Original Research Article

Robustness Test of the Two Stage K-L Estimator in Models with Multicollinear Regressors and Autocorrelated Error Term

Abstract

In a classical multiple linear regression analysis, multicollinearity and autocorrelation are two main basic assumption violation problems. When multicollinearity exists, biased estimation techniques such as Maximum Likelihood, Restricted Maximum Likelihood and most recent the K-L estimator by Kibria and Lukman (2020) are preferable to Ordinary Least Square. On the other hand, when autocorrelation exist in the data, robust estimators like Cochran Orcutt and Prais-Winsten (1954) estimators are preferred. To handle these two problems jointly, the study combines the K-L with the Prais-Winsten's two-stage estimator producing the Two-Stage K-L estimator proposed by Zubair & Adenomon (2021). The Mean Square Error (MSE) and Root Mean Square Error (RMSE) criterion was used to compare the performance of the estimators. Application of the estimators to two (2) real life data set with multicollinearity and autocorrelation problems reveals that the Two Stage K-L estimator is generally the most efficient.

Keywords: Autocorrelation, Autoregressive, K-L estimator, Multicollinearity, Regression

1.0 Introduction

Regression analysis is a classic commonly used prediction tool. Regression analysis explore the relationship between a dependent variable (response variable) and one or more independent variable (explanatory variable). The general single-equation linear regression model can be represented as:

$$y = \beta_0 + \sum_{j=1}^{k} \beta_j X_j + U \qquad (1.1)$$

where y is the dependent variable; $X_1, X_2, X_3, \ldots, X_k$ are the independent variables; $\beta_j, j = 0,1,2 \ldots k$ are the regression coefficients, U is the stochastic disturbance term or error term.

For a sample of n observations,

$$y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots + \beta_k X_{ik} + U_i \qquad (1.2)$$

where i = 1, 2,…n.

Thus,

$$y_1 = \beta_0 + \beta_1 X_{11} + \beta_2 X_{12} + \ldots + \beta_k X_{1k} + U_1$$
$$y_2 = \beta_0 + \beta_1 X_{21} + \beta_2 X_{22} + \ldots + \beta_k X_{2k} + U_2$$
$$\vdots$$
$$y_n = \beta_0 + \beta_1 X_{n1} + \beta_2 X_{n2} + \ldots + \beta_k X_{nk} + U_n$$

In vector form

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ . \\ . \\ . \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & X_{12} & \dots & X_{1k} \\ 1 & X_{21} & X_{22} & \dots & X_{2k} \\ 1 & X_{31} & X_{32} & \dots & X_{3k} \\ & & . & & \\ & & . & & \\ & & . & & \\ 1 & X_{n1} & X_{n2} & \dots & X_{nk} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ . \\ . \\ . \\ \beta_k \end{bmatrix} + \begin{bmatrix} U_0 \\ U_1 \\ U_2 \\ . \\ . \\ . \\ U_n \end{bmatrix}$$

$$n \times 1 \qquad n \times (k+1) \qquad (k+1) \times 1 \qquad n \times 1$$

The general form is:

$$y = X\beta + U \tag{1.3}$$

where y is an $(n \times 1)$ vector of observations of the dependent variable, X matrix is an $n \times (k+1)$ full rank matrix of explanatory variables, β is a $((k+1) \times 1$ vector of unknown parameters to be estimated, U is $(n \times 1)$ vector of random error. The parameter β in a linear regression model are commonly estimated using the Ordinary Least Squares Estimator (OLSE). The OLSE of β is given as:

$$\hat{\beta}_{OLS} = (X'X)^{-1} X'y \tag{1.4}$$

The estimator is generally preferred if there is no violation in any of the assumptions of the linear regression model (Johnston, 1972; Ayinde et al., 2018).

The assumption of uncorrelated errors must be valid for the efficiency of the OLSE. However, violation of this assumption which is called autocorrelation can be encountered in practice especially in time series data. The presence of some independent variables which is not included in the model but should be in the model or non-random measurement errors at the dependent variable may cause autocorrelation problem.

1.1 Autocorrelation Problem in a Linear Model

Autocorrelation in the error terms is a critical problem faced in the linear regression model and it poses serious effects on the OLSE. When autocorrelation is present in the errors the OLSE will not be efficient and the usual estimator of the variance–covariance matrix will be biased. Thus, it become misleading using these variances for the confidence intervals and hypothesis tests (Griffiths et al., 1993).

Alternative estimators to the OLSE were proposed. Some researchers have worked on the methods for detecting the presence of autocorrelation and alternative estimators to estimate the parameters in the linear regression model with autocorrelation error. These include Aitken (1935), Cochran and Orcutt (1949), Durbin and Watson (1950), Hildreth and Lu (1960), Rao and Grilliches (1969), Beach and Mackinnon (1978), Kramer (1980), Busse et al. (1994), Kramer and Hassler (1998), Kleiber (2001), Kramer and Marmol (2002), Butte (2002), Nwabueze (2000), Nwabueze (2005), Olaomi (2004), Olaomi (2006), Olaomi and Ifederu (2006), Grochova and Strelec (2013). In time-series applications, there are many structures of autocorrelation (Olaomi and Ifederu, 2008).

1.2 Multicollinearity Problem in a Linear Model

Another popular assumption is that the independent (explanatory) variables are independent. However, in practice, there may be near to strong linear relationship among the explanatory variables which is referred to as multicollinearity. According to literature, the performance of OLSE drops when there is multicollinearity. The estimator possesses large variance and occasionally the regression coefficient will exhibit wrong sign (Gujarati, 1995; Ayinde et al., 2018; Lukman and Ayinde, 2017).

Various methods of estimating the parameters in linear regression model with multicollinearity are available in the literature. Authors include Hoerl and Kennard (1970), McDonald and Galarneau (1975), Lawless and Wang (1976), Hocking, Speed and Lynn (1976), Dempster, Schatzoff and Wermuth (1977), Wichern and Churchill (1978), Gibbons (1981), Nordberg (1982), Saleh and Kibria (1993), Haq and Kibria (1996), Singh and Tracy (1999), Kibria (2003), Khalaf and Shukur (2005), Alkhamisi, Khalaf and Shukur (2006), Alkhamisi and Shukur (2008), Muniz and Kibria (2009), Dorugade and Kashid (2010), Mansson, Shukur and Kibria (2010), and recently Khalaf (2013), Ghadhan and Mohamed (2014), Dorugade (2014), Kibria and Shipra (2016), Ayinde et al. (2018), Lukman et al. (2017), Lukman et al. (2019a,b), Qasim et al. (2019), Kibria and Lukman (2020), Aslam and Ahmad (2020), Dawoud and Kibria (2020).

1.3 The Joint Violation of Autocorrelation and Multicollinearity Problem in a Linear Model

Literature have recently show that both problems can jointly exist in a linear regression model (Trenkler 1984; Bayhan and Bayhan, 1998; Ayinde et al., 2015; Lukman et al., 2015; Ozkale and Tugba, 2015; Tugba and Ozkale, 2019; Tugba, 2020). Trenkler (1984) proposed the generalized ridge estimator which takes the autocorrelation into account in the general linear regression model. Hussein and Zari (2012) combined the ridge regression estimator and the generalized least squares estimator to mitigate both problem. Recently, Eledum and Zahri (2013) proposed the feasible generalized ridge (FGR) estimator to deal with both the multicollinearity and autocorrelation problems. Dawoud and Kaçıranlar (2015) proposed the feasible generalized Liu (FGL) regression estimator by combining the Liu estimator and the feasible generalized least squares. Ozbay et al. (2016) combined the feasible generalized restricted ridge regression estimator to take account of both problems. Bello et al. (2017) also introduced feasible generalized Ridge Estimators as Alternatives to ridge and feasible generalized least squares estimator.

The Ordinary Least Square (OLS) estimator is popularly employed to estimate the regression parameter in the linear regression model (LRM). The estimator suffers setback in the presence of multicollinearity and/or autocorrelation. It produces inefficient estimates with large variance. Also, the two problems do exist jointly in LRM and in practice estimators to handle them together are rare. Thus, this research attempted to test the robustness of the Two Stage K-L estimator proposed by Zubair & Adenomon (2021) with other popular estimators in literature.

2.0 MATERIALS AND METHOD

2.0.1 Two – Stage K-L Estimator

Consider the Linear Regression Model with autoregressive of order 1, AR (1) given as:

$$y_t = \beta_0 + \beta_1 X_{t1} + \beta_2 X_{t2} + \cdots + \beta_p X_{tp} + U_t \qquad (2.1)$$

where $u_t = \rho u_{t-1} + \varepsilon_t$, ρ is the autocorrelation parameter ($|\rho| < 1$), ε_t is a random term such that $\varepsilon_t \sim N(0, \sigma^2)$, $E(\varepsilon_i \varepsilon_j) = 0$ $(i \neq j)$. Equation (3.1) in matrix form is written as follows:

$$y = X\beta + U \qquad (2.2)$$

Pre-multiplying both sides of equation (1.4) by an $n \times n$ non-singular matrix P, we obtain

$$Py = PX\beta + PU \qquad (2.3)$$

The error term becomes PU with $E(PU) = 0$ and $E(PU'UP)' = \sigma^2 P\Omega P'$. Thus, if it is possible to specify P such that $P\Omega P' = I$ implying that $P'P = \Omega^{-\prime}$, then the OLS estimates of the transformed variable PY and PX in equation (2.3) have all the optimal properties of OLS and so the usual inferences could be valid. Re-defining equation (2.3) as

$$y^* = X^*\beta + U^* \qquad (2.4)$$

where $y^* = Py$, $X^* = PX$ and $U^* = PU$.

The generalized least squares estimator is obtained as follows:

$$\hat{\beta}_{GLS} = (X^{*\prime}X^*)^{-\prime}\, X^{*\prime}y^* = (X'P'PX)^{-\prime}\, X'P'Py$$
$$= (X'\,\Omega^{-\prime}\, X)^{-\prime}\;\; X'\,\Omega^{-\prime}\, y \qquad (2.5)$$

Ω is a known positive definite (p.d.) matrix. However, in practice, Ω is often unknown. A common practice is to use the estimated matrix of Ω in order to find the estimated generalized least square estimator (EGLSE) or Two Stages method estimator that is more efficient than the GLSE. We reform the Two Stages procedure as follows to propose the new estimator.

From equation (2.4) where $E(U^*)=0$, $Cov(U^*)= \sigma^2 I$. Thus, the OLS estimator for model (2.4) is:

$$\hat{\beta}_{TS} = (X^{*\prime}X^*)^{-\prime}\, X^{*\prime}y^* \qquad (2.6)$$

Where:

$$y* = Py = \begin{bmatrix} (1-\rho^2)^{\frac{1}{2}} & 0 & 0 & \cdots & 0 & 0 \\ -\rho & 1 & 0 & \cdots & 0 & 0 \\ 0 & -\rho & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -\rho & 1 \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \cdot \\ \cdot \\ \cdot \\ Y_n \end{bmatrix}$$

$$X^{*}=PX=\begin{bmatrix} (1-\rho^2)^{\frac{1}{2}} & 0 & 0 & \cdots & 0 & 0 \\ -\rho & 1 & 0 & \cdots & 0 & 0 \\ 0 & -\rho & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -\rho & 1 \end{bmatrix} \begin{bmatrix} 1 & X_{11} & X_{12} & \ldots & X_{1k} \\ 1 & X_{21} & X_{22} & \ldots & X_{2k} \\ 1 & X_{31} & X_{32} & \ldots & X_{3k} \\ & & \cdot & & \\ & & \cdot & & \\ & & \cdot & & \\ 1 & X_{n1} & X_{n2} & \ldots & X_{nk} \end{bmatrix}$$

Note that $X^{*\prime}X^{*}=X'P'PX= X'\Omega^{-1}X$ and $X^{*\prime}y^{*}=X'P'Py= X'\Omega^{-1}y$,

Where:

$$\Omega^{-1} = \frac{1}{1-\rho^2}\begin{bmatrix} 1 & -\rho & 0 & \cdots & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & \cdots & 0 & 0 \\ 0 & -\rho & 1+\rho^2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1+\rho^2 & -\rho \\ 0 & 0 & 0 & \cdots & -\rho & 1 \end{bmatrix} \qquad (2.7)$$

After estimating the ρ, we obtained Ω^{-1} and then the two stage is given as Prais Winsten (1954):

$$\hat{\beta}_{TS}= (X'\,\Omega^{-1}\,X)^{-1}\ X'\,\Omega^{-1}\,y \qquad (2.8)$$

The proposed estimators in this study are as follows:

Kibria and Lukman (2020) proposed the K-L estimator to solve the problem of multicollinearity in the LRM.

$$\hat{\beta}_{KL} = (X'X + kI)^{-1}(X'X - kI)\hat{\beta}_{OLS} \qquad (2.9)$$

Following (2.8) and (2.9), is the Two stage K-L estimator by Zubair & Adenomon (2021), as follows:

$$\hat{\beta}_{TKL} = (X'\Omega^{-1}X + kI)^{-1}(X'\Omega^{-1}X - kI_p)\hat{\beta}_{TS} \qquad (2.10)$$

Where:

$$\hat{k} = \min\left[\frac{\hat{\sigma}^2}{2\hat{\beta}_i^2 + \frac{\hat{\sigma}^2}{\lambda_i}}\right].$$

2.0.2 Existing Estimators

Feasible Generalized Least Square Estimators (FGLSE) includes Corchran Orcutt Estimator (CORC), Maximum Likelihood Estimator (MLE), Restricted Maximum Likelihood Estimator (RMLE) and Prais Winsten Estimator.

2.2 Data Description

Two data sets are used in this study to examine the performance of the estimators. The data sets are given in details below:

2.2.1 Hussein Data

The data set used in this study was adopted from the study of Hussein and Zari (2012). The regression model was defined as: where: Y_t is the product value in the manufacturing sector; X_{t1} is the value of the imported intermediate raw materials; X_{t2} is the value of imported capital commodities; X_{t3} is the value of imported raw material.

2.2.2 French Economy Data

The detail about this data set is initially described in Chatterjee and Price (1977) and later available in the following references Malinvard (l980) and Liu (1993). It comprises of one predictand, Imports and three predictor variables (domestic production, stock information and domestic consumption) with eighteen observations.

2.3 Criterion for Investigation and Performance of Estimators

Several authors in literatures have used the Mean Square Error (MSE) to compare the performance of 2 stage K-L regression estimator with the Ordinary Least Square estimator when there is multicollinearity and autocorrelation. These authors include Hoerl and Kennard (1970), Lawless and Wang (1976), Saleh and Kibria (1993), Kibria (2003), Khalaf and Shukur (2005), Alkhamisi and Shukur (2008), Mansson *et al*. (2010), Ozkale (2014), Dawoud and Kaciranlar (2015), Ozbay et al. (2016). For each replicate, the estimated MSE for each of the estimators α^* is obtained as follows:

$$MSE(\alpha^*) = \frac{1}{2000}\sum_{i=1}^{2000}(\alpha^* - \alpha)'(\alpha^* - \alpha), \tag{2.11}$$

where α^* would be any of the estimators earlier listed in Section 2.0. The estimator with the smallest estimated MSE is considered best.

3.0 RESULTS AND DISCUSSION

3.1.1 Result from Hussein Data

The data set used in this study was adopted from the study of Hussein and Zari (2012). The regression model was defined as:

$$Y_t = \beta_1 X_{t1} + \beta_2 X_{t2} + \beta_3 X_{t3} + U_t \tag{2.12}$$

t=1, 2, …, 31

where;

Y_t is the product value in the manufacturing sector
X_{t1} is the value of the imported intermediate raw materials
X_{t2} is the value of imported capital commodities
X_{t3} is the value of imported raw material

The estimated model using the ordinary least squared is:

$$\hat{Y} = 208.88 + 0.611X_1 + 1.256X_2 - 1.217X_3 \tag{2.13}$$

Model (2.13) was diagnosed using the variance inflation factor (VIF) and the Durbin-Watson (DW) test. According to the Kibria and Lukman (2020), there is multicollinearity when the VIF exceeds 10. The variance inflation factor for the variables are: 128.29, 103.43, and 70.87. It is evident from the VIFs that the model suffers from the problem of multicollinearity. The DW value is 0.905 which shows that the error terms are correlated. Therefore, the model suffers from multicollinearity and autocorrelation. The $X'X$ (correlation form) is

$$r = \begin{pmatrix} 1 & 0.9947 & 0.9923 \\ 0.9947 & 1 & 0.9905 \\ 0.9923 & 0.9905 & 1 \end{pmatrix}$$

Table 1: Regression coefficients and mean squared error

Coef	OLS	MLE	CC	RMLE	PW	RP	TSK
B0	208.885	345.318	452.101	365.957	606.62	-0.2365	-316.61
B1	0.613	0.488	0.311	0.44	0.437	0.149	0.742
B2	1.256	1.146	1.572	1.104	1.121	-0.349	1.171
B3	-1.221	-1.418	-1.871	-1.374	-1.403	167.24	-1.385
MSE	12.8075	20.6608	21.321	397.3	9.8659	1.9597	1.9252
RMSE	3.579	4.545	4.617	19.932	3.141	1.400	1.388

The autocorrelation value ρ was obtained using the Cochrane-Orcutt and Prais-Winsten estimator. The rhos are 0.9051 and 0.86387, respectively. The rhos are employed to transform the original data. The results of the model subject to the estimators are in Table 1.

The intercept term of RP and TSK has a negative sign while that of other estimators has a positive sign. This might be due to the multicollinearity effect on other estimators such as OLS, MLE, CC, RMLE, and PW. According to Lukman et al. (2019), the regression coefficient might exhibit a wrong sign when there is multicollinearity. The result in Table 1 shows that the proposed estimator TSK produced the most efficient estimates in terms of lower mean squared error and root mean square. Restricted maximum likelihood estimator has the least performance when there is multicollinearity and autocorrelated error.

3.1.2 French Economy Data

The detail about this data set is initially described in Chatterjee and Price (1977) and later available in the following references Malinvard (l980) and Liu (1993). It comprises of one predictand, Imports and three predictor variables (domestic production, stock information and domestic consumption) with eighteen observations. The variance inflation factors are $VIF_1 = 469.688$, $VIF_2 = 1.047$, $VIF_3 = 469.338$ and the condition number 32612. It is evident that there is multicollinearity in the model. from the VIFs that the model suffers from the problem of multicollinearity. The DW statistic and the p-value are 0.2429 and 0.0000, respectively. The DW result revealed that the model follows an AR(1), that is there is autocorrelation problem. Hence, the model suffers from both problems.

Table 2: Regression coefficients and mean squared error

Coef	OLS	MLE	CC	RMLE	PW	RP	TSK
B0	-19.71	-23.345	-22.221	-23.738	-23.402	18.594	23.387
B1	0.033	-0.031	-0.042	-0.028	-0.031	1.351	0.038
B2	0.406	0.479	0.512	0.474	0.479	-0.309	0.424
B3	0.242	0.365	0.420	0.371	0.366	-1.837	0.122
MSE	7.6574	8.6687	4.321	64274.95	3.464	7.139	3.365
RMSE	2.767	2.944	2.079	253.525	1.861	2.672	1.834

The result in Table 2 also shows that the proposed estimator has the lowest mean squared error and root mean square and is considered best. The result of the restricted maximum likelihood is the worst due to the presence of multicollinearity. The TSK estimator result is also consistent.

4.0 CONCLUSION

Ordinary Least Square (OLS), Maximum likelihood, Restricted maximum likelihood, Prais-Winsten and Cochran Orcutt estimators could not perform well in term of their Mean Squared Error (MSE) in the presence of multicollinearity and autocorrelation. The restricted maximum likelihood was even observed to be the worst of the estimators since has the largest value of mean square error (MSE). It is however observed that in the both instances of the application to real life data set, the Two Stage K-L estimator perform better than the all other six estimators when both problems exist. Finanly, this study confirms the simulation study result in work of Zubair & Adenomon (2021) which also shows that the two stage K-L estimator is most prefer in fitting a linear model when the assumptions of multicollinearity and autocorrelation is violated.

5.0 References

Alheety, M. and Kibria, B. M. (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.

Atiken, A. C. (1935). On the least square and linear combination of observations. Proceeding of Royal Statistical Society, Edinburgh. 52:42–48

Alkhamisi, M., Khalaf, G. and Shukur, G. (2006). Some modifications for choosing ridge parameters. Communications in Statistics - Theory and Methods, 35(11), 2005-2020.

Alkhamisi, M. and Shukur, G. (2008). Developing ridge parameters for SUR model. Communications in Statistics - Theory and Methods, 37(4), 544-564.

Aslam, M. and Ahmad, S. (2020). The modified Liu-ridge-type estimator: a new class of biased estimators to address multicollinearity. Communications in Statistics - Simulation and Computation. doi:10.1080/03610918.2020.1806324

Ayinde, K., Lukman, A. F. and Arowolo, O.T. (2015). Combined parameters estimation methods of linear regression model with multicollinearity and autocorrelation. Journal of Asian Scientific Research, 5(5), 243-250.

Ayinde K., Lukman A. F., Samuel O. O. and Ajiboye S. A. (2018). Some New Adjusted Ridge Estimators of Linear Regression Model. International Journal of Civil Engineering and Technology, 9(11), 2838-2852.

Ayinde, K., Lukman, A. F., Alabi, O. O. and Bello, H. A. (2020). A new approach of principal component regression estimator with applications to collinear data. International Journal of Engineering Research and Technology, 13(7), 1616-1622.

Beach, C. M. and Mackinnon, J. S. (1978). A Maximum Likelihood Procedure regression with autocorrelated errors. Econometrica, 46, 51 – 57.

Cochrane, D. and Orcutt, G. H. (1949). Application of Least Square to relationship containing autocorrelated error terms. Journal of American Statistical Association, 44, 32–61.

Dawoud, I and Kibria, B. M. G. (2020). A New Biased Estimator to Combat the Multicollinearity of the Gaussian Linear Regression Model. Stats, 3, 526–541. doi:10.3390/stats3040033

Dorugade, A. V. and Kashid, D. N. (2010). Alternative method for choosing ridge parameter for regression. International Journal of Applied Mathematical Sciences, 4(9), 447-456.

Dorugade, A. V. (2014). On comparison of some ridge parameters in Ridge Regression. Sri Lankan Journal of Applied Statistics, 15(1), 31-46.

Eledum, H. and Zahri, M. (2013). Relaxation Method For Two Stages Ridge Regression Estimator. International Journal of Pure and Applied Mathematics, 85, 4, 653-667.

Hocking, R., Speed, F. M. and Lynn, M. J. (1976). A class of biased estimators in linear regression. Technometrics, 18 (4), 425-437.

Hoerl, A.E. and Kennard, R.W. (1970). Ridge regression: biased estimation for non-orthogonal problems. Technometrics, 12, 55-67. Hussein, Y. and Zari, M. (2012). Generalized Two Stage Ridge Regression Estimator TR for Multicollinearity and Autocorrelated Errors. Canadian Journal on Science and Engineering Mathematics, 3(3), 79-85.

Johnston, J. (1972). Econometric Methods, 2nd Ed. McGraw-Hill Book Co., Inc., New York.

Kibria, B. M. G. and Shipra, B. (2016). Some Ridge Regression Estimators and Their Performances. Journal of Modern Applied Statistical Methods, 15(1), 206-231.

Kibria, B. M. G. and Lukman, A. F. (2020). A New Ridge-Type Estimator for the Linear Regression Model: Simulations and Applications. Hindawi Scientifica.

Lawless, J. F. and Wang, P. (1976). A simulation study of ridge and other regression estimators. Communications in Statistics A, 5, 307-323.

Lukman. A. F, Osowole, O.I. and Ayinde, K. (2015). Two Stage Robust Ridge Method in a Linear Regression Model. Journal of Modern Applied Statistical Methods, 14(2), 53-67.

Lukman, A. F., and Ayinde, K. (2017). Review and Classifications of the ridge parameter estimation techniques. Hacettepe Journal of Mathematics and Statistics, 46(5), 953-967.

Lukman, A. F., Ayinde, K. and Ajiboye, S. A. (2017). Monte-Carlo Study of Some Classification-Based Ridge Parameter Estimators. Journal of Modern Applied Statistical Methods, 16(1), 428-451.

Lukman, A. F., K. Ayinde, S. K. Sek, and E. Adewuyi. (2019a). A modified new two-parameter estimator in a linear regression model. Modelling and Simulation in Engineering 2019:6342702. doi:doi:10.1155/2019/6342702

Lukman, A. F.; Ayinde, K.; Binuomote, S. and Clement, O. A. (2019b). Modified ridge-type estimator to combat multicollinearity: application to chemical data. Journal of Chemometrics, 33(5), e3125.

Lukman, A. F., Ayinde, K., Aladeitan, B. B. and Rasak, B. (2020). An Unbiased Estimator with Prior Information. Arab Journal of Basic and Applied Sciences, 27:1, 45-55.

Mansson, K., Shukur, G. and Kibria, B. M. G. (2010). A simulation study of some ridge regression estimators under different distributional assumptions. Communications in Statistics- Simulations and Computations, 39(8), 1639 –1670.

Muniz, G. and Kibria, B. M. G. (2009). On some ridge regression estimators: An empirical comparison. Communications in Statistics-Simulation and Computation, 38, 621-630.

Nwabueze, J. C. (2005). Performance of estimators of linear model with autocorrelated error terms when the independent variable is normal. J. Nig. Assoc. Mtah. Phys., 9: 379-384.

Olaomi, J. O. (2004). Estimation of parameters of linear regression models with autocorrelated error terms which are also correlated with the regressor. Unpublished PhD Thesis, University of Ibadan, Nigeria.

Olaomi, J. O. (2006). Estimation of parameters of linear regression models with autocorrelated error terms which are also correlated with the regressor. Global J. Pure Applied Sci., 13(2): 237-242.

Olaomi, J. O. and Ifederu, A. (2006). Estimation of the parameters of linear regression model with autocorrelated error terms which are also correlated with the trended regressor. A paper presented at the 11th Annual African Economic Society (AES), Dakar, Senegal.

Trenkler, G. (1984). On the performance of biased estimators in the linear regression model with correlated or heteroscedastic errors. Journal of Econometrics, 25, 179-190.

Tuğba, S. A. and Özkale, M. R. (2019). Regression diagnostics methods for Liu estimator under the general linear regression model, Communications in Statistics - Simulation and Computation, DOI: 10.1080/03610918.2019.1582781

Tuğba, S. A.. (2020). Identification of Leverage Points in Principal Component Regression and r-k Class Estimators with AR(1) Error Structure. 6(2), 353-363.

Zubair, M.A. and Adenomon, M.O. (2021). Comparison of estimator’s efficiency for linear regressions with joint presence of autocorrelation and multicollinearity. Science World Journal vol. 16(2): 103-109.

Appendix:

Code for the Application to Real-Life Data

```
library(prais)
library(orcutt)
library(nlme)
library(neuralnet)
library(matlib)
library(lmtest)
dd=read.csv(file.choose(),header=T)
hildreth_lu=function(resp,regr){
  rh=seq(-1,1,.1)
  msess=NULL
  for(r in 1:length(rh)){
    xt=regr
    yt=resp
    t=2:length(yt)
    yt[t]=yt[t]+rh[r]*yt[(t-1)]
    xt[t,]=xt[t,]+rh[r]*xt[(t-1),]
    dat=data.frame(yt,xt)
    bhat=coef(lm(yt~.-1,data=dat))
    msess=rbind(msess,sum((yt-xt%*%bhat)^2))
  }
  tab=data.frame(rh,msess)
  tab=subset(tab,msess!=0)
  rho.min=rh[which.min(tab$msess)]

  xt2=regr
  yt2=resp
  t=2:length(resp)
  yt2[t]=yt2[t]+rho.min*yt2[(t-1)]
  xt2[t,]=xt2[t,]+rho.min*xt2[(t-1),]
  dat2=data.frame(yt2,xt2)
  ols2=coef(lm(yt2~.-1,dat2))
  return(list(coefficients=ols2,rho=rho.min, tab=tab))
}

#ordinary regressions

olsall=lm(Y~.,data=dd,x=T,y=T)
ols1=olsall$coefficients
x=olsall$x
    l=t(x)%*%x
    q=as.matrix(eigen(l)$vectors)
    e=eigen(l)$values
sig=summary(olsall)$s

mle <- gls(Y~., data=dd, correlation=corAR1(form=~1), method="ML")
sigm=summary(mle)$s
```

```
mle1=mle$coefficients

model2<-prais_winsten(Y~., data=dd)
sigp=summary(model2)$s

rmle <- gls(Y~., data=dd, correlation=corARMA(p=1), method="REML")
sigrm=summary(rmle)$s
rmle1=rmle$coefficients

praisall=prais_winsten(Y~.,data=dd)
prais1=praisall$coefficients
rhop=tail(praisall$rho,n=1)

cocall=cochrane.orcutt(olsall)
coc1=cocall$coefficients
rhoc=cocall$rho
sigc=111.2992

y=olsall$y
x=olsall$x
hilu=hildreth_lu(y,x)
hilu1=hilu$coefficients
rhoh=hilu$rho

tp=2:length(y)
x_prais=x[tp,]-c(rhop)*(x[(tp-1),])
y_prais=y[tp]-c(rhop)*y[(tp-1)]

x_coc=x[tp,]-c(rhoc)*(x[(tp-1),])
y_coc=y[tp]-c(rhoc)*y[(tp-1)]

x_hlu=x[tp,]-c(rhoh)*(x[(tp-1),])
y_hlu=y[tp]-c(rhoh)*y[(tp-1)]

# Biased Regressions
k=deviance(olsall)*length(ols1)/sum(ols1^2)
k1=sqrt(max(sig/((2*(ols1^2))+(sig/e))))
#k1=sqrt(max(ols1^2/((sig/e)+(ols1^2))))
I=diag(1,length(ols1))
# K-L
kl=solve(t(x)%*%x+k1*I)%*%(t(x)%*%x-k1*I)%*%ols1
kl_pr=solve(t(x_prais)%*%x_prais+k1*I)%*%(t(x_prais)%*%x_prais-k1*I)%*%prais1
kl_coc=solve(t(x_coc)%*%x_coc+k1*I)%*%(t(x_coc)%*%x_coc-k1*I)%*%prais1

mseols=sig*sum(1/e)
mseml=sigm*sum(1/e)
mserml=sigrm*sum(1/e)
```

```
msep=sigp*sum(1/e)
msec=sigc*sum(1/e)

msekl=sig*sum(((e-k1)^2)/(e*(e+k1)^2))+(4*(k1^2))*sum((ols1^2)/(e+k1)^2)
msekl_pr=sigp*sum(((e-k1)^2)/(e*(e+k1)^2))+(4*(k1^2))*sum((ols1^2)/(e+k1)^2)
msekl_coc=sigc*sum(((e-k1)^2)/(e*(e+k1)^2))+(4*(k1^2))*sum((ols1^2)/(e+k1)^2)

res=cbind(mseols,mseml,mserml,msep,msekl,msekl_pr,msekl_coc)
print(res)
res1=cbind(ols1,mle1,rmle1,prais1,kl,kl_pr,kl_coc)
print(res1)
res2=cbind(sig,sigm,sigrm,sigp)
print(res2)
```