

Comparison between the modified two-parameter Liu estimator and the edge regression estimator for the Poisson regression model under the multicollinearity problem with application

Ali Khaled Mohammed Jassim 

A.P.Suhail Najim Abood 

Email: Ali.Khaled2201@coadec.uobaghdad.edu.iq

Email: suhnaj2005@coadec.uobaghdad.edu.iq

Administration and Economics / University of Baghdad

<https://doi.org/10.34009/aujeas.2026.164882.1186>

ABSTRACT

Cite as:

Mohammed, A. K., & Abood, S. N. (2026). Comparison between the modified two-parameter Liu estimator and the edge regression estimator for the Poisson regression model under the multicollinearity problem with application. *AL-Anbar University journal of Economic and Administration Sciences*, 18(2)



©Authors, 2026, College of Administration and economics, University of Anbar. This is an open-access article under the CC BY 4.0 license

(<http://creativecommons.org/licenses/by/4.0>)

Received: 2025 / 09 / 07

Accepted: 2025 / 09 / 25

Published: 2026 / 06 / 30

Print ISSN:1998-8141

Online ISSN:2706-6010

The Poisson regression model is one of the most widely used statistical techniques for analyzing count data. However, this model often suffers from multicollinearity when strong correlations exist among explanatory variables, which inflates variances and reduces the efficiency of the maximum likelihood estimator (MLE). To overcome this issue, several biased estimators have been proposed as alternatives, most notably the two-parameter modified Liu estimator and the ridge regression estimator. This study provides a comprehensive comparison between the two estimators in the context of the Poisson regression model. Theoretical properties were investigated in terms of bias and variance, and an empirical application using real/simulated data was conducted to evaluate their performance. Various levels of multicollinearity and sample sizes were considered to assess the robustness and efficiency of the estimators. The findings indicate that both estimators outperform the MLE under multicollinearity, yet their relative efficiency depends on the severity of correlation among covariates and the sample size. The two-parameter modified Liu estimator demonstrated greater flexibility due to the presence of two shrinkage parameters, allowing for a better trade-off between bias and variance. In contrast, the ridge estimator proved simpler to implement and showed competitive performance under moderate levels of multicollinearity. This research contributes to the literature on biased estimators for count data regression models and offers researchers practical alternatives for addressing multicollinearity within the framework of Poisson regression. Data analysis and model implementation were conducted using Microsoft Excel and MATLAB.

Keywords: Poisson regression, multicollinearity, modified Liu estimator, ridge regression estimator, biased estimators, regularized regression.

Keywords:

modified Liu estimator, ridge regression estimator, biased estimators, regularized regression.

مقارنة بين مقدر ليو المعدل بمعلمتين ومقدر انحدار الحرف لأنموذج انحدار بواسون في ظل وجود مشكلة التعدد الخطي مع التطبيق

الم.د. سهيل نجم عبود 

الباحث. علي خالد محمد جاسم 

الايميل : suhnaj2005@coadec.uobaghdad.edu.iq الايميل : Ali.Khaled2201@coadec.uobaghdad.edu.iq

جامعة بغداد / كلية الادارة والاقتصاد

المستخلص :

يعد أنموذج انحدار بواسون من النماذج الإحصائية الشائعة الاستخدام لتحليل البيانات العدّية، إلا أن هذا الأنموذج قد يواجه مشكلة التعدد الخطي (Multicollinearity) عند وجود ارتباطات قوية بين المتغيرات التوضيحية، مما يؤدي إلى تضخم التباين وضعف كفاءة المقدرات، وللتغلب على هذه المشكلة طُرحت العديد من المقدرات البديلة لمقدر الاحتمال الأعظم التقليدي (MLE)، من أبرزها مقدر ليو المعدل بمعلمتين (Two-parameter Modified Liu Estimator) و مقدر انحدار الحرف (Ridge Regression Estimator)، في هذا البحث تمت دراسة مقارنة شاملة بين المقدرين ضمن سياق أنموذج انحدار بواسون، وذلك من خلال تحليل الخصائص النظرية لهما من حيث التحيز والتباين، بالإضافة إلى إجراء تطبيق عملي على بيانات حقيقية/محاكاة للتأكد من فعاليتهما، تم استخدام عدة مستويات مختلفة من التعدد الخطي وأحجام عينات متنوعة لقياس الأداء الإحصائي للمقدين، أظهرت النتائج أن المقدرين يحققان تحسناً ملحوظاً مقارنة بمقدر الاحتمال الأعظم في ظل التعدد الخطي، إلا أن كفاءة كل مقدر تختلف باختلاف شدة الارتباط بين المتغيرات وعدد الملاحظات، حيث تميز مقدر ليو المعدل بمرونة أكبر نتيجة وجود معاملين للضبط (shrinkage parameters) مما يمنحه القدرة على موازنة أفضل بين التحيز والتباين، بينما كان مقدر انحدار الحرف أكثر بساطة في التطبيق ويؤدي أداءً جيداً في المستويات المعتدلة من التعدد الخطي، يساهم هذا البحث في إغناء الأدبيات الخاصة بمقدرات نماذج الانحدار للبيانات العدّية، ويقدم للباحثين والمهتمين بالتحليل الإحصائي خيارين بديلين فعالين لمعالجة مشكلة التعدد الخطي ضمن إطار انحدار بواسون، تم تحليل البيانات وتنفيذ النمذجة باستخدام برنامجي Microsoft Excel و MATLAB.

الكلمات المفتاحية: انحدار بواسون، التعدد الخطي، مقدر ليو المعدل، مقدر انحدار الحرف، المقدرات المنحازة، الانحدار المنتظم.

First, The Research Methodology

(Introduction)

Count data frequently arise in many scientific disciplines such as medicine, biology, economics, engineering, and the social sciences. A common statistical model for analyzing such data is the Poisson regression model, which relates the mean of a count response variable to a set of explanatory variables through a log-link function. Due to its simplicity and interpretability, Poisson regression has become one of the most widely applied models for studying count-type responses. Despite its usefulness, the maximum likelihood estimator (MLE) employed in Poisson regression is highly sensitive to the problem of multicollinearity, which occurs when the explanatory variables are highly correlated. Multicollinearity leads to inflated variances of the estimated coefficients, unstable estimates, and weak predictive ability. This issue becomes more severe when the sample size is small or when the correlations among predictors are particularly strong. Therefore, addressing multicollinearity is crucial for improving the reliability of parameter estimation in Poisson regression models. To mitigate this problem, several biased estimators have been developed as alternatives to the MLE. Two notable examples are the ridge regression estimator and the two-parameter modified Liu estimator. Ridge regression introduces a penalty term that shrinks the regression coefficients toward zero, thereby reducing variance while allowing a certain degree of bias. The two-parameter modified Liu estimator, on the other hand, extends the original Liu approach by incorporating two shrinkage parameters, offering greater flexibility in controlling the balance between bias and variance.

Although both estimators have been studied extensively in the context of linear regression, their comparative performance in Poisson regression under multicollinearity has not been sufficiently addressed. This gap motivates the present study, which aims to provide both a theoretical and an empirical comparison between the two estimators. By examining their properties under different levels of multicollinearity and across various sample sizes, this study seeks to identify the conditions under which each estimator is more efficient. The findings of this research are expected to enrich the literature on biased estimation in count regression models and to provide applied researchers with practical guidance when dealing with collinear count data.

The Problem of Research

Forecasting mortality rates accurately remains a critical challenge for demographers, policymakers, and actuaries. Traditional methods often fail to capture dynamic changes in mortality trends caused by medical advances, socioeconomic factors, and public health interventions. In this context, the need arises for a robust and statistically reliable model that can predict future mortality rates with reasonable accuracy. The problem this study addresses is how to effectively apply the Lee-Carter model to estimate and forecast mortality rates for the period 2024–2026, while ensuring the model's assumptions and components align with real-world demographic data.

Research Goal

The goal of this research is to compare the performance of the two-parameter modified Liu estimator and the ridge regression estimator within the Poisson regression model in the presence of multicollinearity. Specifically, the study aims to: Examine the theoretical properties of both estimators in terms of bias and variance, Conduct simulation experiments under varying levels of multicollinearity and different sample sizes, Apply the estimators to a real dataset to evaluate their practical performance, Provide recommendations on the conditions under which each estimator performs better, thereby offering applied researchers practical tools for analyzing collinear count data.

Data Description and Analysis:

The empirical part of this study relies on a dataset consisting of 100 observations with one dependent variable and five explanatory variables. The response variable represents the number of visits to a healthcare facility, making the Poisson regression model an appropriate analytical framework. The variables are described as follows:

Dependent Variable (Y): Y_Visits: Number of healthcare visits for each individual.

Explanatory Variables (X):

X1_Age: Age of the individual (in years).

X2_BMI: Body Mass Index, representing weight-to-height ratio.

X3_Chronic: Presence of chronic conditions (e.g., diabetes, hypertension).

X4_Medications: Number of prescribed medications.

X5_Income: Monthly household income level.

Preliminary Analysis

1. Descriptive Statistics:

Summary measures such as mean, standard deviation, minimum, and maximum values were computed for all variables. This provided an initial understanding of the distribution of the data and potential variability across individuals.

2.Multicollinearity Diagnosis:

The correlation matrix revealed moderate to strong correlations among some of the explanatory variables, especially between Age (X1), Chronic conditions (X3), and Medications (X4). The Variance Inflation Factor (VIF) values also confirmed the existence of multicollinearity, with several predictors exceeding the commonly used threshold of 10.

3.Estimation Methods:

The dataset was analyzed using three approaches:

Maximum Likelihood Estimator (MLE) as the benchmark.

Ridge Regression Estimator to introduce shrinkage and reduce multicollinearity effects.

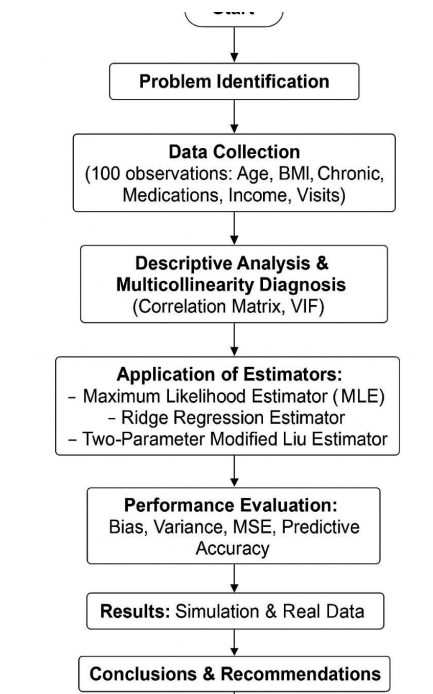
Two-Parameter Modified Liu Estimator for additional flexibility through dual shrinkage parameters.

4.Evaluation Criteria:

The estimators were compared using mean squared error (MSE), bias, and predictive performance. Both simulation-based results and empirical findings from this dataset were reported.

This analysis demonstrates how multicollinearity can affect Poisson regression and highlights the potential of biased estimators to improve efficiency and stability of the results.

Research methodology illustration



Reference Review

Multicollinearity in regression models can reduce the efficiency and stability of traditional estimators like the MLE. Biased estimation methods, such as ridge regression and Liu-type estimators, have been developed to address this issue. While widely studied in linear regression, their application to Poisson regression models has received increasing attention. Previous studies highlight the effectiveness of these estimators in improving accuracy and reducing variance under multicollinearity. This section briefly reviews relevant research on theoretical properties, simulations, and practical applications that inform the current study.

In the year (2023), the researcher (Ramadan) presented a research entitled “A comparison between the letter parameters in the modified two-parameter estimation of the negative binomial regression model with a practical application.” This study aimed to find the best two-parameter estimator to solve the problem of multicollinearity in the negative binomial regression model by combining several combinations of both the letter and Liu parameters into a single estimator. The results obtained showed the superiority of this estimator over classical estimation methods.

In 2024, researcher Ismail presented a research entitled "Employing some shrinkage functions to estimate the parameters of the beta regression model in the presence of the multicollinearity problem with application." This study aimed to compare the parameters of beta regression, the maximum likelihood method, the edge regression method, the Liu method, the modified Liu method, and the method of researchers (Ozkale and Kaciranlar). Through the results of Monte Carlo simulations and relying on the MSE comparison criterion, it was shown that the edge regression estimator outperformed the other methods.

In 2024, researcher Abdul Redha presented a research entitled "Comparison of some methods for estimating the parameters of the Conway-Maxwell-Poisson regression model in the presence of the multicollinearity problem with application." This study aimed to compare the parameters of the Conway-Maxwell-Poisson regression in the presence of the multicollinearity problem, namely the maximum likelihood estimator method, the edge regression method, the Liu estimator method, the modified Liu estimator method, and the Liu method. Two parameters and through the results reached by the researcher and based on the MSE comparison criterion, it was shown that the Liu estimator with two shrinkage parameters is superior.

Second, the theoretical framework

Definition of poisson distribution

(Al-Quraishi et al., 2024: 329-338) The Poisson distribution is a probability distribution named after the French mathematician Simeon Poisson in the year (1873). It is considered one of the important discrete probability distributions that express the number of rare events that appear in a certain period of time or a specific area. The period of time can be a second, minute, hour, day, week, month or year. It represents an approximate case of the negative binomial distribution. Its derivative was published in the year (1837), and its name was given to this distribution. Let us assume that (xi) is a random variable that indicates the number of times a certain event occurs during the period of time. Then (xi) follows the Poisson distribution with its parameter (λ), and it is written in short:

$$Pr(Y/\theta) = \frac{e^{-\lambda} \lambda^Y}{Y!} \quad Y_{i=0,1,2,\dots} \quad (2-1)$$

$$= 0 \quad \text{otherwise}$$

:Where

λ represents the distribution parameter and has a value greater than zero ($\lambda > 0$).

In scientific life, we sometimes encounter some phenomena that fit the binomial distribution. However, these occurrences are rare. This means that the probability of success is very small, approaching zero. Therefore, it can be said that $np = \lambda$, where λ is a constant. Therefore, the probability of failure is large, i.e., it approaches one.

Conditions of the Poisson distribution:

The probability of success (p) is constant, as is the probability of failure (q) in each trial.

The probability of success is small and close to zero, and the probability of failure is close to one.

The number of trials is very large, such that $\lambda = np$ is a constant.

The rate of occurrence is constant, meaning that the rate does not change over time.

Properties of the Poisson distribution:

(Bakhit et al., 2006: 338) & (Kazem et al., 2021: 15)

The Poisson distribution has only one parameter, λ .

The distinguishing property of the Poisson distribution from other discrete distributions is that its mean is equal to its variance and is equal to the parameter value. This property is known as "equal dispersion." This property is restrictive of the Poisson distribution because the variance of numerical variables is often greater than its mean. This property is called "overdispersion." In some cases, the variance may be greater than the mean, indicating overdispersion. In such cases, alternative distributions, such as the negative binomial distribution, can be used. In other cases, the variance is less than the mean. This property is called "underdispersion." This is rare in real data, but it occurs in complex data or data subject to certain constraints. In such cases, another distribution, such as the modified Poisson distribution, can be used.

This distribution tends toward normal distribution if $\lambda \rightarrow \infty$.

The Poisson distribution has a characteristic of being skewed to the right (i.e., positive skewness).

Poisson Regression Model:

(Hameed et al., 2023: 81-89) The Poisson regression model is one of the most important log-linear regression models. It is a form of regression analysis suitable for modeling data when the dependent variable data is discrete and follows a Poisson distribution. This model assumes that the dependent variable (Y_i) is a response variable that follows a Poisson distribution with a parameter of (λ). The random errors in the model also follow a Poisson distribution with a parameter of (λ). This can be expressed using the following formula:

$$Y = e^{X\beta + U} \quad (1)$$

Where:

Y: The vector of the response variable with degree (nx1).

X: The matrix of explanatory variables with degree (nx(px1)).

β : The vector of parameters with degree (p+1)x1).

U: The vector of random errors with degree (nx1).

n: The sample size.

p: The number of explanatory variables. The Poisson regression model is used when the response variable (Y) is in the form of count data and when the response variable (Y) is in the form of equations (rate data).

Multicollinearity Problem

(Al-Tamimi et al., 2014: 239-240) & (Ismail, 2024: 19) The multicollinearity problem arises when there is a linear correlation between two or more explanatory variables. This makes it difficult to separate the effect of these variables on the response variable in practical applications. Multicollinearity also occurs when the value of one of the independent variables is the same for all observations, or when the value of one of the independent variables depends on the value of one or more of the other independent variables in the model. It also occurs when the data is in the form of cross-sectional data or time series data, and the number of parameters to be estimated must be less than the sample size under investigation.

$$\text{Rank}(x) = k+1 < n \quad (2)$$

n: Sample size.

K: Number of explanatory variables in the equation.

K+1: Number of parameters in the equation, including the fixed section.

In the absence of multicollinearity, the explanatory variables are called orthogonal variables, meaning that each explanatory variable affects the dependent variable in a completely separate way. Therefore, the estimation results in multiple regression are completely identical to the estimation results in simple regression, where the Fisher matrix ($X'X$) is

diagonal because the correlation coefficient between the explanatory variables is zero ($r_{X_i X_j} = 0$) or the common correlation between the explanatory variables is zero.

$$\sum_{t=1}^n (x_{ti} - X_i)(x_{tj} - X_j) = 0 \quad \forall i \neq j = 1, 2, 3, \dots, k \quad (3)$$

There are two types of multicollinearity problems, which often occur at varying rates: perfect multicollinearity and semi-perfect multicollinearity.

First: Perfect multicollinearity: This means that The value of the determinant of the information matrix is equal to zero $|X'X|=0$, which makes it difficult to estimate the model parameters using the ordinary least squares (OLS) method. Therefore, the inverse of the information matrix $(X'X)^{-1}$ cannot be calculated, and OLS collapses under the weight of perfect multicollinearity.

Second: Near-perfect multicollinearity: In this case, the value of the determinant of the information matrix is very small, i.e., close to zero. This leads to inaccurate estimation of the model parameters using the ordinary least squares (OLS) method, and the estimated parameters have very large variance. The most prominent method for dealing with this situation is to use the ridge regression method.

Effects of Multicollinearity

(Abdul-Ridha, 2024:20) Multicollinearity does not affect regression estimates, but it does make them inaccurate and unreliable. Therefore, it may be difficult to determine how the independent variables individually affect the dependent variable. This sometimes inflates the standard errors of the regression coefficients, increasing the likelihood of errors.

Ridge Regression Estimators Method:

The Ridge Regression Estimators method is an alternative estimation method when there is a problem of multicollinearity between explanatory variables. This method is constructed using the same method adopted by (Hoerl & Kennard, 1970 : 55-67) which is to add a small positive quantity to the elements of the main diagonal of the information matrix $(X'X)$. The difference between the linear model and the Poisson regression model is that it is a type of linear logarithmic model (Amin et al., 2021: 33)

This method is an alternative estimation method when there is a problem of multicollinearity between explanatory variables in a Poisson regression model (PRM), which causes a problem with the maximum likelihood method (MLE). To address this problem, (Qasim & Månsson, 2020, 1775-1789) tested another estimator, called β_{RR} . The weighted sum of squared errors (WSSE) can be written as follows:

$$u'u = (Y - X\beta_{ML})'(Y - X\beta_{ML}) + (\beta_{RR} - \beta_{ML})'(X'\hat{W}X)(\beta_{RR} - \beta_{ML}) = \phi_{min} + \phi(\beta_{RR}) \quad (4)$$

represents the increase in the weighted mean squared errors.

To find the Poisson regression estimator, we must minimize the length of $(\beta'_{RR}\beta_{RR})$ according to the constraint $\phi_0 = \phi(\beta_{RR})$ This can be formulated as follows:

$$\text{Minimize } F = \beta'_{RR}\beta_{RR} + \left(\frac{1}{K}\right)(\beta_{RR} - \beta_{ML})'(X\hat{W}X)(\beta_{RR} - \beta_{ML} - \phi_0) \quad (5)$$

represents the Lagrange multiplier.

Thus, the weighted sum of squared errors will be:

$$U'U = (Y - X\beta_{ML})'(Y - X\beta_{ML}) + \beta'_{RR}\beta_{RR} + \left(\frac{1}{K}\right)(\beta_{RR} - \beta_{ML})'(X\hat{W}X)(\beta_{RR} - \beta_{ML} - \phi_0) \quad (6)$$

By differentiating the formula (6) for (β_{RR}) and setting the derivative equal to zero, we can find the slope estimator for the Poisson regression model as follows:

$$\hat{\beta}_{RR} = (X\hat{W}X + KI)^{-1}(X\hat{W}X) \hat{\beta}_{ML} \quad (7)$$

$$\hat{\beta}_{RR} = Z\hat{\beta}_{ML} \quad (8)$$

Since:

$$Z = (X\hat{W}X + KI)^{-1}(X\hat{W}X) \quad (9)$$

The regression estimators are biased when $(K > 0)$ and the amount of bias is: (Mansson et al, 2012:1(1)2)

$$\text{Bias}(\hat{\beta}_{RR}) = -kQ \Lambda_K^{-1}\beta \quad (10)$$

$$\begin{aligned} \text{COV}(\hat{\beta}_{RR}) &= E[(\hat{\beta}_{RR} - E(\hat{\beta}_{RR}))(\hat{\beta}_{RR} - E(\hat{\beta}_{RR}))'] \\ &= Q\Lambda_K^{-1}\Lambda\Lambda_K^{-1}Q' \quad (11) \end{aligned}$$

$$\text{MMSE}(\hat{\beta}_{RR}) = W(X\hat{W}X)W' + \text{Bias}(\hat{\beta}_{RR})\text{Bias}(\hat{\beta}_{RR})'$$

$$= Q\Lambda_K^{-1}\Lambda\Lambda_K^{-1}Q' + K^2Q\Lambda_K^{-1}\beta\beta'\Lambda_K^{-1}Q' \quad (12)$$

Where

$$\Lambda_k = \text{diag}(\lambda_{1+k}, \lambda_{2+k}, \dots, \lambda_{r+k})$$

$$\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_r) = Q(X\hat{W}X)Q'$$

Q: is an orthogonal matrix whose columns are eigenvectors of $X\hat{W}X$

Thus, we obtain The mean squared error (MSE) of the Poisson regression model parameters estimated using the letter regression method by applying tr to the formula ()

$$\text{MSE}(\hat{\beta}_{RR}) = \text{tr}(\text{MMSE}(\hat{\beta}_{RR}))$$

$$= \sum_{j=1}^r \frac{\lambda_j}{(\lambda_j + k)^2} + k^2 \sum_{j=1}^r \frac{\alpha_j^2}{(\lambda_j + k)^2} \quad (13)$$

$\alpha = Q' \beta$, which is the letter parameter proposed by (Hoerl & Kennard), and for PRE, it represents a parameter of the Poisson regression model when taking the natural logarithm of the model in the formula (13).

Ridge Parameter Estimators:

There is no specific rule for estimating the k-ridge parameter. However, several methods have been proposed for the ridge regression model.

The first estimator proposed by (Hoerl & Kennard, 1970 : 55-67) :

$$k_1 = \frac{\hat{\sigma}_u^2}{\hat{\alpha}_{max}^2} \quad (14)$$

$\hat{\alpha}_{max}^2$:The largest element in the formula (14)

$\hat{\sigma}_u^2$: represents the estimated mean squared error, which is calculated according to the following formula:

$$\hat{\sigma}_u^2 = \frac{\sum_{i=1}^n (y_i - \hat{\lambda}_i)^2}{n - p - 1}$$

The second estimator proposed by Schaef:

$$k_2 = \frac{1}{2\hat{\alpha}_{max}^2} \quad (15)$$

Several estimators were subsequently proposed, including

$$k_3 = \left(\frac{(\lambda_j)(1-(\hat{\alpha}_j^2))}{1+(\lambda_j)mean(\hat{\alpha}_j^2)} \right) - \hat{d}_2 \quad (16)$$

Modified Liu estimator method for a two-parameter Poisson regression model

(Abdul Wahab et al, 2024: 46). proposed a two-parameter modified Liu estimator for a Poisson regression model. This modified estimator, referred to as PMTPLE, was constructed using a pair of parameters (d, K). The formula for this model is arranged as follows

$$\hat{\beta}_{PMTPL} = (X'WX)^{-1}(X'WX - (K^* + d_0^*)I) \hat{\beta}_{MLE} \quad K^* > 0, \quad 0 < d_0^* < 1 \quad (17)$$

The bias vector, covariance matrix, mean square error matrix, and mean square error matrix :are given as follows

$$Bias(\hat{\beta}_{PMTPL}) = -(K^* + d_0^* + 1)Q(\Lambda + I)^{-1}\beta \quad (18)$$

$$COV(\hat{\beta}_{PMTPL}) = Q(\Lambda + I)^{-1}(\Lambda -$$

$$(K^* + d_0^*)I)\Lambda^{-1}(\Lambda + I)^{-1}(\Lambda - (K^* + d_0^*)I)\Lambda^{-1}$$

$$MMSE(\hat{\beta}_{PMTPL}) = COV(\hat{\beta}_{PMTPL}) + Bias(\hat{\beta}_{PMTPL})Bias(\hat{\beta}_{PMTPL})^{-1}$$

$$= Q(\Lambda + I)^{-1}(\Lambda -$$

$$(K^* + d_0^*)I)\Lambda^{-1}(\Lambda + I)^{-1}(\Lambda - (K^* + d_0^*)I)\Lambda^{-1} + (K^* + d_0^* + 1)^2Q(\Lambda + I)^{-1}\beta\beta'(\Lambda + I)^{-1}Q'$$

$$MSE(\hat{\beta}_{PMTPL}) = trace \left(MMSE(\hat{\beta}_{PMTPL}) \right)$$

$$= \sum_{j=1}^r \frac{(\lambda_j - (K^* + d_0^*))^2}{\lambda_j(\lambda_j + 1)^2} + (K^* + d_0^* + 1)^2 \sum_{j=1}^r \frac{\alpha_j^2}{(\lambda_j + 1)^2} \quad (19)$$

The values of d_0^* and k_0^* are as follows:

$$k_0^* = \left(\frac{\lambda_j(1 - \alpha_j^2)}{1 + \lambda_j \alpha_j^2} \right) - d_0^* \quad (20)$$

$$d_0^* = \frac{\lambda_j(1 - \alpha_j^2) - k(1 + \lambda_j \alpha_j^2)}{1 + \lambda_j \alpha_j^2} \quad (21)$$

Selection of Biasing parameter test

Researchers (Abdul Wahab et al, 2024: 46) suggest using the estimator (d_0^*) for the d_0^* parameter in (PMTPLE). This can be achieved by incorporating the value of (d_0^*) into the formula (22) as shown below

$$\hat{d}_0^* = \left(0, \left(\frac{\lambda_j(1 - \hat{\alpha}_j^2) - \hat{k}(1 + \lambda_j \hat{\alpha}_j^2)}{1 + \lambda_j \hat{\alpha}_j^2} \right) \right) \quad (22)$$

:When

$$\hat{k} = \left(\frac{1}{\hat{\alpha}_j^2} \right) \quad (23)$$

For the parameter K in the two-parameter Liu estimator, we suggest using the estimators in the formulas found in the letter parameter estimator

$$\hat{k}_1^* = \hat{k}_1$$

$$\hat{k}_2^* = \hat{k}_2$$

$$\hat{k}_3^* = \hat{k}_3$$

Third, The Practical Framework

Introduction:

-An application about patient visits to the hospital in Iraq

Hospital visits are the number of times a person visits a healthcare facility, such as a hospital, health center, or affiliated clinic, to receive medical services. These visits include therapeutic, diagnostic, and preventive visits, as well as periodic checkups to monitor chronic diseases. Visits are divided into several types, including planned visits, emergency visits, follow-up visits, and preventive visits. Studying the number of visits is an important indicator of community health, the efficiency of the healthcare system, and the level of health awareness among individuals. In Iraq, these visits are affected by several factors, including the availability of medical services, the economic situation, the level of healthcare infrastructure, the nature of the disease, and the patient's geographical location.

Government hospitals are relatively free but suffer from overcrowding and a lack of resources. Private hospitals provide faster services but at a higher cost. Analyzing the number of visits helps plan the distribution of medical staff, provide supplies, and improve the quality of healthcare services provided to citizens. (Kamal, 1990: 599-649), (Abdullah (2024:265-273) , (Al-(Rashidani ,2006: 48-51)

The scientific, health, and economic dimensions of hospital visits and their role in health planning:

1. Research significance

From a scientific perspective, the rate of hospital visits is considered an important dependent variable in epidemiological and health economic studies, as it reflects the interaction between individual factors such as age, gender, and health status, and systemic factors such as the availability and quality of health services. (Sabry, 2013: 164-188) Recent studies indicate that a high rate of visits may be indicative of deficiencies in primary care, or the presence of complex chronic diseases that require ongoing medical follow-up. (Khorshid, 2018: 466-488).

2. Health determinants of hospital visits

- Demographic factors: Age and gender are among the most important influencing factors, as the rate of visits increases among the elderly due to the accumulation of health problems.
- Pathological factors: The presence of multiple chronic diseases (multimorbidity) leads to an increase in visits, as confirmed by numerous studies, including a study that showed that patients with three or more chronic diseases had a visit rate three times higher than those without.
- Polypharmacy: This indicator reflects the severity and complexity of the medical condition. A study showed that patients who take more than five medications per month are 40% more likely to visit the hospital than others.
- Socioeconomic factors: Income level, education, and insurance coverage directly influence the pattern of utilization of health services (Sabry, 2013:164-188).

3. The economic dimensions of hospital visits

The costs of hospital visits represent a significant portion of total health expenditure in most countries. Data from the World Health Organization (Ramadan, 2023:640-663), indicate that unnecessary hospital visits constitute a financial burden that can be reduced through improved primary care and prevention programs. Readmission rates are also used as an indicator of the quality of care, as they reflect, in some cases, a lack of follow-up or a weak treatment plan (Ismail, 2024:19).

Data Description:

The data used in this section pertain to patient visits to the hospital in Iraq, which were obtained from the Iraqi Ministry of Health. The values of the variables were recorded over a year, with a sample size of (100). The Poisson regression model included: the explanatory variables, especially between Age (X1), Chronic conditions (X3), Medications (X4) X5: and Average monthly income

The dependent variable:

Y: Number of patient visits to the hospital during the year.

Explanatory variables include:

X1: Patient's age.

X2: Body mass index.

X3: Number of chronic diseases.

X4: Number of medications prescribed to the patient per month.

X5: Average monthly income.

Simulation

$P = 5$ and $\sigma^2 = 0.3$.

N		100			200		
P		0.9	0.95	0.99	0.9	0.95	0.99
$\hat{\beta}_k$	K1	0.000495	0.0017597	0.0011864	0.0011953	0.0079366	0.0035319
	K2	0.0004959	0.0020208	0.0017924	0.0012016	0.0094394	0.0036389
	K3	0.0004942	0.0015892	0.0010197	0.0011895	0.0070656	0.0034375
$\hat{\beta}_{k,d}$	K1, do	0.0004947	0.0017381	0.0010785	0.0011935	0.0077237	0.0035134
	K2, do	0.0004955	0.0020127	0.0017860	0.0011989	0.009266	0.0036134
	K3, do	0.0004941	0.0015737	0.0009276	0.0011893	0.0069376	0.0034351

Results of the applied experiment:

Table (1-1) shows the regression parameter estimates for the best two-parameter Liu estimator in the experimental aspect of the Poisson regression model under the multicollinearity problem and its mean square error (MSE). The mean square error (MSE) for the two-parameter Liu estimator was obtained using Equation (19) presented in the theoretical aspect, based on the estimated values ($d_0=1.100$) ($k_3=0$) by writing a number of programs in MATLAB R2013a using data on the number of patient visits to the hospital in Iraq during a year, as shown in Table (1-1).

Table (1-1) Estimation of regression model parameters.

Estimator		β_1	β_2	β_3	β_4	β_5	MSE
$\hat{\beta}_k$	K1	0.23359	0.25266	0.14139	0.15329	0.27995	2.240360
	K2	1.68404	2.71415	1.70510-	2.01437-	4.11609	1.529702
	K3	2.63695	2.89762	2.53639-	1.94527-	3.56378	4.778442
$\hat{\beta}_{k,d}$	K1, do	1.13090	3.32063-	0.00969	0.00637	0.75775-	14.983712
	K2, do	1.13022-	1.61067	0.02791	0.02177	0.53447-	0.464117
	K3, do	1.13296-	1.61834	0.02790	0.02176	0.53292-	0.001355

From Table (1-1), we note that the mean squared error of the two-parameter Liu estimator reached a value of (0.00136). The table shows that the value of parameter β_1 was 1.13325 ($\beta_1 =$, which indicates the existence of an inverse relationship between the variable X_1 (patient's age) and the dependent variable (number of patient visits to the hospital), i.e., increasing the patient's age leads to a decrease in the number of visits. As for parameter β_2 , it was ($\beta_2 = 1.61913$), which reflects the existence of a direct relationship between the variable X_2 (body mass index) and the number of visits, such that visits increase with an increase in the body mass index. The value of parameter ($\beta_3 = 0.02789$) was, which indicates a direct relationship between the variable X_3 (number of chronic diseases) and the number of visits, i.e., an increase in the number of chronic diseases leads to an increase in the number of patient visits. The value of parameter ($\beta_4 = 0.02176$) was, which indicates the existence of a direct relationship between the variable X_4 (number of medications the patient takes monthly) and the number of visits, where visits increase with an increase in the number of medications. Finally, the value of parameter ($\beta_5 = -0.53276$), which indicates an inverse relationship between the variable X_5 (average monthly income) and the number of visits, i.e., an increase in the average monthly income leads to a decrease in the number of patient visits to the hospital.

Fourth, Conclusions and Recommendations

First, The Conclusions

1. Handling Multicollinearity: Reduces the impact of multicollinearity among explanatory variables because it uses instruments to replace correlated regressors. This generally provides more stable estimates compared to ordinary least squares (OLS).

Poisson IV: Sensitive to multicollinearity among explanatory variables. High multicollinearity can increase the variance of estimates and reduce their stability in count data models.

2. Bias and Efficiency: Can be less efficient than standard OLS if instruments are weak, but it addresses bias caused by correlation between explanatory variables and the error term.

Poisson IV: Appropriate for count data as it accounts for the correct distribution of the dependent variable, reducing bias compared to standard Poisson regression. However, multicollinearity can still inflate variance and affect efficiency.

3..Effect of Instrument Choice: Strong and valid instruments are critical. Weak instruments can increase bias and variance.

Poisson IV: Instrument selection is equally important, but evaluation is more complex due to the count nature of the dependent variable, especially under multicollinearity.

Second Recommendations

1.Check for Multicollinearity Before Modeling:

Use Variance Inflation Factor (VIF) or correlation analysis. If multicollinearity is severe, consider using instruments or variable transformations to mitigate its effects.

2. Choose Strong and Valid Instruments:

For both 2SLS and Poisson IV, instruments must be strongly correlated with endogenous regressors and uncorrelated with the error term. Strong instruments improve accuracy and reduce bias.

3. Assess Model Fit and Robustness:

Compare both models using goodness-of-fit measures, overdispersion tests (for Poisson), and sensitivity analysis. This ensures reliable inference in the presence of multicollinearity and validates instrument choice.

References:

1. Abdul-Ridha, Nour Alaa Al-Din (2024) "Comparing some methods of estimating the parameters of the Conway-Maxwell Poisson regression model in the presence of the multicollinearity problem with application" Master's thesis in Statistics, College of Administration and Economics, University of Baghdad pages 20.<https://coadec.uobaghdad.edu.iq>
2. Ismail, Milad Diaa (2024), "Employing some scaling functions to estimate the parameters of the beta regression model in the presence of the multicollinearity problem with application," Master's thesis in Statistics, College of Administration and Economics, University of Baghdad pages 19.
https://www.researchgate.net/publication/377981918_Employing_a_proposed_method_to_estimate_the_parameters_of_a_beta_regression_model_in_the_presence_of_multicollinearity
3. Ramadan, Ahmed Razzaq (2023) "Comparison between letter parameters in the modified two-parameter estimation of the negative binomial regression model with a practical application." Volume 38, pages 1463–1485.
4. Hoerl, A. E., & Kennard, R. W. (1970). "Ridge regression: Biased estimation for nonorthogonal problems." *Technometrics*, 12(1), 55-67.<http://jeasiq.uobaghdad.edu.iq>.
5. Abdelwahab, M. M., Abonazel, M. R., Hammad, A. T., & El-Masry, A. M. (2024). Modified two-parameter Liu estimator for addressing multicollinearity in the Poisson regression model. *Axioms*, 13(1), 46.
<https://www.mdpi.com/2075-1680/13/1/46>
- Qasim, M., Månsson, K., Amin, M., Golam Kibria, B. M., & Sjölander, P. (2020). Biased adjusted Poisson ridge estimators-method and application. *Iranian Journal of Science and Technology, Transactions A: Science*, 44(6), 1775-1789.<https://link.springer.com/article/10.1007/s40995-020-00974-5>
6. Kamal, Ghufra Ismail (1990) "An Evaluative Study of Biased Estimation Methods for Multiple Regression Parameters When Some Hypotheses Are Violated Using the Simulation Method" Master's Thesis, College of Administration and Economics, University of **Baghdad** , Pages 599-649.

7. Abdullah, Mifan Shukor (2024) "Estimating Restricted Gamma Regression Parameters in the Presence of the Multicollinearity Problem with Application" Master's Thesis, College of Administration and Economics, University of Baghdad Journal of Al-Rafidain University College of Science pages 265-273. [HTTPS:](https://doi.org/10.33899/ijjoss.2006.032751)

Al-Rashidani, Khairy Rashid (2006) "Evaluation of the Restricted Liu Method for Addressing the Multicollinearity Problem" College of Computer Science and Mathematics, Department of Statistics Pages 140-150.<https://doi.org/10.33899/ijjoss.2006.032751>

8. Sabry, Hussam Muwaffaq (2013) "Comparison of Poisson Regression Model Parameter Estimation Methods in the Presence of Multicollinearity Problem with Practical Application" PhD Thesis, College of Administration and Economics, University of **Baghdad** , 164-188.

9. Khorshid, Enas Salah (2019) "Comparison of Negative Binomial Regression Model Parameter Estimation Methods in the Presence of Multicollinearity Problem with Practical Application" Master's Thesis in Statistics, College of Administration and Economics, University of **Baghdad** vol.25, No.110 pages 466-488.

Ramadan, Ahmed Razzaq (2023), "Comparison of Letter Parameters in the Modified Two-Parameter Estimation of a Negative Binomial Regression Model with a Practical Application" vol (15) Issue:46-2023 pages 640-663 <https://jeasiq.uobaghdad.edu.iq/home/vol25/iss110/24/>

10. Al-Quraishi, O. A. K., & Atwan, A. F. H. (2024). Comparison of the kernel regression and spline regression methods in estimating the Poisson regression model with a practical application. Warith Scientific Journal, 6(71), 329–338. <https://wsj.uowa.edu.iq/index.php/warith/article/view/205>.

11. Bakhit, H. A., & Fathallah, S. (2006). Econometrics. Amman, Jordan: Dar Al-Yazouri Scientific for Publishing and Distribution, p. 338.

12. Kazem, M. K. (2021). Estimation of the parameters of the generalized Poisson distribution with an applied study (Master's thesis, University of **Baghdad**, College of Administration and Economics), p. 15.

13. Hameed, L. M. A., & Radam, S. A. (2023). Comparison of Poisson Regression and Conway Maxwell Poisson Models Using Simulation. Journal of Economics and Administrative Sciences, 29(136), 81–89. <http://jeasiq.uobaghdad.edu.iq>

14. Al-Tamimi, Z. H., Al-Saadoon, F. G., & Al-Thalibi, S. H. (2014). Regression analysis College of Administration and Economics, University of Basrah. (pp. 239–240).

15. Amin, M., Akram, M. N., & Kibria, B. G. (2021). "A new adjusted Liu estimator for the Poisson regression model Concurrency and computation: Practice and experience, 33(20).<https://doi.org/10.1002/cpe.6340>.

16. Månsson, K., Kibria, B. G., Sjölander, P., & Shukur, G. (2012). Improved Liu estimators for the Poisson regression model. International Journal of Statistics and Probability, 1(1), 2. <https://www.researchgate.net/publication/268042933>