1.024	1.025	1.025	1.257	1.024	1.025	1.028	1.049	1.024
50	1.049	0.840	0.829	0.829	0.840	0.829	0.829	0.829	0.839	0.829	0.829	0.829	0.830	0.829
75	0.790	0.691	0.689	0.689	0.691	0.689	0.689	0.689	0.690	0.689	0.689	0.689	0.689	0.689
100	0.499	0.448	0.447	0.447	0.448	0.447	0.447	0.447	0.448	0.447	0.447	0.447	0.447	0.447
$\rho^2 = 0.99$														
25	20.397	16.949	4.216	2.715	14.459	1.570	2.207	2.640	7.441	1.549	1.841	2.390	1.908	1.547
50	8.965	6.545	2.083	1.883	5.728	1.804	1.967	1.954	4.562	1.804	1.882	2.062	2.055	1.804
75	5.813	3.939	1.797	1.749	3.598	1.747	1.781	1.758	3.157	1.747	1.764	1.812	1.891	1.747
100	4.100	2.709	1.564	1.554	2.537	1.554	1.561	1.554	2.339	1.554	1.558	1.569	1.638	1.554
$\rho^2 = 0.999$														
25	199.81	194.62	63.80	56.93	178.488	1.955	7.518	12.129	55.66	0.596	3.149	7.524	1.949	0.538
50	86.97	82.09	17.52	17.39	73.017	1.029	6.341	8.778	34.64	0.840	2.777	5.605	2.008	0.786
75	56.32	51.64	10.791	9.875	45.377	1.066	4.996	6.550	23.99	1.024	2.469	4.534	2.005	1.001
100	39.62	35.35	7.428	5.947	29.603	1.173	3.598	4.710	17.10	1.113	2.094	3.551	1.924	1.099

of the intercept, a number of explanatory variables, the degree of correlation and the sample size. It is clearly noticed that as the intercept value increase from –1 to 1, the estimated MSE decrease, and the MAPE increase. Moreover, when the number of explanatory variables and the degree of the generated correlation increases, the estimated MSE and MAPE also increases. Furthermore, the estimated MSE and MAPE decrease with the increase in the sample size. However, the estimated MSE of D_{kp4} and D_{q4} are increase with the increase in sample size due to the minimum value of the m_j and h_j. Although, the performance of D_{kp4} and D_{q4} is better than the other estimation methods of the shrinkage parameter d in PLRE and MLE. Overall, the simulated MSE of these estimation methods increases with the increase in n and β_o while it decreases more when ρ^2 increase. These results are supported by the simulated results of Månsson and Shukur [20], Kibria *et al.* [15] and Kibria

Table 4. Estimated MAPE values when $p = 4$.

n	ML	D_1	D_2	D_3	D_4	D_5	D_{kp1}	D_{kp2}	D_{kp3}	D_{kp4}	D_{q1}	D_{q2}	D_{q3}	D_{q4}
							$\boldsymbol{\beta_o} = -1$							
$\rho^2 = 0.90$														
50	85.38	71.84	53.03	51.64	70.19	50.08	51.57	51.77	65.14	50.07	50.87	52.52	53.16	50.04
75	61.37	53.99	46.40	46.08	53.43	46.04	46.34	46.07	52.90	46.04	46.23	46.41	47.72	46.04
100	48.57	44.24	40.86	40.75	44.09	40.75	40.85	40.75	43.96	40.75	40.81	40.85	41.54	40.75
125	41.64	38.65	36.40	36.37	38.56	36.37	36.40	36.37	38.51	36.37	36.39	36.40	36.90	36.37
$\rho^2 = 0.98$														
50	179.77	154.22	84.55	76.26	141.34	58.34	69.96	74.00	111.25	56.15	64.22	72.95	67.63	55.72
75	121.05	101.71	61.14	58.68	96.70	56.85	59.69	59.34	88.69	56.79	58.44	62.21	64.54	56.75
100	93.80	80.76	58.85	57.93	78.78	57.69	58.61	58.00	76.17	57.68	58.22	59.56	62.23	57.67
125	78.98	68.45	54.81	54.01	67.34	53.98	54.66	54.03	65.80	53.98	54.44	54.78	56.91	53.98
$\rho^2 = 0.99$														
50	540.65	520.44	269.04	272.71	477.37	81.22	135.02	168.96	267.24	46.77	98.97	145.15	90.14	41.83
75	377.43	354.74	161.03	164.43	328.53	53.68	98.96	128.49	215.63	44.88	74.32	112.43	81.67	41.58
100	271.37	246.34	108.66	109.68	223.93	49.35	84.54	99.73	165.37	47.32	68.21	92.32	74.52	45.98
125	231.34	206.57	96.69	90.19	189.77	55.24	80.96	89.11	145.40	54.20	69.85	86.25	75.15	53.51
$\rho^2 = 0.999$														
50	1727.27	1719.49	995.37	1198.74	1685.49	282.46	375.54	511.50	814.05	76.38	249.02	423.50	214.12	51.09
75	1114.53	1103.46	551.06	686.35	1080.52	126.73	234.21	342.00	554.17	57.83	149.36	281.42	150.46	39.20
100	849.03	835.59	399.66	488.00	806.45	93.83	213.43	291.21	456.51	54.47	137.52	237.23	129.25	36.73
125	701.86	685.87	346.80	404.24	656.92	77.73	171.85	230.20	369.67	47.85	114.80	190.13	111.31	36.62
							$\boldsymbol{\beta_o} = 1$							
$\rho^2 = 0.90$														
50	30.43	28.96	28.73	28.73	28.95	28.73	28.73	28.73	28.94	28.73	28.73	28.73	28.78	28.73
75	25.49	25.05	25.03	25.03	25.05	25.03	25.03	25.03	25.05	25.03	25.03	25.03	25.03	25.03
100	23.17	22.96	22.96	22.96	22.96	22.96	22.96	22.96	22.96	22.96	22.96	22.96	22.96	22.96
125	20.83	20.72	20.71	20.71	20.72	20.71	20.71	20.71	20.72	20.71	20.71	20.71	20.71	20.71
$\rho^2 = 0.98$														
50	61.46	51.90	47.79	47.68	51.48	47.68	47.77	47.68	50.99	47.68	47.74	47.74	48.50	47.68
75	47.10	42.64	41.87	41.87	42.57	41.87	41.87	41.87	42.54	41.87	41.87	41.87	42.04	41.87
100	38.96	36.84	36.65	36.65	36.83	36.65	36.65	36.65	36.83	36.65	36.65	36.65	36.69	36.65
125	34.91	33.44	33.36	33.36	33.44	33.36	33.36	33.36	33.44	33.36	33.36	33.36	33.38	33.36
$\rho^2 = 0.99$														
50	180.70	150.36	78.82	66.41	129.89	53.97	67.11	70.31	106.33	53.66	62.03	70.31	66.11	53.43
75	132.04	104.58	66.03	61.42	95.08	59.45	63.86	63.04	87.87	59.40	62.22	65.15	66.39	59.37
100	104.48	80.83	63.39	62.07	78.05	61.93	63.03	62.21	75.24	61.93	62.63	63.25	65.23	61.93
125	89.98	70.10	58.96	58.64	68.20	58.57	58.86	58.66	67.10	58.57	58.76	59.07	60.68	58.57
$\rho^2 = 0.999$														
50	578.14	558.57	289.53	308.64	520.64	70.53	142.31	182.22	285.75	45.73	100.50	153.89	91.82	37.42
75	406.13	381.23	178.60	187.89	346.65	57.16	105.27	135.63	212.45	45.10	78.15	117.66	78.96	40.08
100	312.94	283.44	125.93	129.22	258.38	50.75	91.98	112.15	177.40	47.40	72.78	101.84	76.58	45.26
125	265.19	234.33	114.88	105.91	204.27	54.84	82.42	96.76	154.45	53.51	69.04	90.85	76.05	52.55

et al. [16], where the effects of the multicollinearity problem on logistic and PR models are assessed.

It is clear from Tables 234–5, that as the degree of multicollinearity increases, the simulated MSE values are inflated. This increase is especially strong for the MLE. In such conditions, simulated MSEs of the PLREs by using proposed estimators are clearly smaller than the MLE. In our evaluation we found that the performance of all analyzed shrinkage estimators is approximately the same when n is large and ρ^2 is small. However, as multicollinearity levels increase ($\rho^2 \geq 0.90$), the performance of the proposed estimators is very good as compared to the existing estimators. Our proposed Poisson Liu estimators always exhibit a minimum MSE in the presence of multicollinearity – regardless of the value of the other factors (such as number of explanatory variables, value of intercept and size of the sample). We can also see that the simulated MSE of the estimators increases when the value

Table 5. Estimated MAPE values when $p = 8$.

n	ML	D_1	D_2	D_3	D_4	D_5	D_{kp1}	D_{kp2}	D_{kp3}	D_{kp4}	D_{q1}	D_{q2}	D_{q3}	D_{q4}
							$\boldsymbol{\beta_o} = -1$							
$\rho^2 = 0.90$														
50	188.43	152.34	103.78	101.94	147.19	101.84	102.83	102.84	133.96	101.81	102.35	103.97	105.51	101.81
75	123.92	105.25	93.61	93.60	104.24	93.60	93.61	93.60	103.00	93.60	93.60	93.65	94.76	93.60
100	103.53	90.97	85.61	85.61	90.62	85.61	85.61	85.61	90.26	85.61	85.61	85.61	86.19	85.61
125	85.26	76.63	74.84	74.84	76.53	74.84	74.84	74.84	76.43	74.84	74.84	74.84	75.04	74.84
$\rho^2 = 0.98$														
50	397.97	351.67	159.92	135.27	325.80	115.93	130.42	140.39	231.08	115.47	123.27	137.44	127.74	115.38
75	267.45	227.58	131.72	127.32	214.03	125.54	129.28	128.97	194.18	125.47	127.28	132.35	133.72	125.46
100	216.02	181.61	123.92	123.04	174.60	123.03	123.67	123.28	166.09	123.03	123.34	125.04	128.29	123.03
125	175.78	146.42	115.81	115.70	142.51	115.70	115.78	115.70	138.85	115.70	115.74	115.99	118.55	115.70
$\rho^2 = 0.99$														
50	1260.45	1235.00	597.24	554.96	1164.47	95.99	200.81	270.03	588.47	75.83	143.52	225.11	128.75	73.45
75	804.42	770.67	298.60	286.68	706.31	89.75	185.13	230.99	475.83	86.84	136.25	196.40	130.95	85.57
100	655.84	616.92	241.51	221.90	560.38	100.97	174.17	207.03	409.67	99.53	136.74	183.10	134.68	98.89
125	542.91	501.06	199.61	175.27	449.43	110.18	157.53	180.77	345.43	109.90	133.31	168.23	137.32	109.61
$\rho^2 = 0.999$														
50	3901.20	3892.28	2053.39	2176.33	3833.15	196.57	494.04	718.21	1687.49	54.12	310.53	574.47	229.37	41.99
75	2587.47	2574.90	1102.75	1242.34	2514.18	104.89	483.00	664.07	1458.13	59.86	267.34	498.93	208.83	44.64
100	2051.78	2035.64	874.83	957.50	1978.05	92.82	431.58	565.39	1188.51	63.55	244.44	424.66	181.31	51.34
125	1732.72	1714.59	728.35	771.21	1644.86	88.45	358.70	476.17	1012.71	65.10	206.35	361.08	165.97	55.30
							$\boldsymbol{\beta_o} = 1$							
$\rho^2 = 0.90$														
50	67.00	61.33	61.18	61.18	61.32	61.18	61.18	61.18	61.31	61.18	61.18	61.18	61.20	61.18
75	47.69	46.39	46.39	46.39	46.39	46.39	46.39	46.39	46.39	46.39	46.39	46.39	46.39	46.39
100	40.27	39.59	39.59	39.59	39.59	39.59	39.59	39.59	39.59	39.59	39.59	39.59	39.59	39.59
125	32.78	32.39	32.39	32.39	32.39	32.39	32.39	32.39	32.39	32.39	32.39	32.39	32.39	32.39
$\rho^2 = 0.98$														
50	140.14	110.97	99.35	99.31	109.71	99.31	99.33	99.32	107.26	99.31	99.32	99.39	100.28	99.31
75	96.74	85.20	84.73	84.73	85.17	84.73	84.73	84.73	85.13	84.73	84.73	84.73	84.78	84.73
100	81.26	74.52	74.45	74.45	74.52	74.45	74.45	74.45	74.51	74.45	74.45	74.45	74.46	74.45
125	66.89	62.80	62.78	62.78	62.80	62.78	62.78	62.78	62.80	62.78	62.78	62.78	62.78	62.78
$\rho^2 = 0.99$														
50	439.82	389.89	178.39	149.34	354.85	125.30	142.54	155.06	247.86	124.89	134.31	150.82	138.25	124.80
75	297.40	245.96	139.70	134.85	228.39	132.93	137.12	136.86	203.68	132.93	135.14	140.35	141.54	132.93
100	241.08	190.84	130.94	129.70	182.10	129.64	130.56	129.93	170.69	129.64	130.12	131.49	134.68	129.64
125	201.16	158.76	124.38	124.13	153.96	124.13	124.32	124.13	148.25	124.13	124.23	124.54	127.10	124.13
$\rho^2 = 0.999$														
50	1367.62	1343.94	640.13	606.59	1273.43	94.72	207.79	283.35	615.44	72.97	149.99	237.34	130.49	70.68
75	926.42	894.61	353.12	349.79	830.52	89.13	211.31	266.07	541.72	84.86	149.66	221.28	136.87	83.06
100	749.91	711.52	279.67	267.49	655.68	96.69	194.12	233.27	457.00	95.65	145.10	200.91	137.11	94.89
125	626.78	585.03	231.18	207.46	525.97	105.33	168.85	199.66	385.94	104.09	136.84	179.56	137.01	103.66

of intercept becomes lower. This is since the average value of $\hat{\theta}$ decreases with β_o, which leads to a lower value of the diagonal weight matrix of $\hat{V}$. The shrinkage estimators D_{kp4} and D_{q4} are providing the minimum MSE when $\rho^2 = 0.90$. Among these shrinkage estimators, the D_{q4} estimator is the best and provides the minimum MSE in the presence of severe multicollinearity. For a limited case (when $\rho^2 = 0.90$ and $n = 100$), the performance of D_5, D_{kp4}, and D_{q4} is the same. Overall, the proposed shrinkage parameters (D_{kp1}–D_{kp4}) and (D_{q1}–D_{q4}) outperform the PLRE with the parameters suggested by Månsson *et al.* [18]. In addition, these parameters also outperform MLE significantly. Finally, based on the simulated results, we conclude that the PLRE should be used with the shrinkage parameter D_{q4} in the presence of high, but imperfect multicollinearity since it has the lowest estimated MSE and MAPE as compared to the other shrinkage parameters.

4. Empirical application: *Swedish football data*

For the purpose of illustrating the empirical relevance of the proposed methods, we analyze Swedish football data in this empirical section. The proposed and existing estimation methods are elucidated using a dataset regarding the performance of Swedish football teams in the top Swedish league (Allsvenskan) during the year of 2018.[1] The aims of this application are twofold. Firstly, to analyze the number of full-time home team goals (FTHTG) which is illustrated in Table 7. Secondly, to analyze the number of full-time away team goals (FTATG) which is presented in Table 8. This dataset includes $n = 242$ observations and include two dependent variables as described above and six explanatory variables, which are the pinnacle home win odds (x_1), pinnacle away win odds (x_2), maximum oddsportal maximum home win (x_3), oddsportal maximum away win (x_4), average oddsportal home win (x_5) and average oddsportal away win (x_6). The effect of these regressors on FTHTG and FTATG, respectively are demonstrated in the regression. The bivariate correlations among the explanatory variables are demonstrated in Table 6. It is seen from Table 6 that there are high correlations in half of the cases, and moderate correlations among rest of the cases. In addition, the condition number which is the ratio of maximum to the minimum eigenvalues, is $6013.22 > 1000$ which indicates what can be defined as a severe multicollinearity problem in this dataset which originates from Türkan and Özel [24].

The Chi square (χ^2) goodness of fit test is used before applying the Poisson regression (PR) model for the empirical application. As is seen in Tables 7 and 8, these tests confirm that the response variables are well suited to the PR model with p-values corresponding to 0.77 and 0.88. Based on the analysis of the dataset using the standard *glm*{} package in R, the estimated coefficients, standard errors, and values of the MSE and cross validation criterion estimates[2] (CVCE) are summarized in Tables 7 and 8 to assess the performance of the PLRE and MLE. The effect of the estimated coefficients is changed, and the estimated standard errors and the estimated MSE of the PLRE are smaller than the MLE due

Table 6. Correlation matrix.

Variables	x_1	x_2	x_3	x_4	x_5	x_6
x_1	1.0000					
x_2	−0.5746	1.0000				
x_3	0.9972	−0.5681	1.0000			
x_4	−0.5549	0.9938	−0.5488	1.0000		
x_5	0.9952	−0.5949	0.9974	−0.5751	1.0000	
x_6	−0.5885	0.9948	−0.5824	0.9959	−0.6093	1.0000

Table 7. Estimated coefficients, standard errors, MSE and CVCE for FTHTG application[a].

	MLE		PLRE (D_{q4})	
Variables	Coefficients	Standard Errors	Coefficients	Standard errors
x_1	0.7825	0.6392	0.4379	0.4819
x_2	0.0307	0.1317	0.0254	0.1295
x_3	−0.9190	0.8230	−0.4911	0.5627
x_4	−0.1796	0.1718	−0.1667	0.1673
x_5	0.0806	0.7601	-0.0368	0.5324
x_6	0.2562	0.2280	0.2439	0.2178
MSE	1.927087		0.679863	
CVCE	1.576436		1.576388	

[a]The response variable is FTHTG which is well fitted to the Poisson distribution ($\chi^2 = 2.5131, df = 5, p = 0.7745$).

Table 8. Estimated coefficients, standard errors, MSE and CVCE for FTATG application[a].

	MLE		PLRE(Dq4)	
Variables	Coefficients	Standard errors	Estimates	Standard errors
x_1	−0.0085	0.3177	−0.0274	0.2908
x_2	−0.0592	0.2919	−0.0690	0.2702
x_3	−0.3570	0.4911	−0.2618	0.4117
x_4	0.1355	0.2742	0.1162	0.2574
x_5	0.5346	0.4688	0.4416	0.3983
x_6	−0.2052	0.4305	−0.1714	0.3704
MSE	1.097734		0.637463	
CVCE	1.369209		1.369167	

[a]The response variable is FTATG which is well fitted to the Poisson distribution ($\chi^2 = 1.7359, df = 5, p = 0.8843$).

to the high, but imperfect multicollinearity. The different estimators give qualitatively the same results, and in order to save space we focus to analyze the D_{q4} shrinkage estimator since it performs better than the other estimators in the simulation study – given a degree of correlation. Therefore, we used PLRE with D_{q4} but of course full results for all estimators are available from the authors upon request. It is evident from Table 7, based on high standard errors and MSE, that the MLE do not estimate the coefficients very precisely in the presence of multicollinearity. However, on the other hand, the proposed estimation method, estimates the coefficients rather precisely. For instance, theoretically, oddsportal maximum away win have negative effects on the FTHTG, while the MLE shows a positive effect. Meanwhile, proposed method shows negative affect and it is considered a good approach to tackle the problem of multicollinearity. The estimated results of the second model are shown in Table 8, where we can observe that the standard errors and the values of the MSE and CVCE are high when the MLE is used. However, these estimated results are reduced when applying the PLRE with D_{q4} a shrinkage estimator. Hence, the advantage of the proposed method over MLE by means of an empirical application is fairly easily illustrated. However, cross-validation shows little improvement in the predictive power between the methods. The plots of $MSE(\beta_{PLRE})$ and $MSE(\beta_{MLE})$ against different values of d in the interval [0,1] has been presented for FTHTG in Figure 1. It is noted that the

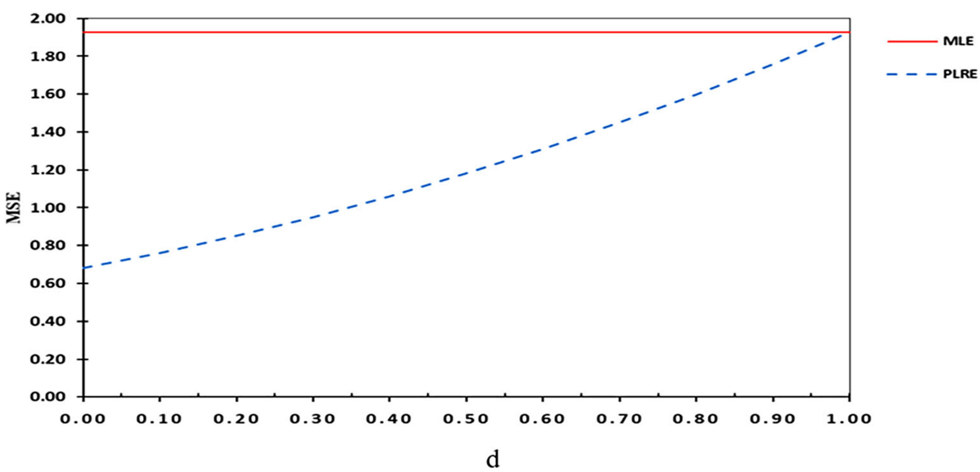

Figure 1. Plot of $MSE(\beta_{PLRE})$ and $MSE(\beta_{MLE})$ against different values of d.

estimated MSE of PLRE equals to MLE when the value of d equals 1 and it decreases as the value of d becomes close to 0. Therefore, we can say that the performance of the PLRE is a function of the values of the shrinkage estimators.

5. Concluding remarks

This article proposed new shrinkage estimators and conducts a comparison with existing estimators by means of a Monte Carlo simulation and an empirical application. The MSE and MAPE are considered as the performance criteria in the evaluation. These estimators are proposed in order to minimize the increase of the MLE caused by multicollinearity. The simulation results illustrated that the estimated MSE and MAPE are clearly affected by changing different factors such as the value of intercept, number of explanatory variables, multicollinearity level and the sample size. However, the general assessment is that the performance of PLRE is superior than the MLE under very different, but empirically relevant, conditions. Based on the Monte Carlo simulations and football dataset, we conclude that the D_{q4} shrinkage parameter should be applied for the PLRE whenever the practitioner, in the presence of considerable multicollinearity, needs to apply the PR model.

Notes

1. The data are publicly available on the webpage www.football-data.co.uk. The data are also available from the authors upon request.
2. The mean squared prediction error is calculated by CVCE as $\sum_{i=1}^{n}(\hat{Y}_{-i} - Y_i)^2/n$, where $\sum_{i=1}^{n}(\hat{Y}_{-i} - Y_i)^2$ is denoted as the prediction sum of squares, and we optimize toward those models which has the smallest CVCE. We apply the leave-one-out cross validation (LOOCV) approach for computation of $\hat{Y}_{-i}$. In LOOCV and fit the model n times. We leave out the ith value at step i and use the resulting fitted model to calculate the predicted value for the leave out ith observation, $\hat{Y}_{-i}$.

Acknowledgments

The authors would like to thank the Editor, associate Editor and anonymous referees for their valuable comments and suggestions that improved the quality of this paper greatly.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

B. M. G. Kibria http://orcid.org/0000-0002-6073-1978

References

[1] M.I. Alheety and B.M.G. Kibria, *On the Liu and almost unbiased Liu estimators in the presence of multicollinearity with heteroscedastic or correlated errors*. Surv. Math. Appl. 4 (2009), pp. 155–167.

[2] M. Amin, M. Qasim, and M. Amanullah, *Performance of Asar and Genç and Huang and Yang's two-parameter*. Iran. J. Sci. Technol. Trans. A Sci. 43 (2019), pp. 2951–2963.

[3] M. Amin, M. Qasim, M. Amanullah, and S. Afzal, *Performance of some ridge estimators for the gamma regression model*. Stat. Pap. (2017), pp. 1–30. doi:10.1007/s00362-017-0971-z.

[4] M. Arashi, B.M.G. Kibria, M. Norouzirad, and S. Nadarajah, *Improved preliminary test and Stein-rule Liu estimators for the ill-conditioned elliptical linear regression model*. J. Multivar. Anal. 126 (2014), pp. 53–74.

[5] M. Arashi, S. Nadarajah, and F. Akdeniz, *The distribution of the Liu-type estimator of the biasing parameter in elliptically contoured models*. Commun. Stat. Theory Methods 46 (2017), pp. 3829–3837.

[6] R. Frisch, *Statistical Confluence Analysis by Means of Complete Regression Systems*, Publication 5, University Institute of Economics, Oslo, 1934.

[7] D.G. Gibbons, *A simulation study of some ridge estimators*. J. Amer. Statist. Assoc. 76 (1981), pp. 131–139.

[8] A.E. Hoerl and R.W. Kennard, *Ridge regression: Biased estimation for nonorthogonal problems*. Technometrics 12 (1970), pp. 55–67.

[9] D. Inan and B.E. Erdogan, *Liu-type logistic estimator*. Commun. Stat.-Simul. Comput. 42 (2013), pp. 1578–1586.

[10] S. Kaçiranlar, *Liu estimator in the general linear regression model*. J. Appl. Stat. Sci. 13 (2003), pp. 229–234.

[11] M.H. Karbalaee, S.M.M. Tabatabaey, and M. Arashi, *On the preliminary test generalized Liu estimator with series of stochastic restrictions*. J. Iran. Stat. Soc. 18 (2019), pp. 113–131.

[12] L. Kejian, *A new class of biased estimate in linear regression*. Commun. Stat.-Theory Methods 22 (1993), pp. 393–402.

[13] B.M.G. Kibria, *Performance of some new ridge regression estimators*. Commun. Stat.-Simul. Comput. 32 (2003), pp. 419–435.

[14] B.M.G. Kibria, *Some Liu and ridge-type estimators and their properties under the ill-conditioned Gaussian linear regression model*. J. Stat. Comput. Simul. 82 (2012), pp. 1–17.

[15] B.M.G. Kibria, K. Månsson, and G. Shukur, *Performance of some logistic ridge regression estimators*. Comput. Econ. 40 (2012), pp. 401–414.

[16] B.M.G. Kibria, K. Månsson, and G. Shukur, *A simulation study of some biasing parameters for the ridge type estimation of Poisson regression*. Commun. Stat.-Simul. Comput 44 (2015), pp. 943–957.

[17] K. Månsson, *Developing a Liu estimator for the negative binomial regression model: method and application*. J. Stat. Comput. Simul. 83 (2013), pp. 1773–1780.

[18] K. Månsson, B.M.G. Kibria, and G. Shukur, *On Liu estimators for the logit regression model*. Econ. Model. 29 (2012), pp. 1483–1488.

[19] K. Månsson, B.M.G. Kibria, P. Sjölander, and G. Shukur, *Improved Liu estimators for the Poisson regression model*. Int. J. Stat. Prob. 1 (2012), pp. 2.

[20] K. Månsson and G. Shukur, *A Poisson ridge regression estimator*. Econ. Model. 28 (2011), pp. 1475–1481.

[21] M. Qasim, M. Amin, and M. Amanullah, *On the performance of some new Liu parameters for the gamma regression model*. J. Stat. Comput. Simul. 88 (2018), pp. 3065–3080.

[22] M. Qasim, M. Amin, and T. Omer, *Performance of some new Liu parameters for the linear regression model*. Commun. Stat.-Theory Methods (2019), pp. 1–19. doi:10.1080/03610926.2019.1595654.

[23] GÜ Şiray, S. Toker, and S. Kaçiranlar, *On the restricted Liu estimator in the logistic regression model*. Commun. Stat.-Simul. Comput. 44 (2015), pp. 217–232.

[24] S. Türkan, and G. Özel, *A new modified Jackknifed estimator for the Poisson regression model*. J. Appl. Stat. 43 (2016), pp. 1892–1905.

[25] J. Wu, Y. Asar, and M. Arashi, *On the restricted almost unbiased Liu estimator in the logistic regression model*. Commun. Stat. Simul. Comput. 47 (2018), pp. 4389–4401.