Estimating the Unrestricted and Restricted Liu Estimators for the Poisson Regression Model: Method and Application

Kristofer Månsson¹ · B. M. Golam Kibria²

Accepted: 16 July 2020 / Published online: 26 August 2020
© The Author(s) 2020

Abstract

This paper considers both unrestricted and restricted Liu estimators in the presence of multicollinearity for the Poisson regression model. It also considers some new estimators of the shrinkage parameter for both unrestricted and restricted Liu estimators. Based on a simulation study and its empirical application, we found that the restricted estimator outperforms the unrestricted one. Further, the restricted Liu estimator also outperforms both the unrestricted Liu and restricted Liu estimators. Hence, this new method is a preferred option when the coefficient vector β may belong to a linear sub-space defined by $R\beta = r$.

Keywords Liu estimator · Maximum likelihood · Monte Carlo simulations · MSE · Multicollinearity · Poisson regression · Restricted estimator

Mathematics Subject Classification Primary 62J07 · Secondary 62J02

1 Introduction

In empirical applications in economics when modeling the gene expression data, the dependent variables are often in the form of non-negative integers of counts, for example, in analyzing the determinants of the number of patents by a firm or in bio-statistics. In such a situation, a common distributional assumption for the dependent variable y_i is that it follows a Poisson distribution. More specifically, it is distributed as $Po(\mu_i)$, where $\mu_i = \exp(\mathbf{x}_i'\boldsymbol{\beta})$, $\mathbf{x}_i$ is the i th row of $\mathbf{X}$ which is a $n \times (p + 1)$ data matrix with p explanatory variables and $\boldsymbol{\beta}$ is a $(p + 1) \times 1$ vector of coefficients. The parameter for the unrestricted Poisson regression model is usually estimated using the iterative weighted least squares (IWLS) algorithm as:

✉ Kristofer Månsson
kristofer.mansson@ju.se

¹ Department of Economics, Finance and Statistics, Jönköping University, Jönköping, Sweden

² Department of Mathematics and Statistics, Florida International University, Miami, FL, USA

$$\beta_{ML} = (\mathbf{X}'\hat{\mathbf{W}}\mathbf{X})^{-1}(\mathbf{X}'\hat{\mathbf{W}}\hat{\mathbf{z}}), \quad (1)$$

where $\hat{\mathbf{W}} = \text{diag}[\hat{\mu}_i]$ and $\hat{\mathbf{z}}$ is a vector where the i th element equals $\hat{z}_i = \log(\hat{\mu}_i) + \frac{y_i - \hat{\mu}_i}{\hat{\mu}_i}$. Hence, it is an iterative maximum likelihood (ML) procedure. As has been shown by Kaçiranlar and Dawoud (2018), Månsson et al. (2012) and Türkan and Özel (2016) among others this estimator is sensitive to multicollinearity. A common solution to this problem is applying the ridge regression method (Hoerl and Kennard 1970). Another widely used method is applying the Liu type estimator introduced by Liu (1993) and introduced in the context of the Poisson regression model by Månsson et al. (2012). This method has certain advantages over the ridge regression method since it is a linear function of the biasing parameter d . Therefore, this method has become more popular during recent years (see, for example, Akdeniz and Kaciranlar 1995; Kaciranlar 2003). The unrestricted Liu type estimator for the Poisson regression model is defined by Månsson et al. (2012) as:

$$\hat{\beta}_d = (\mathbf{X}'\hat{\mathbf{W}}\mathbf{X} + \mathbf{I})^{-1}(\mathbf{X}'\hat{\mathbf{W}}\mathbf{X} + d\mathbf{I})\hat{\beta}_{ML} \quad (2)$$

where $0 \leq d \leq 1$. Different versions of this estimator for count data models have been considered by Asar and Genç (2017), where a two parameter estimator is considered, Asar (2018), where the author discussed different perspectives on this estimator and Türkan and Özel (2018) in which the authors considered a jack-knifed estimator. A review article with detailed literature review is written by Algamal (2018). In some situations, the coefficient vector β is suspected to belong to a linear sub-space:

$$\mathbf{R}\beta = \mathbf{r}, \quad (3)$$

where $\mathbf{R}$ is an $q \times p$ matrix of full rank with $q < p$ and $\mathbf{r}$ is an $q \times 1$ vector. Under this restriction, the restricted estimator is more efficient than the OLS estimator. This method has been considered by several authors for the linear regression model (Alheety and Kibria 2009; Ehsanes Saleh and Golam Kibria 1993). The restricted maximum likelihood estimator (RMLE) of β suggested by Kibria and Saleh (2012) is given by:

$$\hat{\beta}_R = \hat{\beta}_{ML} - \hat{\mathbf{C}}^{-1}\mathbf{R}'[\mathbf{R}\hat{\mathbf{C}}^{-1}\mathbf{R}']^{-1}(\mathbf{R}\hat{\beta}_{ML} - \mathbf{r}), \quad (4)$$

where $\hat{\mathbf{C}} = \mathbf{X}'\hat{\mathbf{W}}\mathbf{X}$. However, the issue of multicollinearity still exists since the inverse of $\hat{\mathbf{C}}$ is instable as it is near singular. Therefore, the estimated coefficient vector has a high variance. As a solution one can use a combination of the Liu type estimator in Eq. (2) and the restricted estimator in Eq. (4). Kaçiranlar et al. (1999) have already done this and it has also been used by many others (Kibria and Saleh 2012; Li and Yang 2010; Yang and Xu 2009). Also for the logit model, Şiray et al. (2015) suggested the following estimator:

$$\hat{\beta}_{R,Liu} = \hat{F}_d \hat{\beta}_{ML} - \hat{F}_d \hat{C}^{-1} R' [R \hat{C}^{-1} R']^{-1} (R \hat{\beta}_{ML} - r). \quad (5)$$

where $\hat{F}_d = (X' \hat{W} X + I)^{-1} (X' \hat{W} X + dI)$. This estimator has later been considered in Månsson et al. (2016), Wu and Asar (2017) and Varathan and Wijekoon (2018). The purpose of our paper is to introduce the restricted estimator and the restricted Liu type estimator in the context of the Poisson regression model. We also introduce some new estimators for the shrinkage parameter d . The new methods are mainly evaluated using a simulation study where one can clearly see the benefits of applying a restricted estimator.

The rest of the paper is organized as follows: Sect. 2 gives the statistical methodology and estimation technique. A simulation study has been conducted in Sect. 3 and Sect. 4 analyzes real life data to illustrate our findings. Some concluding remarks are given in Sect. 5.

2 Statistical Methodology

This section starts with the MSE properties of different estimators and then suggests some estimators for the shrinkage parameter d .

2.1 MSE Properties of Different Estimators

This section first considers the MSE functions of the estimators (a detailed discussion on the derivations of the MSE functions for the Liu model are found in Månsson et al. 2012) and the expressions are presented as:

$$MSE(\hat{\beta}_{ML}) = \sum \frac{1}{\lambda_j} \quad (6)$$

$$MSE(\hat{\beta}_d) = \sum \frac{(\lambda_j + d)^2}{\lambda_j (\lambda_j + 1)^2} + (d - 1)^2 \sum \frac{\alpha_j^2}{(\lambda_j + 1)^2} \quad (7)$$

where α_j^2 corresponds to the j th element of $\gamma \beta$ and γ is the eigenvector defined so that $X' W X = \gamma' \Lambda \gamma$, where Λ equals $diag(\lambda_j)$ and λ_i are eigenvalues of the matrix $X' W X$. Here it is clear that the Liu estimator can be superior to the ML estimator if one can find a value of d such that an increase in MSE due to the bias in the second term in Eq. (7) is smaller than the decrease in the variance in the first term (for details see Månsson 2013). The MSE functions for the restricted estimator and the restricted Liu type of estimator are discussed in detail in Månsson et al. (2016) and shown as:

$$MSE(\hat{\beta}_R) = tr(A) = \sum b_{jj} \quad (8)$$

$$MSE(\hat{\beta}_{R,Liu}) = \sum \frac{(\lambda_j + d)^2}{(\lambda_j + 1)^2} b_{jj} + (d - 1)^2 \sum \frac{\alpha_j^2}{(\lambda_j + 1)^2} \quad (9)$$

where $\mathbf{A} = \mathbf{C}^{-1} - \mathbf{C}^{-1} \mathbf{R}' (\mathbf{R} \mathbf{C}^{-1} \mathbf{R}')^{-1} \mathbf{R} \mathbf{C}^{-1}$ and b_{jj} are the inverse of the eigenvalues of $\mathbf{A}$ matrix. The effect of multicollinearity is that some of the eigenvalues of the $\mathbf{X}'\mathbf{W}\mathbf{X}$ matrix are small which inflates the MSE of the restricted Poisson estimator. However, it is possible to find a value of the shrinkage parameter d such that the restricted Liu estimator outperforms the restricted estimator (see Månsson et al. 2016).

2.2 Estimation of the Shrinkage Parameter

Following Månsson (2013) we suggest the following estimators of the shrinkage parameter d for the unrestricted Liu estimator:

$$D1 = \max \left(0, \frac{\hat{\alpha}_{\max}^2 - 1}{\frac{1}{\hat{\lambda}_{\max}} + \hat{\alpha}_{\max}^2} \right), \quad D2 = \max \left(0, \text{median} \left(\frac{\hat{\alpha}_j^2 - 1}{\frac{1}{\hat{\lambda}_j} + \hat{\alpha}_j^2} \right) \right), \quad D3 = \max \left(0, \frac{1}{p} \sum_j \frac{\hat{\alpha}_j^2 - 1}{\frac{1}{\hat{\lambda}_j} + \hat{\alpha}_j^2} \right)$$

The first one is a version of the Hoerl and Kennard (1970) estimator and the other two are based on Kibria (2003). For the restricted Poisson estimator, we follow Månsson et al. (2016) and suggest the following estimators:

$$D1 = \max \left(0, \frac{\hat{\alpha}_{\max}^2 - 1}{\frac{1}{\hat{\lambda}_{\max}} + \hat{\alpha}_{\max}^2} b_{jj,\max} \right), \quad D2 = \max \left(0, \text{median} \left(\frac{\alpha_j^2 - \lambda_j}{b_{jj} + \alpha_j^2} b_{jj} \right) \right)$$

and $D3 = \max \left(0, \frac{1}{p} \sum_{j=1}^p \frac{\alpha_j^2 - \lambda_j}{b_{jj} + \alpha_j^2} b_{jj} \right).$

We propose the following new estimators: $D4 = \min (D1, D2, D3)$, $D5 = \text{median} (D1, D2, D3)$ and $D6 = \max (D1, D2, D3)$. This is done for both the unrestricted and restricted estimators.

3 Simulation Study

Since a theoretical comparison of the estimators is not possible, in this section we do a simulation study to compare the performance of the shrinkage estimators.

3.1 Simulation Techniques

Since the dependent variable is count, it is generated using pseudo-random numbers from the Poisson distribution with mean, μ_i , i.e. $Po(\mu_i)$ where:

$$\mu_i = \exp(\beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip}), \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, p \quad (10)$$

The parameter values in Eq. (10) are chosen so that sum of squared parameters equals one. The main factor we chose to vary is the degree of correlation between the explanatory variables which is defined as ρ . We consider the following three values of ρ : 0.80, 0.90, and 0.95. Following Kibria (2003), we generate explanatory variables with different degrees of correlation as:

$$x_{ij} = (1 - \rho^2)^{(1/2)} z_{ij} + \rho z_{ip}, \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, p \quad (11)$$

where z_{ij} are pseudo-random numbers generated using the standard normal distribution. To reduce eventual start-up value effects, we discard the first 200 observations. We change the number of independent variables and set $p=4, 8$ and 12 . We also change the number of observations and generate models with 20, 30, 50 and 100 degrees of freedom (denoted by df). The restrictions when $p=4$ are set as:

$$R = [1 \quad 0 \quad -2 \quad 1] \text{ and } r = [0] \quad (12)$$

while the restrictions when $p=8$ are:

$$R = [1 \quad 0 \quad -2 \quad 1 \quad -3 \quad 1 \quad 1 \quad 1] \text{ and } r = [0] \quad (13)$$

Finally, the restrictions when $p=12$ are:

$$R = [1 \quad 0 \quad -2 \quad 1 \quad -3 \quad 1 \quad 1 \quad 1 \quad 1 \quad 0 \quad -2 \quad 1] \text{ and } r = [0] \quad (14)$$

We repeated this experiment 2000 times by generating new pseudo-random numbers and then estimated MSE using the formula:

$$MSE = \frac{\sum_{i=1}^{2000} (\hat{\beta} - \beta)^2}{2000}, \quad (15)$$

where $\hat{\beta}$ is the estimated parameter for the restricted estimator or the restricted Liu estimator and β is the true value of the parameter.

3.2 Results

The simulated MSE values for the unrestricted and restricted estimators are presented in Tables 1 and 2 for $p=4$; in Tables 3 and 4 for $p=8$; and in Tables 5 and 6 for $p=12$. To save space, we only give the MSE values graphically for $p=4$ in Figs. 1 and 2 for unrestricted and restricted estimators respectively. In the appendix the result for the proportion of nonzero values of d are shown in Tables 11, 12, 13. The proportions are the same for the unrestricted and the restricted estimators. One can see that the degree of correlation increases the proportion of nonzero values decreases since the instability increases. The reverse effect may also be noticed for the sample size. The estimator with the least non-zero values is the optimal ones since they tend to induce more bias which is necessary to

Table 1 Estimated MSE for unrestricted Poisson regression model for $p = 4$

	ML	D1	D2	D3	D4	D5	D6
<i>df</i>	$\rho = 0.80$						
20	1.510	0.702	0.726	0.697	0.692	0.702	0.730
30	1.094	0.618	0.625	0.616	0.616	0.618	0.625
50	0.839	0.521	0.521	0.521	0.521	0.521	0.521
100	0.679	0.409	0.409	0.409	0.409	0.409	0.409
<i>df</i>	$\rho = 0.90$						
20	2.624	0.763	0.944	0.766	0.753	0.774	0.945
30	1.862	0.758	0.831	0.760	0.758	0.761	0.831
50	1.191	0.640	0.647	0.640	0.640	0.640	0.647
100	0.852	0.529	0.529	0.529	0.529	0.529	0.529
<i>df</i>	$\rho = 0.95$						
20	5.210	0.762	1.394	0.802	0.747	0.810	1.400
30	3.225	0.777	1.033	0.773	0.771	0.776	1.036
50	2.031	0.774	0.857	0.774	0.774	0.774	0.857
100	1.217	0.665	0.671	0.665	0.665	0.665	0.671

Table 2 Estimated MSE for restricted Poisson regression model for $p = 4$

	ML	D1	D2	D3	D4	D5	D6
<i>df</i>	$\rho = 0.80$						
20	0.991	0.511	0.521	0.510	0.509	0.510	0.522
30	0.710	0.468	0.470	0.468	0.468	0.468	0.470
50	0.517	0.413	0.413	0.413	0.413	0.413	0.413
100	0.380	0.345	0.345	0.345	0.345	0.345	0.345
<i>df</i>	$\rho = 0.90$						
20	1.760	0.540	0.608	0.547	0.539	0.548	0.609
30	1.262	0.550	0.580	0.551	0.550	0.551	0.581
50	0.749	0.469	0.472	0.469	0.469	0.469	0.472
100	0.523	0.417	0.417	0.417	0.417	0.417	0.417
<i>df</i>	$\rho = 0.95$						
20	3.444	0.550	0.804	0.574	0.548	0.573	0.807
30	2.226	0.562	0.658	0.562	0.561	0.562	0.658
50	1.374	0.562	0.597	0.562	0.562	0.562	0.597
100	0.788	0.492	0.494	0.492	0.492	0.492	0.494

combat the multicollinearity problem. The pattern is the same for all p , df and ρ . One can see that the MSE values for all the estimators increase with the number of explanatory variables and the degree of multicollinearity and decrease with an increase in the number of observations. As can be observed from the tables and figures the MSE for the unrestricted ML estimator is always higher than the restricted estimator. There is always (for all sample sizes and different number of explanatory variables) a benefit in using the restricted estimator when

Table 3 Estimated MSE for unrestricted Poisson regression model for $p = 8$

		ML	D1	D2	D3	D4	D5	D6
<i>df</i>	$\rho = 0.80$							
20		6.958	1.185	1.240	1.177	1.177	1.185	1.241
30		5.217	0.972	0.978	0.972	0.972	0.972	0.978
50		4.247	0.696	0.697	0.696	0.696	0.696	0.697
100		3.539	0.389	0.389	0.389	0.389	0.389	0.389
<i>df</i>	$\rho = 0.90$							
20		11.253	1.646	1.904	1.647	1.623	1.659	1.914
30		7.703	1.426	1.481	1.426	1.426	1.426	1.481
50		5.600	1.128	1.134	1.128	1.128	1.128	1.134
100		4.439	0.729	0.730	0.729	0.729	0.729	0.730
<i>df</i>	$\rho = 0.95$							
20		19.397	2.009	3.124	2.035	2.002	2.041	3.125
30		12.445	1.862	2.218	1.870	1.862	1.869	2.219
50		8.374	1.615	1.683	1.615	1.615	1.615	1.683
100		5.750	1.166	1.168	1.166	1.166	1.166	1.168

Table 4 Estimated MSE for restricted Poisson regression model for $p = 8$

		ML	D1	D2	D3	D4	D5	D6
<i>df</i>	$\rho = 0.80$							
20		2.555	1.007	1.028	1.007	1.007	1.007	1.028
30		1.457	0.805	0.807	0.805	0.805	0.805	0.807
50		0.808	0.569	0.570	0.569	0.569	0.569	0.570
100		0.367	0.316	0.316	0.316	0.316	0.316	0.316
<i>df</i>	$\rho = 0.90$							
20		5.573	1.444	1.566	1.452	1.444	1.452	1.566
30		3.195	1.233	1.261	1.233	1.233	1.233	1.261
50		1.735	0.942	0.944	0.942	0.942	0.942	0.944
100		0.836	0.602	0.602	0.602	0.602	0.602	0.602
<i>df</i>	$\rho = 0.95$							
20		11.388	1.849	2.408	1.860	1.849	1.860	2.408
30		6.740	1.673	1.862	1.675	1.673	1.675	1.862
50		3.751	1.418	1.452	1.418	1.418	1.418	1.452
100		1.766	0.999	1.000	0.999	0.999	0.999	1.000

the restriction holds. Further, we also observe that the unrestricted Liu estimator outperforms both the restricted and unrestricted ML estimators in the presence of multicollinearity. One can also see that the performance of the estimators of the shrinkage parameter is fairly similar. However, the new estimator D4 consistently performs better than the other estimators. The potential explanation for this is that by using information from three different estimators we can find one that is usually closer to the optimal value.

Table 5 Estimated MSE for unrestricted Poisson regression model for $p = 12$

		ML	D1	D2	D3	D4	D5	D6
<i>df</i>	$\rho = 0.80$							
20	7.082	1.588	1.639	1.576	1.546	1.582	1.645	
30	3.784	1.427	1.426	1.426	1.406	1.426	1.427	
50	2.073	1.066	1.066	1.066	1.046	1.066	1.066	
100	1.154	0.648	0.648	0.648	0.638	0.648	0.648	
<i>df</i>	$\rho = 0.90$							
20	14.230	1.847	2.081	1.847	1.827	1.847	2.081	
30	4.225	1.003	1.024	1.003	0.990	1.003	1.024	
50	1.853	0.734	0.734	0.734	0.714	0.734	0.734	
100	2.023	1.059	1.059	1.059	1.052	1.059	1.059	
<i>df</i>	$\rho = 0.95$							
20	28.814	1.956	3.454	1.995	1.926	1.995	3.454	
30	15.722	2.136	2.444	2.136	2.046	2.136	2.444	
50	8.210	2.069	2.079	2.069	2.029	2.069	2.079	
100	3.951	1.586	1.586	1.586	1.556	1.586	1.586	

Table 6 Estimated MSE for restricted Poisson regression model for $p = 12$

		ML	D1	D2	D3	D4	D5	D6
<i>df</i>	$\rho = 0.80$							
20	5.149	1.418	1.423	1.417	1.397	1.417	1.424	
30	2.895	1.280	1.280	1.280	1.240	1.280	1.280	
50	1.572	0.974	0.974	0.974	0.954	0.974	0.974	
100	0.746	0.596	0.596	0.596	0.576	0.596	0.596	
<i>df</i>	$\rho = 0.90$							
20	10.470	1.703	1.781	1.703	1.683	1.703	1.781	
30	3.232	0.913	0.922	0.913	0.903	0.913	0.922	
50	1.495	0.667	0.667	0.667	0.647	0.667	0.667	
100	1.432	0.971	0.971	0.971	0.951	0.971	0.971	
<i>df</i>	$\rho = 0.95$							
20	21.666	1.836	2.678	1.855	1.796	1.855	2.678	
30	12.008	1.983	2.151	1.983	1.943	1.983	2.151	
50	6.506	1.907	1.912	1.907	1.887	1.907	1.912	
100	3.109	1.450	1.450	1.450	1.430	1.450	1.450	

4 Empirical Application

In this empirical application, we use a dataset from Cameron and Trivedi (2013) to estimate a recreation demand function. This is survey data on the number of recreational boating trips to Lake Somerville, East Texas in 1980. We use some explanatory variables where we focus on the cost of taking a trip to that particular lake and also to competing or substitute boating attractions at Lake Conroe and Lake Houston. When considering competitors' prices in this demand system, it is expected

Fig. 1 Estimated MSE when $p=4$ for the unrestricted estimator

that we will find multicollinearity since the costs are not independent. The correlation matrix is given in Table 7 in which one can see that all prices are correlated with at least a value of 0.87 as the correlation coefficient. This indicates a multicollinearity problem.

When looking at the coefficients of the unrestricted ML estimator in Table 8 one can see that the value is negative for lakes Conroe and Sommerville and positive for Lake Houston. The negative coefficient for Lake Conroe is not as per theory since an increase in a competitor's cost should increase demand for the others.

Fig. 2 Estimated MSE when $p=4$ for the restricted estimator

Table 7 Correlation matrix

	Lake Conroe	Lake Somerville	Lake Houston
Lake Conroe	1		
Lake Somerville	0.909	1	
Lake Houston	0.963	0.871	1

Table 8 The estimated coefficients

	Unrestricted estimator				Restricted estimator			
	Lake Conroe	Lake Somerville	Lake Houston	Value of d	Lake Conroe	Lake Somerville	Lake Houston	Value of d
ML	-0.655	-1.969	2.646	-	-0.071	-1.976	1.976	-
D1	-0.579	-1.963	2.557	0.33	-0.039	-1.972	1.940	0.27
D2	-0.585	-1.964	2.564	0.27	-0.039	-1.972	1.940	0.27
D3	-0.579	-1.963	2.557	0.33	-0.039	-1.972	1.940	0.27
D4	-0.579	-1.963	2.557	0.32	-0.039	-1.972	1.940	0.27
D5	-0.579	-1.963	2.557	0.33	-0.039	-1.972	1.940	0.27
D6	-0.585	-1.964	2.564	0.29	-0.039	-1.972	1.940	0.27

This may be due to multicollinearity as it may produce a wrong sign (Hoerl and Kennard 1970). Excluding this variable has only a very marginal effect on the Akaike Information Criteria (it increases from 3978 to 3993). When Lake Houston's price is excluded AIC increases to over 5000. The counter intuitive value of the coefficient and the marginal effect that this variable actually has leads us to the following restriction:

$$R = [01 \quad 1] \text{ and } r = [0] \quad (16)$$

Hence, we assume that a substitution effect exists only between Lake Somerville and Lake Houston. The coefficients for the full sample for both the restricted and unrestricted estimators is provided in Table 8. The largest effect is for Lake Conroe where the coefficient remains negative but is pushed towards zero indicating no substitution effect. The estimated coefficients are fairly similar for both ML and shrinkage estimators. This in all probability is due to the large sample size (660 observations). To illustrate the differences in the methods we chose to bootstrap the different observations. This is also done to link the results of our simulation study to empirical applications by changing the sample size. We chose to bootstrap 50, 100 and the full sample which corresponds to 660 observations. The full sample's standard errors are standard ones reported and associated with the coefficients in Table 8. The bootstrapped standard errors for the shrinkage estimators are given in Tables 9 and 10 for the unrestricted and restricted estimators respectively. By bootstrapping different observation (50, 100 and the full sample of 660 observations) one can clearly see that the standard error decreases for all estimators as the sample size increases; this is as expected from theory and the simulation study. We also see that the benefits of using a shrinkage estimator are higher for a small number of observations since a decrease (compared to ML) in standard errors is more significant in such a situation. Further, the standard errors are smaller for all sample sizes for the restricted estimator. The unrestricted Liu estimator is also better than the corresponding unrestricted ML estimator. But

Table 9 Standard errors for the unrestricted estimator

	<i>ML</i>	<i>D1</i>	<i>D2</i>	<i>D3</i>	<i>D4</i>	<i>D5</i>	<i>D6</i>
<i>n = 50</i>							
Lake Conroe	3.168	3.011	3.032	2.569	2.440	2.974	3.091
Lake Somerville	1.880	1.848	1.833	1.604	1.549	1.833	1.851
Lake Houston	2.302	2.173	2.155	1.872	1.828	2.105	2.239
<i>n = 100</i>							
Lake Conroe	2.009	1.887	1.942	1.707	1.623	1.882	1.967
Lake Somerville	1.230	1.218	1.216	1.115	1.095	1.212	1.222
Lake Houston	1.471	1.377	1.394	1.228	1.180	1.349	1.433
Full sample							
Lake Conroe	0.882	0.849	0.866	0.850	0.829	0.851	0.866
Lake Somerville	0.554	0.548	0.554	0.548	0.538	0.548	0.554
Lake Houston	0.460	0.426	0.449	0.428	0.415	0.429	0.449

Table 10 Standard errors for the restricted estimator

	ML	D1	D2	D3	D4	D5	D6
<i>n</i> = 50							
Lake Conroe	0.890	0.833	0.839	0.760	0.693	0.819	0.859
Lake Somerville	1.304	1.281	1.263	1.182	1.108	1.262	1.283
Lake Houston	1.187	1.195	1.170	1.106	1.034	1.184	1.176
<i>n</i> = 100							
Lake Conroe	0.691	0.658	0.668	0.639	0.613	0.658	0.673
Lake Somerville	1.001	0.991	0.987	0.937	0.906	0.983	0.994
Lake Houston	0.848	0.859	0.850	0.793	0.765	0.853	0.849
Full sample							
Lake Conroe	0.421	0.417	0.416	0.425	0.403	0.419	0.416
Lake Somerville	0.523	0.522	0.522	0.518	0.498	0.521	0.523
Lake Houston	0.373	0.370	0.374	0.355	0.337	0.367	0.375

when we observe the results of the simulation study, it is the restricted Liu estimator that has the lowest standard errors. It is always the estimator D4 that minimizes the bootstrapped standard errors.

5 Summary and Concluding Remarks

This paper introduces the unrestricted and restricted Liu estimators for the Poisson regression model. We also introduced some new estimators of d based on different percentiles of existing estimators. Using simulation techniques, we showed that the restricted estimator outperformed the unrestricted one. We also found that the restricted Liu type estimator was superior to the restricted Poisson estimator in the presence of multicollinearity. Our study also shows that the new estimator D4 always has the lowest MSE. Finally, we illustrated the benefit of the new method using an empirical application where the demand for recreational boat trips was studied and once again we found that the proposed D4 estimator performed the best and may be recommended for practitioners.

Funding Open access funding provided by Jönköping University.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendix

See Tables 11, 12 and 13.

Table 11 Proportion of nonzero d for both restricted and unrestricted Poisson regression model for $p = 4$

	D1	D2	D3	D4	D5	D6
<i>df</i>	$\rho = 0.80$					
20	0.05	0.33	0.10	0.03	0.12	0.37
30	0.02	0.31	0.04	0.02	0.04	0.31
50	0.02	0.20	0.02	0.02	0.02	0.20
100	0.01	0.04	0.03	0.01	0.04	0.04
<i>df</i>	$\rho = 0.90$					
20	0.05	0.20	0.08	0.05	0.05	0.24
30	0.03	0.16	0.03	0.02	0.03	0.16
50	0.02	0.08	0.02	0.02	0.02	0.08
100	0.01	0.04	0.03	0.01	0.04	0.04
<i>df</i>	$\rho = 0.95$					
20	0.04	0.12	0.04	0.04	0.04	0.15
30	0.03	0.07	0.03	0.02	0.03	0.07
50	0.02	0.04	0.02	0.02	0.02	0.04
100	0.01	0.02	0.03	0.01	0.04	0.02

Table 12 Proportion of nonzero d for both restricted and unrestricted Poisson regression model for $p = 8$

	D1	D2	D3	D4	D5	D6
<i>df</i>	$\rho = 0.80$					
20	0.04	0.20	0.06	0.03	0.07	0.22
30	0.02	0.17	0.03	0.02	0.03	0.17
50	0.02	0.11	0.02	0.02	0.02	0.11
100	0.01	0.03	0.03	0.01	0.04	0.03
<i>df</i>	$\rho = 0.90$					
20	0.03	0.13	0.05	0.02	0.04	0.15
30	0.02	0.09	0.02	0.02	0.03	0.09
50	0.02	0.05	0.02	0.02	0.02	0.05
100	0.01	0.03	0.03	0.01	0.04	0.03
<i>df</i>	$\rho = 0.95$					
20	0.02	0.03	0.02	0.02	0.02	0.03
30	0.02	0.02	0.02	0.02	0.02	0.02
50	0.02	0.02	0.02	0.02	0.02	0.02
100	0.01	0.02	0.03	0.01	0.04	0.02

Table 13 Proportion of nonzero d for both restricted and unrestricted Poisson regression model for $p = 12$

	D1	D2	D3	D4	D5	D6
df	$\rho = 0.80$					
20	0.01	0.08	0.01	0.01	0.01	0.08
30	0.01	0.04	0.01	0.01	0.01	0.04
50	0.00	0.00	0.00	0.00	0.00	0.00
100	0.00	0.00	0.00	0.00	0.00	0.00
df	$\rho = 0.90$					
20	0.00	0.06	0.00	0.00	0.00	0.06
30	0.00	0.02	0.00	0.00	0.00	0.02
50	0.00	0.00	0.00	0.00	0.00	0.00
100	0.00	0.00	0.00	0.00	0.00	0.00
df	$\rho = 0.95$					
20	0.00	0.04	0.00	0.00	0.00	0.04
30	0.00	0.01	0.00	0.00	0.00	0.01
50	0.00	0.00	0.00	0.00	0.00	0.00
100	0.00	0.00	0.00	0.00	0.00	0.00

References

- Akdeniz, F., & Kaciranlar, S. (1995). On the almost unbiased generalized Liu estimator and unbiased estimation of the bias and MSE. *Communications in Statistics—Theory and Methods*, 24, 1789–1797.
- Algarnal, Z. Y. (2018). Biased estimators in Poisson regression model in the presence of multicollinearity: A subject review. *Al-Qadisiyah Journal for Administrative and Economic Sciences*, 20(1), 37–43.
- Alheety, M. I., & Kibria, B. M. G. (2009). On the Liu and almost unbiased Liu estimators in the presence of multicollinearity with heteroscedastic or correlated errors. *Surveys in Mathematics and its Applications*, 4, 155–167.
- Asar, Y. (2018). Liu-type negative binomial regression: A comparison of recent estimators and applications. In Tez, M., & Von Rose, D. (Eds.) *Trends and Perspectives in Linear Statistical Inference* (pp. 23–39). Springer, Cham.
- Asar, Y., & Genç, A. (2017). A new two-parameter estimator for the poisson regression model. *Iranian Journal of Science and Technology, Transactions A: Science*, 1–11.
- Cameron, A. C., & Trivedi, P. K. (2013). *Regression Analysis of Count Data*, 2nd edn, Econometric Society Monograph No.53, Cambridge University Press, 1998.
- Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: biased estimation for non-orthogonal Problems. *Technometrics*, 12, 55–67.
- Kaciranlar, S. (2003). Liu estimator in the general linear regression model. *Journal of Applied Statistical Science*, 13, 229–234.
- Kaçiranlar, S., & Dawoud, I. (2018). On the performance of the poisson and the negative binomial ridge predictors. *Communications in Statistics—Simulation and Computation*, 47, 1751–1770.
- Kaçiranlar, S., Sakallioğlu, S., Akdeniz, F., Styan, G. P., & Werner, H. J. (1999). A new biased estimator in linear regression and a detailed analysis of the widely-analysed dataset on Portland cement. *Sankhyā: The Indian Journal of Statistics, Series B*, 443–459.
- Kibria, B. M. G. (2003). Performance of some new ridge regression estimators. *Communications in Statistics—Theory and Methods*, 32, 419–435.
- Kibria, B. M. G., & Saleh, A. M. E. (2012). Improving the estimators of the parameters of a probit regression model: A ridge regression approach. *Journal of Statistical Planning and Inference*, 142, 1421–1435.

- Li, Y., & Yang, H. (2010). A new stochastic mixed ridge estimator in linear regression model. *Statistical Papers*, 51, 315–323.
- Liu, K. (1993). A new class of biased estimate in linear regression. *Communications in Statistics-Theory and Methods*, 22, 393–402.
- Månsson, K. (2013). Developing a Liu estimator for the negative binomial regression model. *Journal of Statistical Computation and Simulation*, 83, 1773–1780.
- Månsson, K., Kibria, B. M. G., & Shukur, G. (2016). A restricted Liu estimator for binary regression models and its application to an applied demand system. *Journal of Applied Statistics*, 43, 1119–1127.
- Månsson, K., Kibria, B. M. G., Sjölander, P., & Shukur, G. (2012). Improved Liu estimators for the Poisson regression model. *International Journal of Statistics and Probability*, 1(1), 2.
- Ehsanes Saleh, A. M., & Golam Kibria, B. M. (1993). Performances of some new preliminary test ridge regression estimators and their properties. *Communications in Statistics-Theory and Methods*, 22, 2747–2764.
- Türkan, S., & Özel, G. (2016). A new modified Jackknifed estimator for the Poisson regression model. *Journal of Applied Statistics*, 43, 1892–1905.
- Türkan, S., & Özel, G. (2018). A Jackknifed estimators for the negative binomial regression model. *Communications in Statistics-Simulation and Computation*, 47(6), 1845–1865.
- Varathan, N., & Wijekoon, P. (2018). Liu-Type logistic estimator under Stochastic Linear Restrictions. *Ceylon Journal of Science*, 47(1), 21–34.
- Wu, J., & Asar, Y. (2017). More on the restricted Liu estimator in the logistic regression model. *Communications in Statistics-Simulation and Computation*, 46(5), 3680–3689.
- Yang, H., & Xu, J. (2009). An alternative stochastic restricted Liu estimator in linear regression. *Statistical Papers*, 50, 639–647.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.