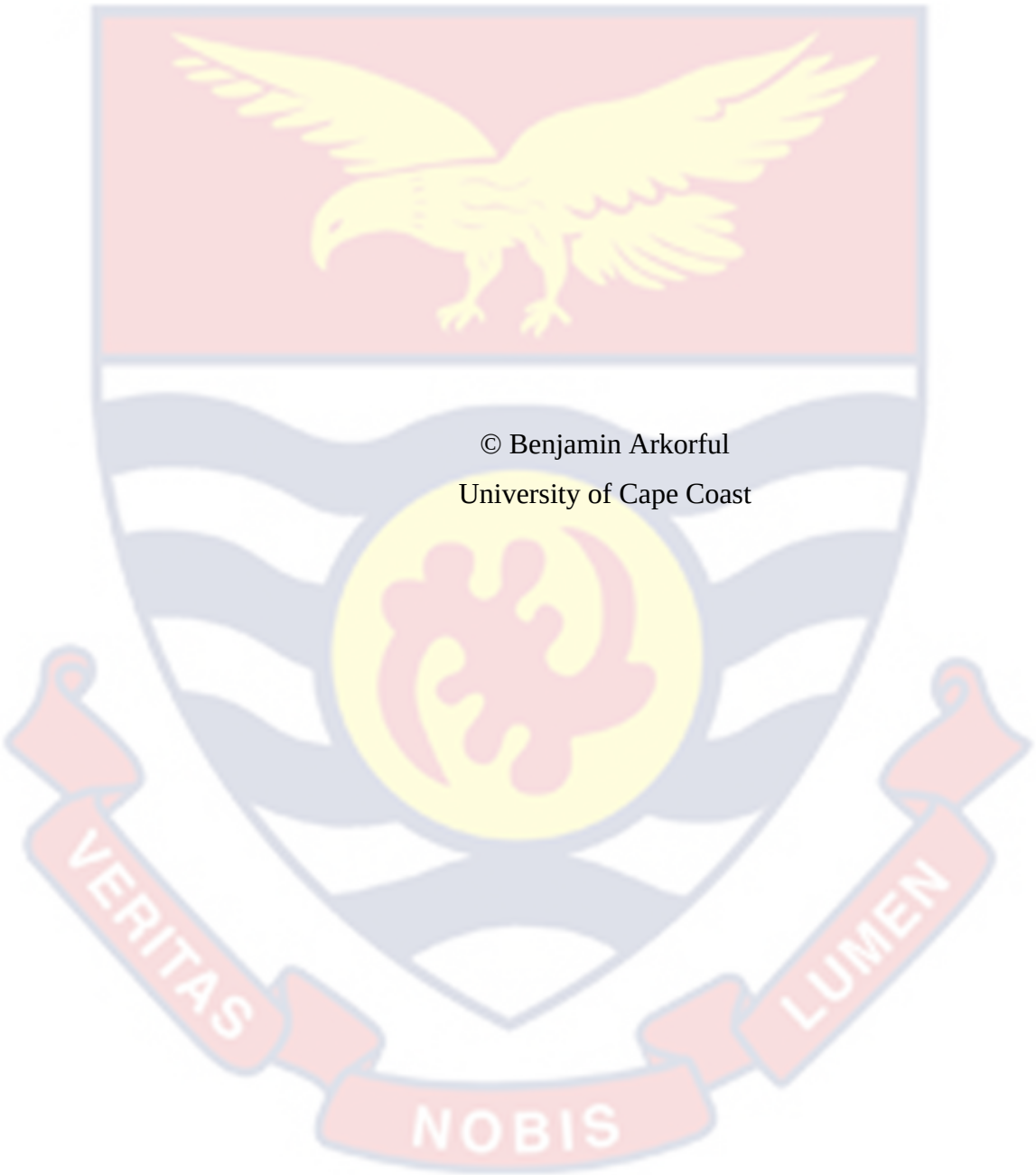
UNIVERSITY OF CAPE COAST

REGULARIZED AND ROBUST REGRESSION METHODS OF LINEAR MODEL WITH MULTICOLLINEAR PREDICTORS AND AUTOCORRELATED ERRORS

BENJAMIN ARKORFUL

2023



UNIVERSITY OF CAPE COAST

REGULARIZED AND ROBUST REGRESSION METHODS OF LINEAR MODEL WITH MULTICOLLINEAR PREDICTORS AND AUTOCORRELATED ERRORS

BY

BENJAMIN ARKORFUL

Thesis submitted to the Department of Statistics of the School of Physical Sciences, College of Agriculture and Natural Sciences, University of Cape Coast, in partial fulfilment of the requirements for the award of Master of Philosophy degree in Statistics

MAY 2023

DECLARATION

Candidate's Declaration

I hereby declare that this thesis is the result of my own original research and that no part of it has been presented for another degree in this university or elsewhere.

Candidate's Signature …… Date ……

Name: Benjamin Arkorful

Supervisor's Declaration

I hereby declare that the preparation and presentation of the thesis were supervised in accordance with the guidelines on supervision of thesis laid down by the University of Cape Coast.

Supervisor's Signature …… Date ……

Name: Prof. Nathaniel Howard

ABSTRACT

Multicollinearity and autocorrelation problems pose a major challenge to the Ordinary Least Squares (OLS) estimate. The purpose of the study is to make a comparative analysis of different regularized and robust regression methods to determine the regression method that best addresses multicollinearity and autocorrelation problems in a linear model. The regularized and robust regression methods used in the study include Ridge Regression (RR), Lasso Regression, Two-Stage Ridge Regression (TR), Two-Stage Lasso Regression (TLasso), Quantile Regression (QR), Ridge Quantile Regression (RQR), Lasso Quantile Regression (LQR), Two-Stage Ridge Quantile Regression (TRQR) and Two-Stage Lasso Quantile Regression (TLQR), as remedies to OLS estimate. The data used for the study consists of simulated and two real datasets with multicollinearity and autocorrelation issues. The Mean Squared Error as the main performance criterion as well as other statistical values (such as regression coefficients, R-squared, adjusted R-squared, Root Mean Squared Error) were used for comparing the performances of the regression methods. The results indicate that the TLQR method with 0.5 quantile level is suitable for handling multicollinearity and autocorrelation problems with many predictor variables. The TR method performs better with few predictor variables. Another important observation is the effect of sample sizes on the regression methods. To effectively build a good model, one should aim at choosing the appropriate samples and predictors with corresponding right regression method for data studies which include possible multicollinearity and autocorrelation issues.

KEY WORDS

Autocorrelation

Lasso Regression

Multicollinearity

Quantile Regression

Ridge Regression

Two-Stage Regression

ACKNOWLEDGEMENTS

The success of this thesis has been possible largely because of the dedication and commitment of the contributors and their efforts are hereby acknowledged. I wish to express my profound gratitude to my supervisor, Prof Nathaniel Howard for his time, patience, guidance, encouragement and contribution that has made this thesis a success. I am also grateful to Dr David K. Mensah, for his various contributions to the success of this thesis. I would like to acknowledge all lecturers of the Department of Statistics, University of Cape Coast for their support throughout the MPhil studies. I also appreciate the support and encouragement from all my colleagues and the non-academic staff of the Department of Statistics of the University. Finally, to all who has contributed to this thesis in diverse ways, I say thank you and may the good God richly reward your efforts.

DEDICATION

To my family

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ABBREVIATIONS

AIC	Akaike Information Criterion
BLUE	Best Linear Unbiased Estimator
DW	Durbin-Watson Test
FBQR	Flexible Bayesian Quantile Regression
FGLS	Feasible Generalized Least Squares
FGRLS	Feasible Generalized Restricted Least Squares
FGRR	Feasible Generalized Restricted Ridge Regression
GLS	Generalized Least Squares
GRR	Generalized Ridge Regression
GTR	Generalized Two-Stage Ridge Regression
JKL	Jackknife Kibria-Lukman
LASSO	Least Absolute Shrinkage and Selection Operator
LQR	Lasso Quantile Regression
MLE	Maximum Likelihood Estimator
MSE	Mean Squared Error
OLS	Ordinary Least Squares
PCR	Principal Component Regression
PLS	Partial Least Squares
QR	Quantile Regression

RJKL	Robust Jackknife Kibria-Lukman
RLS	Restricted Least Squares
RMLE	Restricted Maximum Likelihood Estimator
RMSE	Root Mean Squared Error
RQR	Ridge Quantile Regression
RR	Ridge Regression
SE	Standard Error
STS	Two-Stage Stein
TL	Two-Stage Liu
TLasso	Two-Stage Lasso Regression
TLQR	Two-Stage Lasso Quantile Regression
T-PC	Two-Stage Principal Component Regression
T-PLS	Two-Stage Partial Least Squares
TR	Two-Stage Ridge Regression
TRQR	Two-Stage Ridge Quantile Regression
TS	Two-Stage Regression
VIF	Variance Inflation Factor

CHAPTER ONE

INTRODUCTION

In this introductory chapter, the challenges multicollinearity and autocorrelation pose to linear models has been specified in the background to the study. It is on this background that this study attempts to present different regularized and robust regression methods in handling multicollinearity and autocorrelation problems in linear models. The study also puts into perspective other studies done in the area using other regression methods.

Background to the Study

The Ordinary Least Squares (OLS) estimator is the most widely used estimator for estimating the parameters of a regression model. The OLS estimate has some very interesting statistical properties that have made it one of the most powerful and commonly used estimators of regression models, under certain assumptions. When there is no violation of the classical linear regression model assumptions, OLS outperforms other estimators in parameter estimation. (Lukman, Osowole, & Ayinde, 2015). According to the Gauss-Markov theorem, the OLS estimate has the least mean squared error of all linear estimators with no bias (Greene, 2003; Hastie, Tibshirani & Friedman, 2009). The OLS estimator with $\mathrm{E}(\mathbf{e}) = 0$ and $\mathrm{var}(\mathbf{e}) = \sigma^2 \mathbf{I}_n$ is considered the Best Linear Unbiased Estimator (BLUE), according to the Gauss-Markov theorem. This indicates that, of all the linear estimators, OLS will have the smallest variance, guaranteeing that the regression estimates are unbiased.

There is a biased estimator that has a lower mean squared error. A little bias for a larger reduction in variance would then be traded for by the biased estimator. Therefore, methods that shrinks or sets to zero some of the regression

coefficients of the least squares may produce a biased estimate. These penalized estimation methods were developed to reduce estimate variance and thus improve prediction by means of introducing a slight bias into model estimation (Hastie, Tibshirani, & Friedman, 2009). Recently, much emphasis has been placed on biased estimation of regression coefficients in a linear model. This development is due to the inability of the OLS to provide accurate estimates when the matrix of predictor variables is ill-conditioned. If the predictors are orthogonal, the OLS estimator is optimal among the class of linear unbiased estimators. However, if the regression model contains highly correlated predictors, multicollinearity occurs.

Multicollinearity is one ill-conditioned problem in linear regression models. The term "multicollinearity" was first used in literature by Frisch (1934) on the topic; "statistical confluence analysis by means of complete regression systems". Multicollinearity is "a statistical phenomenon that happens when two or more predictor variables in a regression model are highly correlated" (Corlett, 1990; Jensen & Ramirez, 2013; Johnson, Reimer, & Rothrock, 1973). Multicollinearity may cause issues with the computation of OLS estimations. The multicollinearity issue in fact bedevils the whole subject of linear regression, and is surprisingly a common phenomenon. It is one reason why the multicollinearity problem is 'an art as well as a science' (Bingham & Fry, 2010). The greater a predictor variable's correlation with the other predictor variables in a model, the greater its variance. An estimator will therefore have lower variance when there is a greater variation in the model's predictor variables.

The nature of multicollinearity can be classified into perfect (or exact) and imperfect (or approximate). Multicollinearity is used to mean the presence of a

"perfect" or exact linear relationship between some or all predictor variables in a regression model. It is, however used more broadly to include both the case of perfect multicollinearity and the case where the predictor variables are inter-correlated but not perfectly correlated (Gujarati & Porter, 2009). As a result, its degree must be determined, varying from no collinearity to perfect collinearity. Perfect multicollinearity is an extreme and rare case that occurs when two or more predictor variables in a regression model are linearly dependent. The case of a perfect linear relationship among predictors is a serious failure of the model's assumptions, not of the data. When the predictor variables have an exact relationship, the OLS estimates of the regression coefficients are indeterminate, and the standard error of the estimates becomes infinite. In the case of perfect multicollinearity, the matrix $\mathbf{X}^T\mathbf{X}$ and matrix $\mathbf{X}$ do not have full rank. Consequently, the inverse matrix $(\mathbf{X}^T\mathbf{X})^{-1}$ cannot be computed, therefore $\hat{\boldsymbol{\beta}}_{OLS} = \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{y}$ cannot be solved, and the OLS estimator has no unique solution (Flexeder, 2010; Hashem, 2014).

The most common situation of multicollinearity is when the variables are highly, but not perfectly correlated. In this scenario, regression estimates are determinate but possess large standard error implying that the coefficients cannot be estimated with great precision. When high degree of multicollinearity exists, the matrix $\mathbf{X}^T\mathbf{X}$ is quasi-singular (ill-conditioned). Thus, $\mathbf{X}$ is of full rank and matrix $\mathbf{X}^T\mathbf{X}$ is regular and has a unique solution. However, due to highly correlated predictors in the model, the determinant $\left|\mathbf{X}^T\mathbf{X}\right|$ reaches a value near zero and the computed OLS estimate possesses a very large variance, $var(\hat{\boldsymbol{\beta}}_{OLS}) = \sigma^2(\mathbf{X}^T\mathbf{X})^{-1}$ (Flexeder, 2010; Giacalone, Panarello, & Mattera,

2018). If the determinant defined by $\left|\mathbf{X}^T\mathbf{X}\right| = \prod_{i=1}^{n} \lambda_i$ gives one or more small eigenvalues when multicollinearity is high but imperfect, the distance from OLS vector $(\hat{\boldsymbol{\beta}}_{OLS})$ to true parameter $(\boldsymbol{\beta})$ will tend to be large, that is, $\|\boldsymbol{\beta}_{OLS}\| \geq \|\boldsymbol{\beta}_{true}\|$ (Hoerl & Kennard, 1970).

The presence of multicollinearity would misleadingly inflate in an excessive amount the standard error, and the regression coefficients may not appear significant. Despite OLS having the very desirable property of being BLUE under the usual model conditions, OLS estimators can have extremely large mean squared errors in the presence of multicollinearity. Many approaches have been developed in detecting as well as solving problems associated with multicollinearity in regression analysis. When multicollinearity is detected, the simplest approach is to drop the variable(s) or the problematic variables from the model (Demaris, 2004; Giacalone, Panarello, & Mattera, 2018; Weisberg, 2005), increase the number of observations or add new data (Demaris, 2004; Gujarati & Porter, 2009), combine time series and cross-sectional data (Gujarati & Porter, 2009), incorporate variables into a scale (Demaris, 2004), utilize variable transformation (Demaris, 2004; Gujarati & Porter, 2009), and apply centering for multicollinearity caused by cross-product terms (Demaris, 2004). However, there are times when none of these solutions are satisfactory (Demaris, 2004).

Several methods in literature have been proposed to handle multicollinearity problem in regression analysis. For instance, Stein (1956) proposed stein estimator; Least Absolute Shrinkage and Selection Operator (LASSO) developed by Tibshirani (1996); Hoerl and Kennard (1970) developed Ridge regression; Zou and Hastie (2005) designed the Elastic net, by combining the

L_1-penalty (Lasso) and the L_2-penalty (Ridge); Massy (1965) suggested Principal Component Regression estimator (PCR); and Wold (1966) introduced Partial Least Squares to handle multicollinearity problem.

The autocorrelation problem is a violation of the error term assumption of independence in the classical linear regression model. When the Gauss-Markov assumption of uncorrelated error terms is violated, autocorrelation occurs. The correlation between elements of a time-ordered series of observations, as in time series data, is characterized as autocorrelation (Gujarati, 2003). In the context of regression, the classical linear regression model implies that autocorrelation does not occur in the errors. When this assumption fails, we have an autocorrelation problem. Violation of the no autocorrelation assumption on error terms, will lead to inefficiency of the OLS estimates. The variance of the regression coefficients estimated will be biased and inconsistent, and it will be larger than the variances calculated by other methods. As a result, among all linear unbiased estimators, the OLS estimators will no longer have the least variance. If the errors are interrelated, OLS may underestimate the standard error of the coefficients. Standard errors that are underestimated can make predictors appear significant when indeed they are not. This leads to wrong standard errors for the regression coefficient estimates, therefore, testing of hypotheses is no longer valid. The *F*-statistic and *t*-statistic will tend to be higher and therefore, may not be valid (Damoder, 2006). Furthermore, the OLS estimates are not asymptotic if the error terms are correlated. In other words, estimated β s are not asymptotically efficient (Greene, 2003; Gujarati, 2003).

Several corrective procedures based on variable transformations have been proposed, to correct for autocorrelation. They are, the use of Generalized Least Squares (GLS) or Feasible Generalized Least Square (FGLS) techniques such as the Prais-Winsten estimator, Cochrane-Orcutt estimator, Hildreth and Lu estimator, Durbin estimator, Thornton estimator, Theil's estimator, Maximum Likelihood estimator and Maximum Likelihood Grid estimator (Cochrane & Orcutt, 1949; Hildreth & Lu, 1960; Paris & Winstein, 1954; Thornton, 1982; Theil, 1971; Beach & Mackinnon, 1978).

Inevitably, both problems (autocorrelation and multicollinearity) can coexist in a linear regression model (Hussein, 2009), hence proven in literature (Ayinde *et al.*, 2015; Bayhan & Bayhan, 1998; Lukman *et al.*, 2015; Ozkale & Tugba, 2015; Trenkler 1984; Tugba, 2020; Tugba & Ozkale, 2019). For instance, the generalized ridge estimator (Trenkler, 1984), the feasible generalized ridge estimators (Bello *et al*., 2017) and the combination of the generalized least squares and the ridge regression (Hussein & Zari, 2012) to mitigate both problems. The design of the feasible generalized ridge (FGR) estimator to correct for both problems (Eledum & Zahri, 2013). The combination of the Liu estimator and the feasible generalized least squares to develop the feasible generalized Liu (FGL) estimator (Dawoud & Kaçıranlar, 2016).

Quantile regression (QR) analysis is used to overcome unsatisfactory assumptions, such as autocorrelation, no multicollinearity, normality assumptions and variance homogeneity (Yanuar, Yozza, Firdawati, Rahmi, & Zetra, 2019). QR estimates the conditional median (or other quantiles) of response variable, whereas OLS estimate the response variable's conditional

mean across predictor variables. When estimating the various quantiles of a conditional distribution, QR is used. QR models outperform linear regression (mean regression) models in terms of robustness to outliers (Reed, 2011). Regularization (for instance, Ridge and Lasso) in quantile regression has been proven to improve prediction accuracy (Bager, 2018; Li & Zhu, 2008; Li, Lin, & Xi, 2010; Suhail, Chand, & Kibria, 2020; Wu & Liu, 2009). As a result, it is vital to explore more regularized and robust regression methods in handling multicollinearity and autocorrelation problems, hence, the need for this study.

Statement of the Problem

The general linear regression model makes the fundamental assumption that there is no correlation (or multicollinearity) between the predictors and that there is no autocorrelation. When assumption of no correlation (or multicollinearity) between predictor variables is violated, OLS estimates become unstable, have large variances, and may have an incorrect sign (Greene, 2012). Furthermore, when the multicollinearity degree gets higher, the OLS estimate becomes imprecise, the model may have insignificant tests, wider confidence intervals, parameters with less interpretability, and the OLS being the BLUE does not hold anymore. The existence of multicollinearity makes estimating the unique effects of distinct variables in the regression model unfeasible.

Again, when the assumption of no autocorrelation is not met, the OLS estimator, although linear and unbiased no longer have minimum variance among all linear unbiased estimators (Gujarati, 2003). The OLS estimates are consistent and unbiased, even with correlated error terms. The problem is the efficiency of these estimates. Forecasts based on OLS in the presence of

autocorrelation will be unbiased, but inefficient due to inefficient estimates of the regression parameters. The estimated variance of the regression parameters will be biased, resulting in an erroneous test of hypotheses. Due to the fact that $cov(\mathbf{e}_t, \mathbf{e}_s) \neq 0$, the BLUE that minimizes variance will not be the same as OLS estimate. Therefore, OLS estimator is not BLUE and hence is inefficient.

The traditional approaches of dealing with multicollinearity problems in regression analysis have some limitations. When dropping problematic variables from a model, it can be difficult to decide which variable(s) to be dropped from the model. However, the dropped variable(s) may reduce model's predictive power, and sometimes the variable(s) may be too significant to leave out of the analysis. It may also lead to a problem of bad-specification, specifically, specification bias or specification error, where in some case there is no assurance of whether the model will exhibit less multicollinearity (Gujarati & Porter, 2009). The approach of increasing the number of observations may help in reducing the sampling variance of estimates, but economic constraints may also not allow increasing the number of observations by collecting additional data. Also, the addition of new data may not match with the earlier data collected and may be unusual. Increasing the sample size, according to Chatelain and Ralf (2014), may result in a false inference of strongly correlated classical suppressors. The technique of merging time series and cross-sectional data may create interpretation problems of regression models (Gujarati & Porter, 2009).

Our interest is to squeeze out maximum information from the multicollinearity and autocorrelation data at our disposal and this has motivated the study of the performance of some regression methods. This study, therefore,

suggests a suitable alternative to the least squares estimator to mitigate both problems. One approach in addressing the problem is to apply regularization to robust regression methods. It is therefore relevant to explore from a different perspective by combining a robust regression method and regularized methods. We therefore make a comparative study to evaluate performances of different regularized and robust regression methods when multicollinearity and autocorrelation problems are present under conditions of different degrees of autocorrelation, different sample size, different number of predictors and different multicollinearity levels.

Purpose of the Study

The main purpose of the study is to make a comparative analysis of different regularized and robust regression methods when autocorrelation and multicollinearity problems are present in a dataset. The aim is to determine if regularized and robust regression methods reduce total model error and which of the methods under consideration is the most effective in handling autocorrelation and multicollinearity problems.

Research Objectives

Specifically, we seek to:

1. Compare the performance of Ridge regression, Lasso regression, Two-Stage Ridge regression and Two-Stage Lasso regression methods as remedies to OLS method in the presence of autocorrelation and multicollinearity problems.
2. Develop the Two-Stage Ridge Quantile regression and Two-Stage Lasso Quantile regression methods.

3. Compare the performance of Two-Stage Ridge Quantile regression and Two-Stage Lasso Quantile regression with Ridge Quantile regression and Lasso Quantile regression on multicollinear and autocorrelated data. Objectives (1) and (3) are carried out under conditions of different number of predictor variables, different sample size, different multicollinearity levels and different degrees of autocorrelation.

Significance

The findings of the study are expected to make a significant contribution to the development of suitable regularized and robust regression methods for handling autocorrelation and multicollinearity problems in linear model. The findings of the study will help us to understand the degree of multicollinearity for which OLS should be preferred over regularized regression and robust regression methods.

The study will suggest the best estimation method that gives the least mean squared error under multicollinearity and autocorrelation conditions for a given number of predictor variables, at different sample sizes, for different multicollinearity levels and different autocorrelation degrees. This will further help to know the effect of sample size and samples to be collected from a highly homogeneous population with multicollinearity problem when a particular regression method is to be used.

The study will be helpful to future researchers by providing them views about the best regression estimator to handle multicollinearity and autocorrelation problems.

Delimitations

This study focuses on solving autocorrelation and multicollinearity problems in linear regression analysis using regularized and robust regression methods for real and simulated data. The study focused on the use of regularized and robust methods on a simulated dataset to compare their performance at different multicollinearity levels, different degrees of autocorrelation, different number of predictor variables and different sample sizes.

The difficulty in getting real data with different degrees of multicollinearity, different degrees of autocorrelation, different number of predictor variables and different sample sizes resulted in the used of data simulations to help us compare the performance of the regularized and robust regression methods.

Firstly, comparison is made between the regularized regression methods as single estimators to evaluate their performance on multicollinearity and autocorrelation dataset using MSE criteria. Secondly, comparison is made between the combinations of different regularized regression methods and quantile regression method as combined estimators, to know which one performs better on multicollinearity and autocorrelation dataset.

Limitations

The study focused on multicollinearity and autocorrelation problems regardless of other issues, such as outliers and heteroscedasticity problems. Again, since this is a simulation study, data has been generated in an attempt to enable for generalization to practical circumstances. Also, with regards to Quantile Regression method used in combination to regularized methods, there are other possible members of the robust regression method that may be used to construct the title of combined regularized and robust regression methods.

Organisation of the Thesis

The thesis is organized into five chapters. Chapter One, the introduction of the study highlights the background, statement of the problem, purpose of the study, research objectives, significance, delimitations and limitations of the study. Chapter Two presents the literature review. It covers both theoretical frameworks and empirical reviews related to the concepts under study. Chapter Three deals with the research methodology employed to accomplish the study. The chapter discussed the mathematical and statistical methods and procedures used in the data analysis. Chapter Four covers the results and discussions of the study. The final chapter, Chapter Five, covers the summary, conclusions and recommendations, as well as suggestions for further studies.

Chapter Summary

This chapter described multicollinearity and autocorrelation, and highlighted the implications of multicollinearity and autocorrelation on the regression coefficient of OLS estimators. When multicollinearity and autocorrelation are present, the performance of the OLS estimate was found to be unsatisfactory. The traditional approaches of handling autocorrelation and multicollinearity problems in regression model have some limitations, and there are times when none of these approaches are satisfactory. We give the main motivation that has been captured in the problem statement together with the research objectives. Therefore, this study introduces regularized and robust regression methods to address the problems of autocorrelation and multicollinearity. We concluded the chapter by presenting an outline for the thesis.

CHAPTER TWO

LITERATURE REVIEW

Introduction

This chapter discusses the challenges multicollinearity and autocorrelation pose to linear models. The chapter delves into some works done by other researchers and authors in addressing multicollinearity and autocorrelation problems. The chapter is structured as follows; overview of OLS, nature, source and effect of multicollinearity and autocorrelation, test for multicollinearity and autocorrelation, solutions to models with multicollinearity and autocorrelation, regularization and regularized regression methods, regularized estimators in robust regression methods, and literature on addressing multicollinearity and autocorrelation problems.

Overview of Ordinary Least Squares

The OLS estimate is considered as Best Linear Unbiased Estimator (BLUE). It is useful for investigating the linear relationships between variables of interest. OLS regression is based on assumptions, and when those assumptions hold true, the OLS regression produces the best estimates. When the assumptions are met, the OLS generates better estimates than any other linear model estimating methods, according to the Gauss-Markov theorem (Greene, 2003; Hastie, Tibshirani, & Friedman, 2009). One of the most important goals of regression analysis is to explain the relationship between predictors and response variable.

The effectiveness of regression analysis is highly dependent on the structure of correlations between predictive variables. If the predictors are orthogonal, OLS estimator is optimal among the class of linear unbiased estimators.

Multicollinearity develops when the regression model contains highly correlated predictors. Again, when the uncorrelated error terms assumption is violated, autocorrelation occurs.

Autocorrelation leads to inefficiency of the OLS estimates. Consider the multiple linear regression model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}$$

where $\boldsymbol{\beta}$ is a $(p+1)\times 1$ vector of unknown parameters and $\boldsymbol{\varepsilon}$ is an $n\times 1$ vector of random errors with $\mathrm{E}(\mathbf{e}) = 0$ and $\mathrm{var}(\mathbf{e}) = \sigma^2\mathbf{I}_n$. The response variable is arranged in the $n\times 1$ vector $\mathbf{y}$ and the data for the predictor variables are in the $n\times(p+1)$ matrix $\mathbf{X}$. Multicollinearity violates the assumption that the design matrix $\mathbf{X}$ is given the full rank, rendering OLS estimation unfeasible. When a model does not have full rank, inverse of matrix $\mathbf{X}$ cannot be determined.

Nature of Multicollinearity and Autocorrelation

Multicollinearity develops when a linear relationship exists between two or more than two variables; in contrast to multicollinearity, collinearity can refer to either the general scenario of a linear dependence among predictors or a linear relationship between only two predictors. The nature of multicollinearity can be classified into perfect (or exact) and imperfect (or near) multicollinearity. If the predictor variables in a model exhibit perfect correlation, statistical software will fail to fit the model and produce an error notice. Statistical software, on the other hand, can fit OLS regression models with imperfect but strong relationships between predictors. However, correlations can cause issues if they are strong enough.

In the case of perfect multicollinearity, the matrix $\mathbf{X}$ and matrix $\mathbf{X}^T\mathbf{X}$ lack full rank. Consequently, the inverse matrix $(\mathbf{X}^T\mathbf{X})^{-1}$ cannot be computed, therefore $\hat{\boldsymbol{\beta}}_{OLS} = \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{y}$ cannot be solved and the OLS estimator has no unique solution (Flexeder, 2010; Hashem, 2014). In the case of imperfect multicollinearity, the matrix $\mathbf{X}^T\mathbf{X}$ and matrix $\mathbf{X}$ have full rank, however, it is not far from being rank-deficient. The inverse of matrix $(\mathbf{X}^T\mathbf{X})^{-1}$ can be computed, therefore $\hat{\boldsymbol{\beta}}_{OLS} = \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{y}$ can be solved and the OLS estimator has a unique solution. However, the computed OLS estimator possesses a very large variance. A high multicollinearity level can either result in an erroneous inversion or prohibit the statistical software from completing a matrix inversion while calculating the regression coefficients.

The importance of multicollinearity stems from the fact that, estimates of OLS regression coefficients may vary unpredictably in response to slight changes in predictors. As a result, in linear regression analysis, the ability to detect multicollinearity is critical. Such detection entails two interconnected steps: first determining the existence of a multicollinearity problem, and then determining its severity or strength.

The classical model assumes that the error term associated with any observation is unaffected by any other observation's error term. Autocorrelation occurs when an important assumption that all error terms for all observations are independent of each other, is violated during an OLS regression analysis. It can cause issues in conventional analyses that assume observation independence such as OLS regression. Because it refers to a lack of independence between values, autocorrelation is problematic for most statistical

tests. First-order autocorrelation, which can be positive or negative, is the most common type of autocorrelation.

Sources and Effects of Multicollinearity and Autocorrelation

There are two basic sources of multicollinearity, that is, Structural and Data-based multicollinearity. When a researcher creates new predictor variables or a model term using other terms, structural multicollinearity occurs. In other words, it is a result of the model specified by the researcher rather than being present in the data itself. It results from the use of predictor variables that are powers or linear combinations of another set of variables. For instance, if term X is squared to model curvature by researcher, there is clearly a relationship between term X^2 and X. Multicollinearity may increase with any combination of the original variables if the sampling subspace for the predictor variables is narrow. Data-based multicollinearity is inherent in the data rather than results of researcher's model. Data-based multicollinearity is caused by poorly designed experiments, observational data issues, or non-manipulable data collection methods, and is more common in observational experiments. It also occurs in data collection methods when the data is assembled from a small subspace of the predictor variables. This is multicollinearity caused by poor sampling methodology.

Researchers face issues when predictors are highly, but not perfectly, correlated. One issue with data multicollinearity might be that small variations in the dataset generate large swings in estimations of parameters. Multicollinearity also results in coefficients having very high standard errors. Due to multicollinearity, coefficients may result in implausible magnitudes or "wrong" sign. The confidence interval for coefficient estimators will be wider

due to multicollinearity (Mela & Kopalle, 2002). Increased standard error of the coefficients may result in some predictor variables being insignificant (Gujarati & Porter, 2009), whereas these same coefficients could have been found to be significant with lower standard errors and no multicollinearity (Greene, 2012).

Although multicollinearity reduces the model's power to identify statistically significant predictor variables and also makes it difficult to interpret coefficient of regression estimates, it has no effect on predictions, precision of predictions, or goodness-of-fit statistics (Montgomery, Peck, & Vining, 2012).

There are several sources of autocorrelation. Autocorrelation can occur whenever there is some ordering of sampling units. The effect of removing some variables is another source of autocorrelation. Mis-specifying the type of relationship can introduce autocorrelation into the data. If there are exponential or log terms in the model, the model's linearity is put into question, and this may cause autocorrelation in the data. Carryover of effect is a significant source of autocorrelation, at least in part. Autocorrelation can also be caused by systematic measurement errors.

Ordinary regression suffers when the independent errors assumption is violated. The OLS estimates of the regression coefficients are still consistent and unbiased when the error terms are autocorrelated, but the minimum variance property is not satisfied. R-squared will be overestimated in the majority of cases and hypothesis testing is no longer valid. The traditional F and t significance tests are no longer valid, and if used, are likely to produce significantly misleading results with respect to the statistical significance of the computed regression parameters. Another effect is the OLS estimates are not asymptotic when the error terms are autocorrelated. Because the OLS estimates

will be inefficient, they will no longer be BLUE (Damoder, 2006; Greene, 2003; Gujarati, 2003).

Test for Multicollinearity

A variety of methods can be used to detect multicollinearity. Some of the most common methods are discussed below.

Correlation Matrix

A correlation matrix is a table that illustrates the correlation coefficients among variables. Multicollinearity can be determined by inspecting a predictor variable correlation matrix for high correlation.

Variance Inflation Factor (VIF)

The VIF measures how much a given predictor's linear correlation with the other predictors raises the variance of its coefficient estimate in comparison to the baseline case of no correlation, and is defined as

$$VIF_i = \frac{1}{1-R_i^2}$$

where R_i^2 is the coefficient of determination of a regression of predictor i on all the other predictors. A VIF of 5 or 10 and above implies a multicollinearity issue (O'Brien, 2007). The denominator $1-R_i^2$ is known as tolerance.

Eigenvalues, Condition Indexes and Condition Number

For eigenvalues, multicollinearity is indicated by smaller eigenvalues of $\mathbf{X}^T\mathbf{X}$ or its related correlation matrix. One or more small eigenvalues that are smaller than other eigenvalues indicate that the variables have near-linear dependencies.

The condition indexes (CI) of the $\mathbf{X}^T\mathbf{X}$ matrix is defined as,

$$CI_i = \frac{\max(\lambda_i)}{\lambda_i},\ i = 1,2,\cdots p$$

Where λ_i are eigenvalues. The number of near-linear dependences in $\mathbf{X}^T\mathbf{X}$ is determined by the number of "large" condition indexes. Condition indexes ranging from 5 to 10 shows weak dependencies, whereas condition indexes ranging from 30 to 100 shows moderate to strong dependencies (Belsley, 1991).

The condition number (CN) of the $\mathbf{X}^T\mathbf{X}$ matrix is defined as,

$$k = \frac{\max(\lambda_i)}{\min(\lambda_i)},\ i = 1,2,\cdots p$$

Where $\max(\lambda_i)$ and $\min(\lambda_i)$ are maximal and minimal eigenvalues respectively, and λ_i are the eigenvalues. The condition number is most commonly used to determine whether the inversion of $\mathbf{X}^T\mathbf{X}$ may result in numerical issues. From no serious, moderate to strong, and severe multicollinearity problems, respectively, are associated with condition numbers less than 100, between 100 and 900, and greater than 900 (Belsley, 1991).

Farrar χ^2

Farrar χ^2 is a Chi-square test for detecting the degree of multicollinearity across the entire set of predictors and is defined as $\chi^2 = -\left(n - 1 - \frac{1}{6}(2p+5)\right)\log\left(\left|\mathbf{X}^T\mathbf{X}\right|\right) \sim \chi^2\left(-\frac{1}{2}p(p-1)\right)$.

The multicollinearity exists among predictors if $\chi^2 > \chi^2\left(-\frac{1}{2}p(p-1)\right)$ (Farrar & Glauber, 1967).

Klein's rule

According to Klein's rule, multicollinearity is a major issue if the degree of inter-correlation is greater than the overall multiple correlation degree. That is, $R_i^2 > R_y^2$, where R_i^2 is from a regression of predictor variable i on all of the others and R_y^2 is from a regression of response variable y on all predictor variables (Klein, 1962).

Corrected VIF (CVIF)

The corrected VIF (CVIF) as a new multicollinearity index was defined by Curto and Pinto (2011):

$$CVIF_i = VIF_i \times \frac{1-R^2}{1-R_o^2}$$

Where $R_o^2 = R_{yx1}^2 + R_{yx2}^2 + \cdots + R_{yxp}^2$, and R_{yx1}^2, R_{yx2}^2, R_{yx3}^2, $\cdots$, R_{yxp}^2 are the coefficient of correlation between **y** and x_i with $i = 1,2,\cdots p$. The corrected VIF measure is used to evaluate the influence of the predictor variables correlation on the variance of the OLS estimators. Collinearity exists if $CVIF_i \geq 10$.

Test for Autocorrelation

There are several methods for detecting autocorrelation, the most common of which are the graphical method and the Durbin-Watson test.

Graphical Method

The estimated residuals are plotted against time in a graph. The residual plot's clustering effect is used to determine the presence of autocorrelation. The residuals should be randomly distributed around the zero line, if the data are independent. However, if a discernible pattern emerges (especially one that is

cyclical), dependency is most likely present, or if successive residuals tend to cluster on one side or the other of the zero line, it is a pictorial sign that autocorrelation exists.

Durbin-Watson Test

The Durbin-Watson test only takes into account first-order autocorrelation. Durbin-Watson can test the hypothesis:

$H_o : \rho = 0$ (The errors are not correlated)

$H_A : \rho \neq 0$ (The errors are correlated)

The Durbin-Watson test statistic is defined as follows;

$$D = \frac{\sum_{t=2}^{n}(e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2}$$

where $e_t = y_t - \hat{y}_t$ are the residuals from the OLS fit. The Durbin-Watson tests produces statistic ranging from 0 to 4, where values ranging from 0 to 2 indicate positive autocorrelation, 2 indicate zero autocorrelation, and values ranging from 2 to 4 indicate negative autocorrelation.

Solutions to models with Multicollinearity and Autocorrelation

Multicollinearity issues can be addressed by removing a redundant or problematic variables from the model, increasing the number of observations or adding new data, transforming or recoding multicollinear variables, incorporating variables into scale or centering if multicollinearity arises from cross-product terms, and using regularization methods (Demaris, 2004; Giacalone, Panarello, & Mattera, 2018; Gujarati & Porter, 2009). When multicollinearity problem exists in a regression model, it is critical to determine

which predictor variables are causing the problem, and the simplest approach is to remove the model variables that are problematic. Also, if a variable is redundant, dropping it is simply correcting a specification error. Also, to simply correct a specification error, remove redundant variables from the model. Stepwise regression procedures are formal methods that accomplish essentially the same thing.

Increasing the sample size improves an estimator's precision and reduces the negative effects of multicollinearity. Even if the variance of the OLS estimates is inflated by high correlation among predictors, the variance tends to decrease; additionally, a larger sample size reduces over-fitting and tends to reduce R-squared. Multicollinearity problems are reduced by re-specifying the model, such as by combining multicollinear variables. The approach is effective in controlling multicollinearity when it was caused by experiment design or model specification rather than a sampling problem.

Regularization methods such as Ridge and Lasso, are frequently employed to "solve" the problem of multicollinearity. Regularization methods add restriction information on the value of β or some combination of regression coefficients. These restrictions can be applied to mitigate the effects of multicollinearity. Principal Component Analysis is used in Principal Component Regression (PCR) to convert the original variables into a new set of uncorrelated linear variables. Principal component analysis is also employed in Partial Least Squares (PLS) regression to generate a set of uncorrelated components to include in the model.

Autocorrelation issues can first be addressed by investigating the absence of any other key predictor variable. The variables are then transformed when such

predictor variable cannot be identified to eliminate autocorrelation from the model. The dataset is transformed so that the resulting modified model's error term conforms to the OLS assumption of no autocorrelation. One popular technique is to simply proceed with a first difference model. Redesigning the variables with a Generalized Least Squares (GLS) or Feasible Generalized Least Square (FGLS) methods such as the Prais-Winsten estimator, Cochrane-Orcutt estimator to address autocorrelation problems (Cochrane & Orcutt, 1949; Paris & Winstein, 1954).

Regularization and Regularized Regression Methods

Generally, the OLS estimate performs poorly when there is multicollinearity. However, the linear model suffers from high variance while having a low bias. This is where a popular Goldilocks solution called "regularization" and sometimes called "shrinkage" come to rescue (Berk, 2020; Hastie, Tibshirani, & Friedman, 2009). They are approaches to addressing the bias-variance trade-off in the context of statistical learning. Shrinkage, on the other hand, is regarded as a subset of regularization in which the loss function of ordinary linear regression is modified to incorporate a complexity penalty (Berk, 2020; Hastie, Tibshirani, & Friedman, 2009).

Regularization is essential in modern data analysis. This allows for a reduction in variance at the expense of some bias, ultimately lowering the total error of the model. The process regularizes regression coefficients and, as a result, the fitted values. Conventional regression methods cannot handle data with more predictors (p) than observations (n) (i.e. $p >> n$). Some of the coefficients are reduced to zeros using these methods. When there are a large number of parameters, regularization is also required to avoid over-fitting.

Regularized regression methods for linear regression have evolved to outperform the shortcomings of OLS regression in terms of prediction precision. Regularizing the model is a way of reducing undesirable effects or structural parameters where weights are constrained. Regularization methods are also used as a variable selection method for determining the most important variables in a model. They are used to control the regression coefficients, which can reduce variance and sample error and thus solve the multicollinearity problem.

In general, two main techniques are employed in improving the OLS estimation procedure. One method focuses on identifying biased estimators with lower MSE than OLS estimators. Lasso regression and Ridge regression are examples. The second technique, on the other hand, deals directly with the predictor variables' dependency nature, such as PLS regression and PCR. There have been studies comparing regularized models (Gnat, 2020; Melkumova & Shatskikh, 2017).

According to a study conducted by Gnat (2020), regularization enhances outcomes and allows for reduced average absolute percentage errors. In the study, two regularization methods (Ridge regression and Lasso regression) were used in assessing their impact on errors. The study also confirms the effectiveness of regularization in removing the problem of model overfitting or multicollinearity. According to studies from literature, the less data you have, the more beneficial it is to use regularization (Gnat, 2020).

Regularized estimators in Robust Regression Methods

According to Giacalone *et al*. (2018), some robust estimators are preferred in the remedy to problems of multicollinearity, because it handles both multicollinearity and outliers simultaneously. Since OLS is highly sensitive to

outliers, robust regression methods are intended to address some of the shortcomings of OLS estimate when outliers are present. Addressing these issues, there have been several robust regression estimators developed, such as Huber's M-estimators (Huber, 1964), S-estimators (Rousseeuw & Yohai, 1984), Least Median of Squares estimators and Least Trimmed Squares estimators (Rousseeuw, 1984), MM-estimators (Yohai, 1987), Robust and Efficient Weighted Least Squares estimators (Gervini & Yohai, 2002), τ-estimators (Yohai & Zamar, 1988), and Quantile regression methods (Koenker & Bassett, 1978). Other methods, such as the Bisquare Ridge LTS estimators (BRLTS) (Pati *et al.*, 2016) and Ridge Least Absolute Value (RLAV) (Pfaffenberger & Dielman, 1985), have combined regularized methods and robust regression methods, which should be preferred, according to Giacalone *et al*. (2018).

The majority of techniques in literature are centered on mean regression, and outliers are particularly sensitive to mean-centered regressions (Aguinis, Gottfredson, & Joo, 2013). Regularization (for example, Ridge and Lasso) in Quantile regression (QR) has been shown to improve prediction accuracy (Bager, 2018; Li & Zhu, 2008; Li, Lin, & Xi, 2010; Suhail, Chand, & Kibria, 2020; Wu & Liu, 2009). Koenker (2004) is the first to use regularization in QR, where the Lasso penalty on random effects was applied in a mixed-effect QR model to shrink individual effects towards a common value. Bayesian technique by incorporating an L_2-penalty into the check function of standard QR to shrinking random effects to a common value was also proposed by Yuan and Yin (2010). Yu, Alkenani, and Alhamzawi (2012) proposed adaptive Lasso quantile regression and a flexible Bayesian Lasso approach. The proposed

methods were compared to Lasso quantile regression (LQR), Flexible Bayesian quantile regression (FBQR), and QR methods. They found that the proposed methods outperform the alternative methods.

Zaikarina, Djuraidah, and Wigena (2016) researched on "Lasso and ridge quantile regression using cross validation to estimate extreme rainfall". They used quantile regression with Lasso and ridge regularizations and modified percentile cross validation. Local monthly rainfall and precipitation data from the "Global circulation model (GCM)" were used in the study at Indramayu, Indonesia, from 1981 to 2013. They found that the modified percentile method criteria for selecting optimal Lasso and ridge coefficients provide good prediction on both Lasso and ridge quantile regressions. Their results showed that quantile regression with ridge regularization performed better than Lasso. Adlouni, Salaou, and St-Hilaire (2018) considered five penalties (Lasso, Ridge, SCAD0, SCAD1, SCAD2, Ridge and Lasso) combined with quantile regression in Bayesian framework. They found that the Lasso for quantile estimation had the best performance according to the Relative Mean-Error (RME) and Relative Mean Bias (RMB) criterion, particularly when dealing with heavily distributed errors.

Related Works on addressing Multicollinearity Problems

In a study conducted by Irfan, Javed, and Raza (2013), Ordinary Least Squares (OLS), Principal Component Regression (PCR), Partial Least Squares (PLS) and Ridge Regression (RR) methods were compared to address the multicollinearity problem on Pakistani GDP data from 1973 to 2011. On the basis of the Cross Validation Parameter (CVP), Root Mean Square Error Cross Validation (RMSECV), RMSE and R-Squared, the PLS regression approach

was discovered to produce superior data findings than the other regression methods. In a study performed by El-Fallah and El-Salam (2014), PLS regression, PCR and RR methods were compared in their study on multicollinearity. The effectiveness of these three regression methods was assessed using Monte Carlo simulation study. The MSE criterion was used to assess the efficiency of the three regression methods. The simulated data used consist of p = 2, 4, 6 and 25 predictor variables for n = 20, 30, 40, 50 and 60 sample sizes. Their findings show that the RR performs best with small number of predictors (p = 2), while the PLS performs best with moderate number of predictors (p = 4, 6) and high (p = 25). They also found that, PLS and RR outperformed PCR in all cases.

Firinguetti, Kibria, and Araya (2017) carried out a simulation study to compare Ridge regression and PLS to OLS. The findings of their study revealed that both the RR and PLS methods result in significant reductions in MSE over the OLS method. They also found that as multicollinearity increases, neither method appears to be preferred over the other. However, with large error variance, the RR performs better, and with model that contains more variables, the PLS method works best. Goktas and Akkus (2020) compared the performance of RR, PLS and PCR methods based on MSE values in simulated research. Simulations were run for six different levels of multicollinearity (0.0, 0.3, 0.5, 0.7, 0.9 and 0.99), six different sample sizes (30, 50, 100, 200, 500 and 1000) and three different number of variables (4, 7 and 9). The simulation was run 10,000 times with three different number of predictors (4, 7 and 9). Their results show that the number of predictors, sample size and level of multicollinearity all have an effect on each prediction method. Based on the

obtained findings, it was observed that when there is severe multicollinearity among predictors, the PCR method produces significantly superior estimates.

Herawati, Nisa, Setiawan, Nusyirwan, and Tiryono (2018) used data simulation to assess the performances of OLS, RR, PCR and Lasso methods in addressing severe multicollinearity problems. The data simulated show severe multicollinearity between predictors ($\rho = 0.99$), five different number of predictor variables ($p = 4, 6, 8, 10, 20$) and five different sample sizes ($n = 25, 50, 75, 100, 200$) with 100 iterations using the R package. The Akaike Information Criterion (AIC) and Average Mean Square Errors (AMSE) criterion were employed to assess the methods performances. They found that the RR and PCR methods can overcome severe multicollinearity problems. Overall, they concluded that PCR is the best method for estimating regression coefficients on data with severe multicollinearity.

Adepoju and Ojo (2018) proposed a Bayesian estimation method for linear regression models with the use of informative priors (conjugate) in existence of multicollinearity. The Bayesian method was compared to that of the OLS method using a simulation study, with Confidence Intervals (CI), Standard Error (SE), and MSE as criterion for evaluation and comparison. The simulated data were generated with two different multicollinearity degrees (0.80 and 0.95) and sample sizes were varied between 30, 200 and 300 with three predictor variables. According to the simulation results, the Bayesian method with an informative prior outperformed the OLS method. As a result, they recommended the Bayesian estimation method for dealing with multicollinearity, especially when there is sufficient prior information and the degree of multicollinearity is high. Another study on Bayesian method with the

application of informative prior was carried out by Jaya, Tantular, and Andriyana (2019). They then used the MSE, Bias and power of the test criterion to reveal that the Bayesian method outperformed the RR method. They also showed that, in multicollinear regression analysis, the Bayesian method can be used to efficiently address hypothesis testing.

Suhail, Chand, and Kibria (2020) developed a new type of ridge estimator based on quantile and compared its performance with other RR and OLS methods in a simulation study. Their results showed that the proposed method outperforms alternative approaches, particularly when error variance and multicollinearity are high. They also showed that increasing the sample size and a large number of predictor variables results in a corresponding reduction in estimator MSE values. Davino, Romano, and Vistocco (2022) researched on "handling multicollinearity in quantile regression through the use of principal component regression". They applied PCR principles to quantile regression by proposing Quantile Principal Component Regression (QPCR). The study used both simulated and empirical data to investigate collinearity's effect on QR. The simulation study employed three predictors and 100 observations with varying degrees of correlation. For each design grid, they ran 1000 simulations. Their results reveal that as collinearity rises, standard errors in OLS increase, while those in QR increase more. However, when PCR is used, the results' stability in the various scenarios is confirmed. The study's main findings are the importance of collinearity in QR and the application of PCR as a viable remedy.

Jegede, Lukman, Ayinde, and Odeniyi (2022) combined the Jackknife Kibria-Lukman (JKL) estimator and the M-estimator to form the Robust Jackknife Kibria-Lukman (RJKL) estimator to address outliers and

multicollinearity. In the simulation, multicollinearity degrees were set to 0.70, 0.80, 0.90, and 0.99, and sample sizes were varied between 50, 100, 200, and 250 with three different values of σ (1, 5, and 10). They included an outlier of 10%, 20%, and 30% in each sample size considered in the simulation study. The simulation was repeated 2000 times, each time with a different predictor variable (p = 3 and p = 7). They also used the Hussein data from Eledum and Alkhaklifa (2012) study for data analysis. The performances of ridge, M-estimator, ridge M-estimator, Kibria-Lukman (KL) estimator, robust KL estimator, JKL estimator, and RJKL estimator were investigated. The RJKL estimator performed well in both the simulation research and the real-life data analysis. Based on the MSE criterion of each regression estimator, they concluded that the RJKL estimator outperforms the other existing estimators.

According to studies in the literature, researchers have devised a combined-estimator technique to handling multicollinearity problem in regression analysis that outperforms the single-estimator approach. For instance, the $r-k$ class estimator which combines ridge regression (RR) and PCR (Baye & Parker, 1984), the Liu estimator which combines the idea of RR and stein estimator (Liu, 1993), the $r-d$ class of estimator which combines the PCR and Liu estimator (Kaciranlar & Sakallioglu, 2001), a new two-parameter estimator which combines the OLS, RR, Liu, and contraction estimators in special cases (Ozkale & Kaciranlar, 2007), and another new two-parameter estimator by combining the OLS, RR, and the Liu estimators in a different manner (Yang & Chang, 2010).

Principal component two-parameter (PCTP) estimator is proposed by Chang and Yang (2012), which combines PCR and another two-parameter estimator

developed by Yang and Chang (2010). Ozkale (2012) developed a general class of estimator called the $r-(k,d)$ class estimator, and the developed estimator was a combination of a number of regression estimators (OLS, RR, contraction, PCR, Liu, $r-d$ and $r-k$ class estimators) into a single-estimator. The superiority of the $r-(k,d)$ class estimator was obtained over the OLS, RR, PCR, $r-d$ class, $r-k$ class, $\hat{\beta}(k,d)$, Liu and contraction estimators under the scalar MSE criterion. The concept of transitioning to a combined-estimator technique from a single-estimator technique is a gradual process aimed at resolving multicollinearity issues in regression analysis models (Ayinde, Alabi, & Nwosu, 2021).

Related Works on addressing Autocorrelation Problems

Beach and MacKinnon (1978) indicated that estimating the regression parameters with first-order autocorrelation, the widely used Cochrane-Orcutt and Hildreth-Lu procedures typically ignore the first observation. As a result, they proposed the maximum likelihood estimator (MLE), takes into consideration the first observation. They explained that the maximum likelihood procedure functions like the Cochrane-Orcutt procedure but seems to be more computationally efficient. They presented Monte Carlo results and concluded that the maximum likelihood procedure is theoretically superior to the conventional ones.

Kramer (1980) investigated the "finite sample efficiency of ordinary least squares in the linear regression model with autocorrelated errors". The OLS, Prais-Winsten, and Durbin estimators were investigated. In 100 or 500 Durbin and Prais-Winsten sample runs, efficiency is calculated analytically for OLS.

The obtained results strongly demonstrated that the Durbin and Prais-Winsten methods outperform OLS.

Tanizaki (2003) compares the Maximum Likelihood estimator (MLE) to the Bayes estimator (BE) for estimating models with first-order autocorrelated errors. The Metropolis-Hastings algorithm and Gibbs sampler algorithm are employed in the Bayes estimator. The study used Monte Carlo studies to compare BE to MLE, with the standard error (SE), arithmetic average (AVE), and RMSE as comparison criterion. The Bayes estimator was found to be more efficient and less biased than the maximum likelihood estimator in the AVE, SE, and RMSE criteria. The results showed the superiority of BE over MLE because BE values for both the variance of the error term and autocorrelation coefficient are closer to true values than MLE values. Overall, the results show that the Bayes estimator outperformed the maximum likelihood estimator.

Desviona and Yanuar (2020) conducted a study on "simulation study of autocorrelated error using Bayesian quantile regression". They compared the ability of the classical QR and Bayesian quantile regression methods to estimate linear models with autocorrelated error problems. The posterior distribution for the Bayesian quantile regression was estimated using Gibbs sampler algorithm by applying Markov Chain Monte Carlo method. The simulated data employed in the study consisted of two predictor variables each generated from N (0,1) with 150 observations. For parameter estimation, the method with the smallest confidence interval and least absolute value of bias is considered best. The Bayesian quantile method, according to the findings, produces lower confidence intervals and absolute bias values than the QR method. They concluded that the Bayesian quantile regression method outperforms the QR method.

Literature on addressing Autocorrelation and Multicollinearity Problems

Eledum and Alkhalifa (2012) developed the generalized two stages ridge regression (GTR) by combining the generalized ridge regression (GRR) with the two-stage procedure (TS) for dealing with linear models that suffer from multicollinearity and autocorrelation. They derived some statistical properties of the GTR biased estimator. Real data example was employed to achieve the study purpose, with the data representing "the product in the manufacturing sector, capital commodities, imported intermediate and imported raw materials in Iraq from 1960 to 1990". The GTR regression estimator was compared to the OLS, RR, GRR, two-stage regression, and two-stage ridge regression estimators. They found that the GTR estimator has the smallest MSE and thus outperforms the other regression estimators based on the MSE criterion of each regression estimate.

Eledum and Zahri (2013) introduced a two-stage ridge regression (TR) for autocorrelated errors and multicollinear data. After adjusting it with the RR, they applied the two-stage regression (TS) using a mixed method. They also derived the statistical properties of the TR biased method and use an application example to compare the TR estimator's performance with OLS, TS, and RR estimators. They concluded that the TR estimator is unbiased, with lower MSE, variance, and bias than the RR estimator. Eledum and Awadallah (2021) also researched the two-stage ridge regression method for dealing with autocorrelation and multicollinearity problems in linear models. The TR method's performance is evaluated using Simulation study, with MSE used as a criterion for evaluation. Three different multicollinearity degrees (0.70, 0.80, and 0.99), three different autocorrelation degrees (0.5, 0.7, and 0.99), and four

different sample sizes (10, 30, 70, and 100) were simulated. The simulation was run 10,000 times. The results showed that the TR method outperforms the RR, and the regularization parameter values for the TR are always less than those for the RR.

Ayinde, Lukman, and Arowolo (2015) conducted a study that combined Principal Component estimator with two Feasible Generalized Least Square Estimators (Cochrane-Orcutt and Maximum Likelihood estimators) to handle problems of autocorrelation and multicollinearity in linear models. The following estimators were compared in performance: OLS, ridge, PC1, PC12, ML, MLPC1, MLPC12, MLPC123, CORC, CORCPC1, CORCPC12 and CORCPC123. The MSE of estimators at each multicollinearity level, autocorrelation, and parameter was ranked in Monte Carlo studies. The degrees of multicollinearity were 0.9, 0.95, and 0.99, while the degrees of autocorrelation were 0.7, 0.8, 0.9, 0.95, and 0.99. The simulation was run 1000 times with four different sample sizes (10, 20, 30, and 50). The findings revealed that the MLPC1 estimator is generally the best, despite the fact that PC1 and CORCPC1 compete favorably with it. Furthermore, the MLPC12 and CORCPC12 are often the best with larger sample sizes.

In research conducted by Dawoud and Kaçıranlar (2016), they proposed the Two-Stage Liu (TL) method to address autocorrelation and multicollinearity issues in linear models. The performance of the TL method over the others was investigated using real data example and Monte Carlo simulation studies. Three different multicollinearity degrees (0.70, 0.80, and 0.90), five different autocorrelation degrees (0.1, 0.3, 0.5, 0.7, and 0.99), and two different sample sizes (20 and 60) with six different values of σ (0.1, 0.5, 1, 4, 9, and 20) were

simulated. In their study, they also used the Portland cement data from Woods, Steinour, and Starke (1932). The MSE criterion was employed to evaluate the following methods performances: TL, OLS, RR, TS, TR, Liu, Stein, and Two-Stage Stein (STS). The real data and simulation results revealed that TL method has a lower MSE than the other methods. However, when the σ increases, the STS method produces similar or better results over TL method.

A study conducted by Arowolo, Adewale and Kayode (2016) compared the Two-Stage Principal Component regression (T-PC) and Two-Stage Partial Least Square (T-PLS) when a linear model suffers from problems of autocorrelated error and multicollinearity. The study used a real-life dataset from the publication of World Bank (2015) on Gross Domestic Product (GDP), export, import, government capital formation and foreign direct investment for the year 1981 to 2013. The RMSE and RMSE cross validation was employed as criterion for comparison of methods. They found that T-PLS and T-PC perform similarly; however, T-PLS has a higher predictive power.

Ozbay, Kaçıranlar and Dawoud (2017) proposed Feasible Generalized Restricted Ridge regression (FGRR) method to address autocorrelation and multicollinearity issues in linear models at the same time. They derived some statistical properties of FGRR estimator and used a Monte Carlo simulation to compare FGRR estimator to others using matrix MSE criterion. The simulated data were generated with three different multicollinearity degrees (0.7, 0.8, and 0.9), and the errors were obtained by either an MA (1) process or an AR (1) process, where ρ is considered as -0.8, -0.4, 0.0, 0.4, 0.8 with $p = 4$ and $n = 20$. They used τ to control the true or false restriction, and it is chosen as 1, 1.05, or 1.10 and remains constant throughout the simulation. The simulation was

replicated 5000 times. The performance of the following estimators was compared: RR, Restricted Least Squares (RLS), Feasible Generalized Restricted Least Squares (FGRLS) and FGRR. In comparison to the other estimators, they found that FGRR estimator has lower MSE values and therefore outperforms the other estimators.

Oyewole and Agunbiade (2020) carried out research on "regression techniques in the presence of multicollinearity and autocorrelation phenomena: a Monte Carlo approach". They proposed the generalized least squares-ridge (GLS-R) regression method by combining GLS and ridge regression to deal with autocorrelation and multicollinearity problems. However, because $\boldsymbol{\Omega}$ is unknown, parameter estimates for the biased GLS-R estimator are impossible to obtain in practice. To approximate those parameters, they combined Durbin's two-step method with RR. The study's data was simulated using Monte Carlo. The performances of four different estimation methods were investigated: OLS, Ridge regression, Lasso regression, and the GLS-R method. According to their findings, the GLS-R regression method has the lowest AIC and MSE value of the four methods. As a result, they found that the GLS-R regression method is more predictive and provides the preferred estimator in estimating all of the model's parameters based on the AIC and MSE criterion, and thus it is preferred over the OLS, ridge, and Lasso regression methods.

Zubair and Adenomon (2021a) proposed the two-stage $K-L$ estimator to address autocorrelation and multicollinearity issues in linear regression models. The performance of the two-stage $K-L$ estimator is evaluated using Monte Carlo Simulation, with MSE used as a criterion for evaluation. Autocorrelation and multicollinearity degrees were both set to 0.6, 0.8, and 0.9, and sample sizes

ranged from 25, 50, 100, 250, and 500. The simulated data used included six predictor variables, and the process was repeated 1000 times. The Prais-Winsten, OLS, Cochrane-Orcutt, Restricted Maximum Likelihood estimator (RMLE) and MLE estimators were compared to the two-stage $K-L$ estimator. They found that the two-stage $K-L$ estimator performs similarly to MLE and RMLE under severe collinearity and autocorrelation conditions. They further found that, under moderate collinearity and severe autocorrelation conditions, whatever the sample size, the two-stage $K-L$ estimator outperforms all other estimators and appears to perform better with larger sample sizes. The study suggested that, to avoid erroneous inferences, the multicollinearity level and autocorrelation degree between variables be considered when estimating regression model parameters.

Another study was conducted by Zubair and Adenomon (2021b) on the "robustness test of the two-stage $K-L$ estimator in models with multicollinear regressors and autocorrelated error term". The estimators were applied to two real data sets with autocorrelation and multicollinearity issues (Hussein data and French economy data). The two-stage $K-L$ estimator was compared to the OLS, Prais-Winsten, Cochrane-Orcutt, MLE and RMLE under the MSE and RMSE criterion. The study's findings revealed that two-stage $K-L$ estimator outperforms the other estimators in both applications to real data. The findings from the real data confirm the findings of the simulation study in work of Zubair and Adenomon (2021a), which also found that two-stage $K-L$ estimator is preferred in fitting a linear model with assumptions of multicollinearity and autocorrelation are violated.

Chapter Summary

The literature has revealed the extent to which multicollinearity and autocorrelation problems pose a challenge to OLS regression. The chapter has described the nature and sources of multicollinearity and autocorrelation. The literature is clear on the effect of multicollinearity and autocorrelation on regression models. The chapter identified ways of detecting multicollinearity and autocorrelation in a linear model. The chapter has also reviewed some approaches in handling autocorrelation and multicollinearity problems. The literature established that the traditional approaches of handling autocorrelation and multicollinearity problems in regression model have some limitations, and there are times when none of these approaches are satisfactory.

Regularization is essential in modern data analysis and therefore, regularization methods were introduced as possible methods in addressing multicollinearity issues. It was noted that regularization in quantile regression improve prediction accuracy. Regularized and robust regression methods have clear advantage over the regularized regression methods, because it handles both multicollinearity and outliers simultaneously. The chapter further reviewed literature on addressing multicollinearity and autocorrelation problems. Simulation studies have mainly been used in studying performances of regression methods in addressing multicollinearity and autocorrelation. The results suggest that the degrees of multicollinearity and autocorrelation have effect on performance of regression methods. Furthermore, sample size and number of predictors were found to have effect on MSE values of regression methods.

CHAPTER THREE

RESEARCH METHODOLOGY

Introduction

In this chapter, we review the methods employed in the analysis of the data. The chapter describes the data and regression methods employed in addressing the problems of multicollinearity and autocorrelation. We review the L_1 and L_2 regularization methods employed to correct for multicollinearity and the two-stage method adopted to correct for autocorrelation. We further present the formulation of the regularized and robust regression methods. Applying these regression methods to linear models with multicollinearity and autocorrelation problems, the goal is to find the regression method that best addresses multicollinearity and autocorrelation problems in a dataset.

Linear Regression Model

A linear regression model describes the relationship between one or more predictor variables, $X_1, X_2, \cdots, X_p$, and a response variable, y. The model relating the response, y, to p predictor variables, $x_1, x_2, \cdots, x_p$, is given by

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} + e_i .$$

Now for n independent observations of y and associated values of x_i, the model becomes

$$\begin{aligned} y_1 &= \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \cdots + \beta_p x_{1p} + e_1` \\ y_2 &= \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \cdots + \beta_p x_{2p} + e_2 \\ &\vdots \\ y_n &= \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \cdots + \beta_p x_{np} + e_n \end{aligned}$$

In matrix form, the system of equations above is written as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e} \qquad (1)$$

where,

$$\mathbf{y}=\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix},\ \mathbf{X}=\begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{12} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix},\ \boldsymbol{\beta}=\begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} \text{ and } \mathbf{e}=\begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}.$$

When the model errors are normally and independently distributed, the maximum-likelihood estimators for the model parameters in multiple linear regression are also least squares estimators. The normal density function for the errors is

$$f(\boldsymbol{e}_i)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{\boldsymbol{e}_i^2}{2\sigma^2}\right)$$

The likelihood function is the joint density of $e_1, e_2, \ldots, e_n$ or $\prod_{i=1}^{n} f(e_i)$. Therefore, the likelihood function is

$$L(\boldsymbol{e},\beta,\sigma^2)=f(\boldsymbol{e}_i)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{1}{2\sigma^2}\boldsymbol{e}^T\boldsymbol{e}\right).$$

Now since we can write $\boldsymbol{e}=\mathbf{y}-\mathbf{X}\beta$, the likelihood function becomes

$$L(\mathbf{y},\mathbf{X},\beta,\sigma^2)=\frac{1}{(2\pi)^{n/2}\sigma^n}\exp\left(-\frac{1}{2\sigma^2}(\mathbf{y}-\mathbf{X}\beta)^T(\mathbf{y}-\mathbf{X}\beta)\right).$$

According to Stuart (2011), if the assumptions of the error terms are met, that is, the $e_i \sim N(0,\sigma^2\mathbf{I}_n)$, then the least squares estimator is the maximum likelihood estimator for β, maximizing

$$\frac{1}{(2\pi)^{n/2}\sigma^n}\exp\left(-\frac{1}{2\sigma^2}(\mathbf{y}-\mathbf{X}\beta)^T(\mathbf{y}-\mathbf{X}\beta)\right) \qquad (2)$$

over β, which is equivalent to maximizing the logarithm of Equation (2) over β: $\ln L(\mathbf{y},\mathbf{X},\beta,\sigma^2)=-\frac{n}{2}\ln(2\pi)-n\ln(\sigma)-\frac{1}{2\sigma^2}(\mathbf{y}-\mathbf{X}\beta)^T(\mathbf{y}-\mathbf{X}\beta)$.

This corresponds to minimising $(\mathbf{y}-\mathbf{X}\beta)^T(\mathbf{y}-\mathbf{X}\beta)=\sum_{i=1}^{n}\boldsymbol{e}_i^2$ since σ is a constant. Therefore, the minimization of the residual sum of squares $\sum_{i=1}^{n}\boldsymbol{e}_i^2$ is the least squares estimate $\hat{\beta}$. Therefore, the maximum-likelihood estimator of β under normal errors is equivalent to the least squares' estimator $\hat{\beta}=(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$.

Review of Ordinary Least Squares (OLS)

The OLS method is used to estimate the regression coefficients in that the residual sum of squares is to be minimised.

It minimizes the least squares loss function

$$\|\mathbf{y}-\mathbf{X}\boldsymbol{\beta}\|^2 \tag{3}$$

The least squares estimate aims to minimize the sum of the square residuals as

$$\begin{aligned}\sum_{i=1}^{n}\boldsymbol{e}_i^2 &= \boldsymbol{e}^T\boldsymbol{e}\\ &= (\mathbf{y}-\mathbf{X}\beta)^T(\mathbf{y}-\mathbf{X}\beta)\\ &= \mathbf{y}^T\mathbf{y}-\mathbf{y}^T\mathbf{X}\beta-\beta^T\mathbf{X}^T\mathbf{y}+\beta^T\mathbf{X}^T\mathbf{X}\beta\end{aligned}$$

By minimizing the errors

$$\begin{aligned}\frac{\partial}{\partial\beta}\left(\sum_{i=1}^{n}\boldsymbol{e}_i^2\right) &= \frac{\partial}{\partial\beta}\left(\mathbf{y}^T\mathbf{y}-\mathbf{y}^T\mathbf{X}\beta-\beta^T\mathbf{X}^T\mathbf{y}+\beta^T\mathbf{X}^T\mathbf{X}\beta\right)\\ &= 0-\mathbf{X}^T\mathbf{y}-\mathbf{X}^T\mathbf{y}+2(\mathbf{X}^T\mathbf{X})\beta\\ &= 0-2\mathbf{X}^T\mathbf{y}+2(\mathbf{X}^T\mathbf{X})\beta\end{aligned}$$

Now, the least squares estimator $\hat{\beta}$ is the solution to

$$\mathbf{X}^T\mathbf{X}\hat{\beta}=\mathbf{X}^T\mathbf{y}$$

Therefore, when $\mathbf{X}^T\mathbf{X}$ is non-singular the least squares estimate can be evaluated directly from the data by

$$\hat{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y} \quad (4)$$

Linear Model Regularization

Regularization is performed by applying weight constraints on the model. These constraints are implemented in various ways by different algorithms. Ridge and Lasso regression are two forms of regularization used in this study. Model weights in regression models are determined by minimizing the model's residual sum of squares (RSS). The key stage for both methods of regularization is the selection of the regularization parameter, k. For ridge regression, a regularization term equal to $k\sum_{j=1}^{p}\beta_j^2$ is added to RSS loss function of Equation (3). The regularization parameter k controls how much of the model should be regularized. If $k=0$, then Ridge regression is simply an OLS. If k is large, all weights are very close to zero, resulting in a flat line passing through the mean of the data. Setting k is thus an important stage in the creation of a model in order to produce good quality outputs.

The model weights are regularized in the case of Lasso regression by adding an expression $k\sum_{j=1}^{p}|\beta_j|$ to loss function of Equation (3). An important feature of Lasso regularization, is the elimination of the least significant variables from models. Assume we have a large dataset (say, 1000 predictors) and some predictors are correlated. Then, both Ridge and Lasso penalty terms are used. In the Ridge regression method, all 1000 predictors are retained in the model, but their coefficients are shrunk toward zero (but not exactly zero) to retain the model's complexity. In contrast, Lasso regression will reduce some of the coefficients to exactly zero, therefore reducing the model's complexity while

improving its prediction accuracy. As a result, by maintaining only the important variables in the model, the Lasso penalty function naturally achieves the optimal model.

Ridge Regression

According to Hoerl and Kennard (1970), the potential instability in employing OLS estimates could be improved by adding a small constant k to the $\mathbf{X}^T\mathbf{X}$ matrix diagonal entries before taking its inverse. When the predictors are highly correlated, the Ridge regression approach is utilized to improve the estimation of regression parameters. Because of the $\|\boldsymbol{\beta}_{OLS}\| \geq \|\boldsymbol{\beta}_{true}\|$ problem in Equation (4) when the matrix $\mathbf{X}^T\mathbf{X}$ is ill-conditioned, Ridge regression's penalty term is designed in such a way that the length of the parameter vector β is restricted.

The estimates for the Ridge regression parameters are obtained by minimizing the residual sum of squares subject to an L_2-penalty on the coefficients.

$$\hat{\beta}_{Ridge} = \min_{\beta}\left\{\sum_{i=1}^{n}(\mathbf{y}_i - \sum_{j=1}^{p}\mathbf{X}_{ij}\beta_j)^2\right\},\ s.t.\ \sum_{j=1}^{p}\beta_j^2 \leq t,\ t \geq 0 \qquad (5)$$

Equivalently, the following minimization problem defines Ridge regression

$$\hat{\beta}_{Ridge} = \min_{\beta}\left\{\sum_{i=1}^{n}(\mathbf{y}_i - \sum_{j=1}^{p}\mathbf{X}_{ij}\beta_j)^2 + k\sum_{j=1}^{p}\beta_j^2\right\},\ k \geq 0 \qquad (6)$$

where $i = 1, 2, \ldots, n;\quad j = 1, 2, \ldots, p$ and the amount of shrinkage is controlled by the regularization parameter, k. The parameter t in Equation (5) is clearly related to the parameter k in Equation (6). This means that for a given value k, there exists a value t for which the estimation Equations (5) and (6) yield the same result.

Rewriting the Equation (6) in matrix form

$$\hat{\beta}_{Ridge} = \min_{\beta} \|\mathbf{y} - \mathbf{X}\beta\|_2^2 + k\|\beta\|_2^2$$

The function can be restated as

$$\hat{\beta}_{Ridge} = \min_{\beta} \left\{(\mathbf{y} - \mathbf{X}\beta)^T(\mathbf{y} - \mathbf{X}\beta) + k\beta^T\beta\right\}$$

Simplifying gives

$$\hat{\beta}_{Ridge} = \min_{\beta} \left\{\mathbf{y}^T\mathbf{y} - 2\beta^T\mathbf{X}^T\mathbf{y} + \beta^T\mathbf{X}^T\mathbf{X}\beta + k\beta^T\beta\right\}$$

Taking the derivative with respect to β, we obtain

$$\frac{\partial}{\partial\beta}\left(\hat{\beta}_{Ridge}\right) = \mathbf{X}^T\mathbf{X}\beta + \mathbf{X}^T\mathbf{X}\beta - 2\mathbf{X}^T\mathbf{y} + 2k\beta$$

$$= 2\mathbf{X}^T\mathbf{X}\beta - 2\mathbf{X}^T\mathbf{y} + 2k\beta$$

Setting the derivative to zero

$$\frac{\partial}{\partial\beta}\left(\hat{\beta}_{Ridge}\right) = 0 \Rightarrow \mathbf{X}^T\mathbf{X}\beta + k\beta = \mathbf{X}^T\mathbf{y}$$

$$\Rightarrow (\mathbf{X}^T\mathbf{X} + k\mathbf{I}_p)\beta = \mathbf{X}^T\mathbf{y}$$

Therefore, the regularized solution is given as

$$\hat{\beta}_{RR} = (\mathbf{X}^T\mathbf{X} + k\mathbf{I}_p)^{-1}\mathbf{X}^T\mathbf{y} \quad (7)$$

where $\mathbf{I}$ is the $p \times p$ identity matrix. By adding $k\mathbf{I}_p$ to $\mathbf{X}^T\mathbf{X}$, this results in a regular and invertible matrix. For strictly positive k Ridge solution in Equation (7) is well-defined despite multicollinearity or $p >> n$ (van Wieringen, 2021).

The Ridge regression method, in contrast to the OLS estimates, is biased. As a result, this regularization method accepts some bias in order to lower the variance and MSE of the estimates and possibly improve the prediction accuracy. As a result, the goal of Ridge regression is to choose a k that is large enough to overcome the problem of near singularity of the matrix of cross-products but not so huge that it introduces a lot of bias. Ridge regression is very popular for non-orthogonal regression issues due to its ability to obtain descent estimates in the presence of multicollinearity (Duzan & Sharif, 2015).

The way one chooses k is an important aspect of Ridge regression. There are many procedures for determining the biasing parameter (k) value. The procedures are classified as either subjective (such as ridge trace, VIF trace, *df* trace, plotting of variance, bias, and MSE) or objective (such as cross-validation, generalized cross-validation, *m*-scale, Information criteria – Bayesian Information Criterion (BIC), AIC) methods. Many different techniques for estimating k have been suggested or proposed by various researchers. To mention a few notable ones, there are: K_{HKB} (Hoerl & Kennard, 1970), K_{LW} (Lawless & Wang, 1976), K_{TH} (Thisted, 1976), K_{DS} (Dwividi & Shrivastava, 1978), K_{AM} (Kibria, 2003), K_{KD} (Dorugade & Kashid, 2010), $K_{4(AD)}$ (Dorugade, 2014). The common ways of doing this is to use the ridge trace graphical method introduced by Hoerl and Kennard (1970), generalized cross-validation and cross-validation method. We used the generalized cross-validation and cross-validation methods in determining the parameter (k).

Lasso Regression

The Lasso is another regularization method which imposes an L_1-penalty on the regression coefficients. The Lasso minimizes the residual sum of squares subject to the sum of the absolute value of the coefficients being less than a constant.

The Lasso estimate $\hat{\beta}$ is defined by

$$\hat{\beta}_{Lasso} = \min_{\beta}\left\{\sum_{i=1}^{n}(\mathbf{y}_i - \sum_{j=1}^{p}\mathbf{X}_{ij}\beta_j)^2\right\},\ s.t.\ \sum_{j=1}^{p}|\beta| \le t,\ t \ge 0 \qquad (8)$$

An equivalent form of the Lasso is

$$\hat{\beta}_{Lasso} = \min_{\beta}\left\{\sum_{i=1}^{n}(\mathbf{y}_i - \sum_{j=1}^{p}\mathbf{X}_{ij}\beta_j)^2 + k\sum_{j=1}^{p}|\beta|\right\},\ k \ge 0 \qquad (9)$$

where $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, p$ and k is the regularization parameter.

In matrix form

$$\hat{\beta}_{Lasso} = \min_{\beta} \|\mathbf{y} - \mathbf{X}\beta\|_2^2 + k\|\beta\|_1 \quad (10)$$

Tibshirani (1996) defines $t \geq 0$ (or k) as a regularized parameter that can be chosen using generalized cross-validation or cross-validation. There is a choice of k for every choice of t that produces the same result.

The Lasso estimation is a convex optimization issue that can be addressed using a quadratic programming algorithm for a given k. Therefore, solving Equation (9) does not provide a closed form expression. When the regularization parameter k is set to a high value, the Lasso solution is usually sparse as a result of the L_1-penalty. Lasso regression has the same bias penalizing effects, variance rewarding effects, and shrinking capabilities as Ridge regression. However, as k increases, more and more coefficients in Lasso regression will be identically zero (Hastie *et al.*, 2009). Lasso regression can be suitable in handling models with high degree of multicollinearity.

Two-Stage Regression Method

When error terms have autocorrelation, the OLS estimate is still unbiased, but it is no longer a minimal variance estimate, making it inefficient. Variable transformation is one method for addressing estimates in linear autocorrelation models. This study looks at the first-order autoregressive structure AR (1). To deal with autocorrelation, we employed a two-stage method. The two-stage method is used to correct for autocorrelation and the regularized method to correct for multicollinearity.

We used the two-stage procedure outlined below.

1. Obtain an estimate of ρ using the OLS residuals. The estimate of ρ is obtained by

$$\rho = \frac{\sum_{i=2}^{n} e_i e_{i-1}}{\sum_{i=1}^{n} e_i^2}$$

2. The ρ obtained from the OLS was used to transform the data.
3. Two-stage estimator $\hat{\beta}_{TS}$ is obtained by applying OLS to the transformed data. The Durbin-Watson test is then used to test whether error terms for the transformed model are uncorrelated.
4. The regularized and robust regression methods was finally applied to the autocorrelation-free transformed data.

Various transformation approaches have been proposed by different authors. There are those who employ the $\mathbf{P}^*$ matrix and those who employ the $\mathbf{P}$ matrix of transformation. The $\mathbf{P}$ transformation matrix is derived from the AR (1) model for the error term, whereas the $\mathbf{P}^*$ transformation matrix uses a lag definition and loses the first observation. The sole difference between $\mathbf{P}^*$ matrix and $\mathbf{P}$ matrix of transformation is how the first observation is handled. The $\mathbf{P}$ transformation matrix keeps the first observation. The $\mathbf{P}$ transformation matrix is obtained by inserting a new first row with $\sqrt{1-\rho^2}$ in the first position and zero in the other positions in the transformation matrix. It has been shown that with large sample size, the difference between the use of the $\mathbf{P}^*$ matrix and $\mathbf{P}$ matrix is minimal, but with small sample size, the difference is significant (Olaomi, 2008). These approaches are categorized into those that utilize $\mathbf{P}^*$ matrix for transformation, for instance the Hildreth and Lu (HILU) and

Cochrane-Orcutt (CORC) methods, and those that use $\mathbf{P}$ matrix for transformation, for instance the Prais-Winstein method (Olaomi, 2008). We therefore adopted the $\mathbf{P}$ transformation matrix.

The $\mathbf{P}$ transformation matrix is given by

$$\mathbf{P}_{n\times n} = \begin{bmatrix} \sqrt{1-\rho^2} & 0 & 0 & \cdots & 0 \\ -\rho & 1 & 0 & \cdots & 0 \\ 0 & -\rho & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

We consider the linear regression model in Equation (1), that is, $\mathbf{y} = \mathbf{X\beta} + \mathbf{e}$. Pre-multiplying both sides of Equation (1) by the $\mathbf{P}$ transformation matrix, we get an equivalent linear model

$$\mathbf{Py} = \mathbf{PX\beta} + \mathbf{Pe} \tag{11}$$

Let $\mathbf{y}^* = \mathbf{Py}$, $\mathbf{X}^* = \mathbf{PX}$ and $\mathbf{e}^* = \mathbf{Pe}$, then $\mathrm{E}(\mathbf{e}^*) = 0$ and $\mathrm{cov}(\mathbf{e}^*) = \sigma^2\mathbf{I}_n$.

The transformation model is given by

$$\mathbf{y}^* = \mathbf{X}^*\boldsymbol{\beta} + \mathbf{e}^* \tag{12}$$

Model (12) satisfies the error assumption $e^* \sim N(0, \sigma^2\mathbf{I}_n)$.

Therefore, the least squares estimate for the Model (12), which is called the Two-stage is

$$\hat{\beta}_{TS} = (\mathbf{X}^{*T}\mathbf{X}^*)^{-1}\mathbf{X}^{*T}\mathbf{y}^* \tag{13}$$

where $\mathbf{y}^* = \mathbf{Py} = \begin{bmatrix} \sqrt{1-\rho^2} & 0 & 0 & \cdots & 0 \\ -\rho & 1 & 0 & \cdots & 0 \\ 0 & -\rho & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix}$

$$\mathbf{X}^{*} = \mathbf{PX} = \begin{bmatrix} \sqrt{1-\rho^2} & 0 & 0 & \cdots & 0 \\ -\rho & 1 & 0 & \cdots & 0 \\ 0 & -\rho & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} 1 & X_{11} & X_{12} & \cdots & X_{1k} \\ 1 & X_{21} & X_{22} & \cdots & X_{2k} \\ 1 & X_{31} & X_{32} & \cdots & X_{3k} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & X_{n1} & X_{n2} & \cdots & X_{nk} \end{bmatrix}, \; k = p-1$$

The matrix $\mathbf{P}$ can be specified such that $\mathbf{P}\Omega\mathbf{P}^{T} = \mathbf{I}$, and Ω must be positive definite such that $\mathbf{P}^{T}\mathbf{P} = \Omega^{-1}$. Then the OLS estimate of the transformed variable $\mathbf{PX}$ and $\mathbf{Py}$ in Equation (11) have all of the optimal OLS properties, and the usual inferences could be true.

Now, $\mathbf{X}^{*T}\mathbf{X}^{*} = \mathbf{X}^{T}\mathbf{P}^{T}\mathbf{PX} = \mathbf{X}^{T}\Omega^{-1}\mathbf{X}$ and $\mathbf{X}^{*T}\mathbf{y}^{*} = \mathbf{X}^{T}\mathbf{P}^{T}\mathbf{Py} = \mathbf{X}^{T}\Omega^{-1}\mathbf{y}$

where Ω^{-1} is given by

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

After obtaining Ω^{-1}, Two-stage is given by $\hat{\beta}_{TS} = (\mathbf{X}^{T}\Omega^{-1}\mathbf{X})^{-1}\mathbf{X}^{T}\Omega^{-1}\mathbf{y}$

We can find $\hat{\Omega}^{-1}$ after estimating ρ, then the Two-stage can be given by

$$\hat{\beta}_{TS} = (\mathbf{X}^{T}\hat{\Omega}^{-1}\mathbf{X})^{-1}\mathbf{X}^{T}\hat{\Omega}^{-1}\mathbf{y} \qquad (14)$$

Now, to correct for multicollinearity, the regularized regression methods are needed. Because the rank of $\mathbf{X}^{*}$ is equal to the rank of $\mathbf{X}$, the multicollinearity in the datasets still affects the Two-stage method. The regularized methods will be required to handle the multicollinearity problem.

Two-Stage Ridge Regression

The two-stage process used to alter the data is now applied to Ridge regression to produce the Two-stage Ridge regression (TR) method. We

replaced $\mathbf{y}$ and $\mathbf{X}$ in Equation (6) by $\mathbf{y}^*$ and $\mathbf{X}^*$, respectively. The solution of the coefficients can be written as

$$\hat{\beta}_{TR} = \min_{\beta} \left\{ \sum_{i=1}^{n} (\mathbf{y}_i^* - \sum_{j=1}^{p} \mathbf{X}_{ij}^* \beta_j)^2 + k \sum_{j=1}^{p} \beta_j^2 \right\}, \; k \geq 0 \tag{15}$$

where $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, p$ and k is the regularization parameter.

In matrix form

$$\hat{\beta}_{TR} = \min_{\beta} \left\| \mathbf{y}^* - \mathbf{X}^* \beta \right\|_2^2 + k \left\| \beta \right\|_2^2$$

Solving Equation (15) has closed form solution, and the rank of $\mathbf{X}^*$ is equal to the rank of $\mathbf{X}$. Therefore, regularized solution is given by

$$\hat{\beta}_{TR} = (\mathbf{X}^{*T} \mathbf{X}^* + k\mathbf{I}_p)^{-1} \mathbf{X}^{*T} \mathbf{y}^* \tag{16}$$

Following Equations (7) and (14), is the Two-stage Ridge regression proposed by Eledum and Zahri (2013)

$$\hat{\beta}_{TR} = (\mathbf{X}^T \hat{\Omega}^{-1} \mathbf{X} + k\mathbf{I}_p)^{-1} \mathbf{X}^T \hat{\Omega}^{-1} \mathbf{y} \tag{17}$$

Two-Stage Lasso Regression

The two-stage procedure which was used to arrive at the transformed data is now applied to Lasso regression to get the Two-stage Lasso regression (TLasso) method. The complete Lasso cost function is convex and non-differentiable. We replaced $\mathbf{y}$ and $\mathbf{X}$ in Equation (9) by $\mathbf{y}^*$ and $\mathbf{X}^*$, respectively. The solution of the coefficients can be written as

$$\hat{\beta}_{TLasso} = \min_{\beta} \left\{ \sum_{i=1}^{n} (\mathbf{y}_i^* - \sum_{j=1}^{p} \mathbf{X}_{ij}^* \beta_j)^2 + k \sum_{j=1}^{p} |\beta| \right\}, \; k \geq 0 \tag{18}$$

where $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, p$ and k is the regularization parameter.

In matrix form

$$\hat{\beta}_{TLasso} = \min_{\beta} \left\| \mathbf{y}^* - \mathbf{X}^* \beta \right\|_2^2 + k \left\| \beta \right\|_1$$

Instead of running the Lasso on $\mathbf{y}$ vector of response variable and $\mathbf{X}$ matrix of predictors, we therefore run it on $\mathbf{y}^*$ vector of response variable and $\mathbf{X}^*$ matrix of predictors.

Quantile Regression Method

The Quantile regression (QR) method is widely used to describe the distribution of a response variable given a set of predictor variables. QR, as introduced by Koenker and Bassett (1978), is an extension of classical least squares estimate of conditional mean models to the estimation of an ensemble of models for different conditional quantile functions when the conditions of linear regression are not met. QR allows estimating the entire distribution of the response variable's conditional quantiles. The QR estimates are more robust against outliers in response variables, which is one advantage of QR over OLS estimate. However, the QR extends beyond this advantage and is useful when dealing with conditional quantile functions. In contrast to traditional linear regression, QR does not make any assumptions about the error distribution or assume homogeneous residual variance. In this regard, it is a more adaptable technique than traditional linear regression. When all of the assumptions of the conventional linear regression model are met, QR should produce the same results as the traditional model (Tareghian & Rasmussen, 2013).

A typical QR model is formulated as

$$Q_{\tau}(\mathbf{y} \mid \mathbf{X}) = \mathbf{X}^T \boldsymbol{\beta}_{\tau} \tag{19}$$

where $Q_{\tau}(\cdot \mid \cdot)$ is the conditional quantile function for the τth conditional quantile with $0 < \tau < 1$, τ determines the quantile level, $\mathbf{X}$ is a vector of

predictor variables and $\boldsymbol{\beta}_\tau$ is a vector of parameters related to the τth QR. QR provides separate models for each conditional quantile τ of interest.

Regression coefficients β_τ can be estimated by

$$\hat{\beta}_\tau = \min_{\beta_\tau} \sum_{i=1}^{n} \rho_\tau(\mathbf{y}_i - \mathbf{X}_i^T \boldsymbol{\beta}_\tau) \tag{20}$$

where $\rho_\tau(\cdot)$ is the check function defined by

$$\rho_\tau(y) = \left[(1-\tau)\mathrm{I}(y<0) + \tau \mathrm{I}(y>0)\right]|y|$$

A weighted sum of absolute deviations is used as the check function, where a τ weight is used for the positive deviations and a $(1-\tau)$ weight is assigned to the negative deviations. As a result, QR estimates for Equation (20) can be formulated as

$$\hat{\beta}_\tau = \min_{\beta_\tau} \sum_{y_i < \mathbf{X}_i^T \beta_\tau} (1-\tau)\left|\mathbf{y}_i - \mathbf{X}_i^T \boldsymbol{\beta}_\tau\right| + \sum_{y_i \geq \mathbf{X}_i^T \beta_\tau} \tau \left|\mathbf{y}_i - \mathbf{X}_i^T \boldsymbol{\beta}_\tau\right|$$

QR allows the vector $\hat{\beta}_\tau$ to variate on different τ, the median case ($\tau = 0.5$) being equivalent to minimize the sum of absolute values of the residuals. We employed 0.25, 0.5 and 0.75 quantile levels in the analysis of the QR models.

Regularized and Quantile Regression Methods

The Ridge type and Lasso type penalties have been proposed in the literature to improve the prediction accuracy and interpretability of statistical models in both regularization and variable selection (Hoerl & Kennard, 1970; Tibshirani, 1996; Wang, Li, & Jiang, 2007). The presence of multicollinearity causes large variance, which leads to poor inference and prediction. As a solution to the multicollinearity problem, QR with the Ridge penalty has been proposed (Kibria, 2003). QR with Ridge and Lasso regularization is developed by a QR model using Ridge and Lasso coefficients. We therefore employed QR with

Ridge and Lasso regularization to develop QR models to study their performances on multicollinearity and autocorrelation datasets.

Ridge Quantile Regression Method

The Ridge Quantile Regression (RQR) use the ridge coefficients to build the QR model. The RQR is achieved by adding L_2-penalty to the quantile loss function.

The RQR estimate β is given by

$$\hat{\beta}_{RQR} = \min_{\beta}\left\{\sum_{i=1}^{n}\rho_{\tau}(\mathbf{y}_i - \mathbf{X}_i^T\beta_{\tau}) + k\sum_{j=1}^{p}(\beta_j)^2\right\},\ k>0 \tag{21}$$

where $i=1,2,\ldots,n;\quad j=1,2,\ldots,p$ and k is the ridge parameter.

Lasso Quantile Regression Method

Li and Zhu (2008) introduced Lasso Quantile Regression (LQR) for simultaneous estimation and variable selection in QR models. The L_1-penalty is added to the quantile loss function to formulate the LQR method.

The LQR estimate β is given by

$$\hat{\beta}_{LQR} = \min_{\beta}\left\{\sum_{i=1}^{n}\rho_{\tau}(\mathbf{y}_i - \mathbf{X}_i^T\beta_{\tau}) + k\sum_{j=1}^{p}|\beta_j|\right\},\ k>0 \tag{22}$$

where $i=1,2,\ldots,n;\quad j=1,2,\ldots,p$ and $k>0$ is the regularization parameter controlling the amount of penalty. The second term in Equation (22) is the L_1-penalty, which is required for the Lasso to succeed. As the regularized quantile loss function is convex and piecewise linear, the LQR method can be computed by linear programming.

Two-Stage Regularized and Robust Regression Methods

The robust regression method employed in the study is the QR method. We then combined the two-stage method with regularized and quantile regression methods. The regression methods formulated are Two-stage Ridge Quantile regression and Two-stage Lasso Quantile regression. These regression methods were used to estimate the linear model with autocorrelation and multicollinearity problems.

Two-Stage Ridge Quantile Regression Method

The two-stage process used to alter the data is now used for Ridge Quantile regression to produce the Two-stage Ridge Quantile regression (TRQR) model. We replaced $\mathbf{y}$ and $\mathbf{X}$ in Equation (21) by $\mathbf{y}^*$ and $\mathbf{X}^*$, respectively.

The TRQR estimate β is given by

$$\hat{\beta}_{TRQR} = \min_{\beta}\left\{\sum_{i=1}^{n}\rho_{\tau}(\mathbf{y}_i^* - \mathbf{X}_i^{*T}\beta_{\tau}) + k\sum_{j=1}^{p}(\beta_j)^2\right\},\ k>0 \qquad (23)$$

$i = 1,2,\ldots,n;\quad j = 1,2,\ldots,p$ and $k>0$ is the regularization parameter.

We therefore run the Ridge Quantile regression on $\mathbf{y}^*$ vector of response variable and $\mathbf{X}^*$ matrix of predictors, instead of running it on $\mathbf{y}$ vector of response variable and $\mathbf{X}$ matrix of predictors.

Two-Stage Lasso Quantile Regression Method

The two-stage procedure which was used to arrive at the transformed data is now applied to Lasso quantile regression to produce the Two-stage Lasso Quantile regression (TLQR) method. We now replace $\mathbf{y}$ and $\mathbf{X}$ in Equation (22) by $\mathbf{y}^*$ and $\mathbf{X}^*$, respectively.

The TLQR estimate β is given by

$$\hat{\beta}_{TLQR} = \min_{\beta}\left\{\sum_{i=1}^{n}\rho_{\tau}(\mathbf{y}_i^* - \mathbf{X}_i^{*T}\beta_{\tau}) + k\sum_{j=1}^{p}|\beta_j|\right\},\ k>0 \quad (24)$$

$i = 1,2,\ldots,n;\quad j = 1,2,\ldots,p$ and $k>0$ is the regularization parameter.

We now run the Lasso Quantile regression on $\mathbf{y}^*$ vector of response variable and $\mathbf{X}^*$ matrix of predictors, instead of running the Lasso Quantile regression on $\mathbf{y}$ vector of response variable and $\mathbf{X}$ matrix of predictors.

Performance Criterion

We shall use the Mean Squared Error, MSE, as the main criteria to compare the performances of the regression methods. Also, other performance criteria such the estimating regression coefficients of methods, R-squared, adjusted R-squared, root mean squared error were considered to evaluate the performances of the regression methods. The estimated MSE for each of the regression methods is given by

$$MSE(\hat{\beta}) = \frac{1}{500}\sum_{r=1}^{500}(\hat{\beta}_r - \beta)^T(\hat{\beta}_r - \beta)$$

where $r = 1,2,\ldots,500$, $\hat{\beta}_r$ denote parameter estimated for the r^{th} replication of the experiment and β is the true parameter value. The regression method with the smallest MSE value is considered best.

Data Description

We used both simulated and real datasets. We considered two real datasets in this study: The Portland cement dataset and the Historical dataset. We first checked for the presence of multicollinearity and autocorrelation in the datasets by using the VIF and Durbin-Watson tests. The Portland cement dataset was originally used by Woods, Steinour and Starke (1932). The dataset consists of 13 observations. The response variable is the heat evolved after 180 days of

curing and measured in calories/gram of cement with 40 percent water at 35 degrees Celsius. The four predictor variables considered are the clinker compounds (CALCD) defined as tricalcium aluminate, tricalcium silicate, tetracalcium aluminoferrite and dicalcium silicate.

The Historical dataset was discussed by Bayhan and Bayhan (1998). The dataset consists of 15 observations. The data includes the weekly quantities of shampoos sold as response variable. The two predictor variables are weekly list prices (averages from selected supermarkets) of the firm's shampoos and weekly list prices of a certain brand of soap, substituted from shampoos. Due to the difficulty in obtaining real data with multicollinearity and autocorrelation problems with sample sizes ranging between 25 to 500, we therefore used the two real datasets with sample sizes smaller than 25.

Simulation Design

We evaluated the performance of the regression methods through simulation. Following McDonald and Galarneau (1975) and Kibria (2003), we generated predictor variables as follows

$$x_{ij} = (1-\gamma^2)^{\frac{1}{2}} z_{ij} + \gamma z_{i,p+1} \qquad \text{for } i = 1,2,\ldots,n \quad \text{and} \quad j = 1,2,\ldots,p$$

where z_{ij} is an independent standard normal pseudo random number, and γ is specified so that the theoretical correlation between any two predictor variables is given by γ^2.

The response variable was generated by the equation

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_p x_{ip} + \mu_i .$$

The model parameters are assumed to be unity and β_0 is taken to be zero. The μ_i are generated from AR (1) process as

$$\mu_i = \rho\mu_{i-1} + e_i, \quad i = 1, 2, \ldots, n$$

where e_i are independent normal pseudo random numbers and ρ is autoregressive coefficient such that $|\rho| < 1$. The error terms were generated by assuming $e_i \sim N(0,1)$ and using the distributional property of the autocorrelated error terms $\mu_i \sim N\left(0, \frac{\sigma_e^2}{1-\rho^2}\right)$.

The design was performed by varying multicollinearity levels, the number of predictor variables, sample sizes and degrees of autocorrelation as shown:

1) Multicollinearity levels (0.7, 0.9 and 0.99)
2) Number of predictor variables (2, 4 and 8)
3) Sample size (25, 50, 200 and 500)
4) Degrees of autocorrelation (0.1 and 0.9)

We want to see the effect of the sample sizes, number of predictor variables, multicollinearity levels and degrees of autocorrelation on the performance of the regression methods discussed in this study. The simulation was replicated 500 times in order to obtain the approximate distribution.

Chapter Summary

Various regression methods that would be used in the study have been reviewed. These include OLS, RR, Lasso, TR, TLasso, QR, RQR, LQR, TRQR and TLQR. The concept of regularization of linear models has also been explained. In line with this, techniques of regularization processes have also

been reviewed. These are the L_1 and L_2 penalties of regularization. A discussion of Quantile regression as a robust regression method was done. The study employed three quantile levels (0.25, 0.5 and 0.75). Quantile regression with Ridge and Lasso regularization to develop quantile regression models were formulated. It is anticipated that quantile regression with regularization could improve models.

Two-stage procedure was introduced and $\mathbf{P}$ transformation matrix was adopted to be used to transform the datasets to correct for autocorrelation. To deal with multicollinearity and autocorrelation simultaneously, the chapter presents the combination of two-stage method with regularized and quantile regression methods. The data employed in the study consists of simulated and two real datasets. The simulation design was presented. Various performance criteria have been stated, which would be needed to determine the method that best addresses the problems of multicollinearity and autocorrelation simultaneously. These are MSE, root mean squared error, adjusted R-squared and R-squared. The MSE happened to be the main performance criteria for comparing the performances of the regression methods.

CHAPTER FOUR

RESULTS AND DISCUSSION

Introduction

This chapter analyses and discusses the results from our simulated and real application datasets. The study examines different regularized and robust regression methods for handling multicollinearity and autocorrelation problems. It presents results of performances of different regression methods, namely, Ordinary Least Squares Regression (OLS), Ridge Regression (RR), Lasso Regression, Two-Stage Ridge Regression (TR), Two-Stage Lasso Regression (TLasso), Quantile Regression (QR), Ridge Quantile Regression (RQR), Lasso Quantile Regression (LQR), Two-Stage Ridge Quantile Regression (TRQR), and Two-Stage Lasso Quantile Regression (TLQR). The analyses were done in two parts. Firstly, appropriate data was simulated and then the regression models were fitted using different regression methods. The MSEs of the different regression methods were computed and compared to know the best regression method for the simulated dataset. Secondly, the regression models were fitted to the real datasets described earlier. The MSEs of the different regression methods as well as other statistical values (such as regression coefficients, R-squared, adjusted R-squared, RMSE) were computed and compared to know the best regression method by applying real dataset.

Statistical Packages

The R software program was used for the analysis of the data at various stages of the study. Algorithms for data simulations were written in R language. The R software program was used to simulate data, and for analysing the simulated and real datasets. R is a software environment and programming

language for statistical computing. R has been widely used in research data analysis. The data analysis in R is done by writing scripts and functions. The following packages (or libraries) in R were employed in the analysis of the simulated dataset and the real dataset; MASS, lmridge, islasso, quantreg, hqreg, glmnet, lrmest, and the lmtest package.

The ordinary least squares regression was performed by the 'MASS' R package. For ridge regression (and two-stage ridge regression) we employed the R package 'lmridge', for Lasso regression (and two-stage Lasso regression) we employed the R package 'islasso', for quantile regression we employed the R package 'quantreg', for ridge quantile regression (and two-stage ridge quantile regression) and Lasso quantile regression (and two-stage Lasso quantile regression) we employed the R package 'hqreg'. The 'lrmest' R package was used to check the multicollinearity degree present in the dataset and the 'lmtest' R package was used to perform the Durbin-Watson test for autocorrelation of errors.

Simulation Study

Autocorrelation and multicollinearity are two of the most common regression analysis problems. We usually think of these two problems separately; however, they can occur concurrently in applied situations. A Monte Carlo simulation was designed to assess the regularized and robust regression methods performances. We designed dataset with 0.7, 0.9 and 0.99 multicollinearity levels with 0.1 degree of autocorrelation and another dataset with 0.7, 0.9 and 0.99 multicollinearity levels with 0.9 degree of autocorrelation. For both settings we use two, four and eight predictor variables with different

sample sizes; n = 25, 50, 200 and 500. The number of replications used for the simulation is 500 Monte Carlo simulations.

Autocorrelation degrees of 0.1 and 0.9 were utilized, resulting in autocorrelation for some models and no autocorrelation for others. The sensitivity of the methods was assessed by varying the multicollinearity degrees from 0.7 to 0.9 and 0.99. Sample sizes of 25, 50, 200, and 500 were utilized to ensure that the sample size ranged from small to moderate to large. The simulation runs of 500 does not take too much time to run and it is considered sufficient to produce stable estimates of the results. The same setting is used for all the methods and their mean square error (MSE) computed. All methods may thus be compared on the same level, and their MSEs can be directly compared.

Simulation Results for Two Predictor Variables

The MSE values as well as the regression coefficients of the regression methods are computed. Due to the voluminous nature of the regression coefficients generated for all the methods, only few are presented in the work and the rest in Appendix B. The Table 1 to Table 4 are some of the simulation results of regression coefficients for two predictor variables.

Table 1: **Regression Coefficients when $\gamma^2 = 0.99$, $\rho = 0.1$ and $n = 200$**

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	0.009	0.009	0.009	0.009	0.009
β_1	0.980	0.980	0.975	0.976	0.972
β_2	1.001	0.996	0.991	0.999	0.994

Table 2: **Regression Coefficients when $\gamma^2 = 0.7$, $\rho = 0.1$ and $n = 500$**

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	-0.005	-0.005	-0.006	-0.005	-0.006
β_1	0.999	0.997	0.992	0.998	0.992
β_2	1.008	1.006	1.001	1.006	1.001

Table 3: **Estimated Coefficients when $\gamma^2 = 0.70$, $\rho = 0.9$ and $n = 500$**

τ	Coefficient	QR	RQR	LQR	TRQR	TLQR
0.25	β_0	-1.489	-1.497	-1.498	-1.497	-1.497
	β_1	0.957	0.938	0.944	0.937	0.944
	β_2	1.093	1.056	1.077	1.055	1.076
0.5	β_0	0.022	0.026	0.024	0.024	0.023
	β_1	1.028	1.004	1.015	1.005	1.015
	β_2	1.011	0.986	0.996	0.986	0.997
0.75	β_0	1.332	1.345	1.351	1.351	1.356
	β_1	1.127	1.088	1.110	1.089	1.110
	β_2	0.914	0.904	0.908	0.908	0.910

Based on 500 Monte Carlo Simulations

Table 4: **Estimated Coefficients when $\gamma^2 = 0.70$, $\rho = 0.1$ and $n = 500$**

τ	Coefficient	QR	RQR	LQR	TRQR	TLQR
0.25	β_0	-0.704	-0.708	-0.709	-0.707	-0.709
	β_1	1.017	1.007	1.009	1.007	1.009
	β_2	1.000	1.001	1.003	1.001	1.004
0.5	β_0	-0.027	-0.023	-0.022	-0.022	-0.022
	β_1	0.979	0.971	0.975	0.973	0.975
	β_2	0.973	0.969	0.972	0.969	0.972
0.75	β_0	0.659	0.673	0.673	0.674	0.674
	β_1	1.098	1.070	1.072	1.067	1.070
	β_2	0.899	0.914	0.913	0.916	0.915

Based on 500 Monte Carlo Simulations

Table 5: **Estimated MSE for Two Predictor Variables when ρ = 0.1**

γ^2	*n*	OLS	RR	Lasso	TR	TLasso
0.7	25	0.92991	0.88680	0.93337	**0.88272**	0.92903
	50	0.99950	0.97748	1.00059	**0.97631**	0.99942
	200	1.00534	1.00020	1.00551	**1.00011**	1.00544
	500	1.05310	1.05097	1.05325	**1.05096**	1.05323
0.9	25	0.93267	0.88549	0.93365	**0.88186**	0.92902
	50	0.98313	0.95980	0.98322	**0.95848**	0.98172
	200	1.00534	1.00007	1.00553	**0.99999**	1.00545
	500	1.05286	1.05071	1.05303	**1.05069**	1.05301
0.99	25	0.93526	0.87858	0.92268	**0.87584**	0.91808
	50	0.98373	0.95488	0.97629	**0.95384**	0.97523
	200	1.06706	1.06030	1.06627	**1.06020**	1.06614
	500	1.05254	1.05014	1.05262	**1.05012**	1.05261

Based on 500 Monte Carlo Simulations

Table 5 shows the simulation results of the estimated MSEs for two predictor variables when ρ = 0.1. Based on the MSE of each regression method, the TR has the least MSE as compared to the other methods for all categories. Consequently, RR and TR methods are better than OLS when the multicollinearity problem exists in a data with TR being the best in this case for the two predictor variables. It clearly indicates that the estimated mean squared errors for TR is always less than that of RR method. The simulation results show that the TR method performs better than OLS, RR, Lasso and TLasso for all samples sizes (*n* = 25, 50, 200 and 500) and across the different levels of

multicollinearity (0.7, 0.9 and 0.99). In addition, we observe that the performance of Lasso and TLasso methods to the OLS method is not satisfactory with large sample size. The results show that the estimated MSE for the OLS method competes with Lasso and TLasso methods with large sample sizes ($200 \leq n \leq 500$) for two predictor variables. The findings from comparisons of the OLS, RR, Lasso, TR and TLasso methods show that the TR method is better than other methods (OLS, RR, Lasso and TLasso), and much better than OLS method.

Table 6: **Estimated MSE for Two Predictor Variables when ρ = 0.9**

γ^2	n	OLS	RR	Lasso	TR	TLasso
0.7	25	3.09038	2.94315	3.11001	**2.92748**	3.09169
	50	6.21790	6.07838	6.21422	**6.06744**	6.20470
	200	4.78922	4.76461	4.78946	**4.76418**	4.78901
	500	4.99732	4.98721	4.99748	**4.98713**	4.99740
0.9	25	3.08228	2.92430	3.07193	**2.91079**	3.05419
	50	6.21827	6.06692	6.19399	**6.05674**	6.18414
	200	6.29188	6.25867	6.29115	**6.25811**	6.29064
	500	4.99369	4.98348	4.99387	**4.98339**	4.99378
0.99	25	2.61518	2.44835	2.56965	**2.44126**	2.55847
	50	6.21861	6.03361	6.15648	**6.02524**	6.14837
	200	6.29241	6.25361	6.28074	**6.25315**	6.28065
	500	4.99047	4.97904	4.98852	**4.97896**	4.98839

Based on 500 Monte Carlo Simulations

Table 6 presents the simulation results of the estimated MSEs for two predictor variables when ρ = 0.9. Based on the MSE of each regression method, the TR has the least MSE among the other regression methods for all sample sizes (n = 25, 50, 200 and 500) and across the different levels of multicollinearity (0.7, 0.9 and 0.99). Consequently, RR and TR methods are better than OLS, Lasso and TLasso, with TR being the best in this case for the two predictor variables. It clearly indicates that the estimated mean squared errors for TR is always less than that of RR when the multicollinearity and autocorrelation problems exist in a dataset. We observe that the performance of Lasso and TLasso as compared to the OLS is not satisfactory for large sample size (n = 500). The results show that the MSE for OLS method is relatively smaller than the Lasso and TLasso methods with large sample size. The estimated MSEs of Lasso and TLasso decreases as the multicollinearity level increases for a decreased sample size. The results of comparison indicated that, the TLasso method outperforms the OLS for combinations of multicollinearity and autocorrelation problems when the degree of multicollinearity is high with small sample sizes. The estimated MSEs of TLasso method gets closer to the estimated MSEs of TR method for increased sample size. However, the TR method seems to be an appropriate alternative to other methods in handling autocorrelation and multicollinearity problems.

Table 7: **Estimated MSE for Two Predictor Variables when τ = 0.25**

ρ	γ^2	*n*	QR	RQR	LQR	TRQR	TLQR
		25	1.67303	1.66693	1.64995	1.59754	**1.57718**
	0.7	50	1.47071	1.46677	1.41505	1.42172	**1.41410**
		200	1.53069	1.51364	1.51310	**1.51181**	1.51205
		500	1.54791	1.54309	1.54397	**1.54216**	1.54393
		25	1.68017	1.62394	1.63584	1.59982	**1.57222**
0.1	0.9	50	1.51393	**1.44162**	1.44203	1.45574	1.44470
		200	1.53160	**1.51069**	1.51253	1.51174	1.51104
		500	1.54864	1.54399	1.54732	**1.54345**	1.54717
		25	1.68117	1.58152	1.88607	1.56626	**1.55670**
	0.99	50	1.51386	1.43914	1.43767	**1.43337**	1.43766
		200	1.55362	**1.51480**	1.52803	1.51528	1.52650
		500	1.54971	1.54264	1.54931	**1.54128**	1.54862
		25	5.35151	5.36602	5.59986	**5.25951**	5.38336
	0.7	50	9.24295	8.85666	8.81489	8.81102	**8.80087**
		200	7.90937	7.78684	7.78254	**7.75910**	7.76405
		500	7.22085	7.19784	7.19921	**7.19683**	7.19935
		25	5.32352	5.39459	5.38209	12.44507	**5.14952**
0.9	0.9	50	9.22574	8.83476	8.82583	8.80307	**8.76999**
		200	9.29791	9.15282	**9.14251**	9.15644	9.14642
		500	7.21351	7.19055	7.19745	7.19321	7.19729
		25	4.65485	3.88326	3.96025	**3.86494**	3.91002
	0.99	50	9.21678	8.89440	8.82273	8.83673	**8.79430**
		200	9.29719	9.16282	9.15287	9.16444	**9.15178**
		500	7.21022	**7.17773**	7.19117	7.18025	7.19138

Based on 500 Monte Carlo Simulations

In Table 7, we provide the results of estimated MSEs of the methods for two predictor variables when τ = 0.25. When ρ = 0.1, RQR, TRQR and TLQR methods outperforms the other methods. Comparing TRQR and TLQR methods, TRQR performs better than TLQR with large sample sizes (*n* = 500).

Whiles, TLQR produce better results than TRQR with small sample sizes (n = 25). When ρ = 0.9, TRQR and TLQR methods outperforms the other methods except on two cases where RQR and LQR methods produced better MSE results.

Table 8: **Estimated MSE for Two Predictor Variables when τ = 0.5**

ρ	γ^2	n	QR	RQR	LQR	TRQR	TLQR
		25	1.01189	**0.94039**	0.94678	0.95230	0.94541
	0.7	50	1.02653	0.97606	**0.97516**	0.97681	0.97517
		200	1.01361	1.00206	1.00024	1.00210	**1.00004**
		500	1.05860	1.05331	1.05241	1.05308	**1.05240**
		25	1.01519	0.94382	**0.92823**	0.94040	0.93435
0.1	0.9	50	1.02199	0.96585	0.96871	**0.96552**	0.96610
		200	1.01304	1.00233	0.99963	1.00234	**0.99961**
		500	1.05832	1.05324	**1.05151**	1.05333	1.05161
		25	1.01514	**0.91707**	0.92139	0.94047	1.03871
	0.99	50	1.02256	0.96641	0.96540	0.96632	**0.96331**
		200	1.07900	1.06225	1.06148	1.06223	**1.06137**
		500	1.05790	1.05363	**1.05068**	1.05322	1.05082
		25	3.32752	2.94866	2.96282	**2.93685**	2.94384
	0.7	50	6.51798	6.24601	6.22953	**6.19736**	6.20087
		200	4.85857	4.83668	4.82008	4.82880	**4.81209**
		500	5.00895	4.98316	4.98045	4.98325	**4.98022**
		25	3.32630	2.93558	2.93385	3.11302	**2.91792**
0.9	0.9	50	6.53255	6.21143	6.20211	**6.17466**	6.18157
		200	6.35187	6.30817	6.29031	6.30568	**6.28719**
		500	5.00513	4.98207	4.97660	4.98178	**4.97647**
		25	2.83886	2.54820	2.55223	2.57032	**2.53253**
	0.99	50	6.53412	6.20382	6.18190	**6.16772**	6.17820
		200	6.35244	6.30329	6.29353	6.29950	**6.28559**
		500	5.00170	4.97696	4.97353	4.97688	**4.97319**

Based on 500 Monte Carlo Simulations

Table 8 is the estimated MSEs of the methods for two predictor variables when $\tau = 0.5$. The two scenarios considered in the simulation studies are when $\rho = 0.1$ and $\rho = 0.9$. Based on the mean square errors for each model, we noted that all regression methods with $\tau = 0.5$ are better than regression methods when $\tau = 0.25$ and $\tau = 0.75$ (comparing Table 7, Table 8 and Table 9) for $\rho = 0.1$ and $\rho = 0.9$.

The simulation results from the first scenario when $\rho = 0.1$ indicated that, the MSE values of TLQR methods at $\tau = 0.5$ are relatively smaller than that of the TRQR methods at $\tau = 0.5$ in all the cases, except two cases where the MSE values of TRQR methods at $\tau = 0.5$ are relatively smaller than TLQR methods at $\tau = 0.5$. Also, RQR produce better results than LQR with less predictor variables and small sample size.

The simulation results from the second scenario when $\rho = 0.9$ indicated that, TRQR and TLQR methods at $\tau = 0.5$ outperforms the other methods in all the cases. Comparing TRQR and TLQR methods at $\tau = 0.5$, the TRQR produce better results than TLQR with less predictor variables and small sample sizes whiles TLQR performs better with larger sample size.

Table 9: **Estimated MSE for Two Predictor Variables when τ = 0.75**

ρ	γ^2	*n*	QR	RQR	LQR	TRQR	TLQR
		25	1.53290	**1.31314**	1.35788	1.35180	1.33929
	0.7	50	1.49271	1.39874	1.40523	**1.39821**	1.40153
		200	1.50195	1.47498	1.47467	1.47461	**1.47430**
		500	**1.50661**	1.51549	1.51571	1.51663	1.51661
		25	1.53669	1.37181	1.41571	**1.32701**	1.35836
0.1	0.9	50	1.51217	**1.46355**	1.46952	1.46446	1.46493
		200	1.50161	1.47543	**1.47520**	1.47578	1.47539
		500	**1.50884**	1.51301	1.51544	1.51391	1.51564
		25	1.53687	**1.29625**	1.32649	1.33246	1.31398
	0.99	50	1.51213	1.45056	1.44827	1.44072	**1.43896**
		200	**1.55332**	1.55523	1.55469	1.55757	1.55741
		500	1.50866	**1.50827**	1.51472	1.50918	1.51508
		25	5.25091	5.07891	5.03045	5.06845	**4.94730**
	0.7	50	9.72631	9.51762	9.55344	**9.48918**	9.53751
		200	7.30032	**7.21249**	7.24436	7.22166	7.25403
		500	**6.82334**	6.82719	6.83741	6.84147	6.85131
		25	5.24440	5.07615	5.24310	4.99809	**4.90922**
0.9	0.9	50	9.69809	9.49986	9.49538	9.49295	**9.43714**
		200	9.95328	9.68758	9.68559	**9.64254**	9.65360
		500	**6.83170**	6.83566	6.85179	6.84689	6.86236
		25	4.85473	4.74489	4.68336	4.56022	**4.46491**
	0.99	50	9.70719	9.43089	9.37521	9.40251	**9.35997**
		200	9.97164	9.69224	9.69806	**9.65311**	9.66381
		500	**6.83560**	6.85268	6.85304	6.86284	6.86532

Based on 500 Monte Carlo Simulations

In Table 9, we provide the simulation results of estimated MSEs of the methods for two predictor variables when τ = 0.75. The results show that QR method competes with the other methods with large sample sizes (*n* = 500).

Table 10: **Comparison of Methods MSEs for Two Predictor Variables**

ρ	γ^2	*n*	TR	TLasso	RQR	LQR	TRQR	TLQR
					$\tau = 0.5$	$\tau = 0.5$	$\tau = 0.5$	$\tau = 0.5$
0.1	0.7	25	**0.88272**	0.92903	0.94039	0.94678	0.95230	0.94541
		50	0.97631	0.99942	0.97606	**0.97516**	0.97681	0.97517
		200	1.00011	1.00544	1.00206	1.00024	1.00210	**1.00004**
		500	**1.05096**	1.05323	1.05331	1.05241	1.05308	1.05240
	0.9	25	**0.88186**	0.92902	0.94382	0.92823	0.94040	0.93435
		50	**0.95848**	0.98172	0.96585	0.96871	0.96552	0.96610
		200	0.99999	1.00545	1.00233	0.99963	1.00234	**0.99961**
		500	**1.05069**	1.05301	1.05324	1.05151	1.05333	1.05161
	0.99	25	**0.87584**	0.91808	0.91707	0.92139	0.94047	1.03871
		50	**0.95384**	0.97523	0.96641	0.96540	0.96632	0.96331
		200	**1.06020**	1.06614	1.06225	1.06148	1.06223	1.06137
		500	**1.05012**	1.05261	1.05363	1.05068	1.05322	1.05082
0.9	0.7	25	**2.92748**	3.09169	2.94866	2.96282	2.93685	2.94384
		50	**6.06744**	6.20470	6.24601	6.22953	6.19736	6.20087
		200	**4.76418**	4.78901	4.83668	4.82008	4.82880	4.81209
		500	4.98713	4.99740	4.98316	4.98045	4.98325	**4.98022**
	0.9	25	**2.91079**	3.05419	2.93558	2.93385	3.11302	2.91792
		50	**6.05674**	6.18414	6.21143	6.20211	6.17466	6.18157
		200	**6.25811**	6.29064	6.30817	6.29031	6.30568	6.28719
		500	4.98339	4.99378	4.98207	4.97660	4.98178	**4.97647**
	0.99	25	**2.44126**	2.55847	2.54820	2.55223	2.57032	2.53253
		50	**6.02524**	6.14837	6.20382	6.18190	6.16772	6.17820
		200	**6.25315**	6.28065	6.30329	6.29353	6.29950	6.28559
		500	4.97896	4.98839	4.97696	4.97353	4.97688	**4.97319**

Based on 500 Monte Carlo Simulations

The comparison of the performances of the methods in term of their MSE values over two predictor variables at various levels of multicollinearity and sample size is given for both ρ = 0.1 and ρ = 0.9 in Table 10. The simulation findings when ρ = 0.1 revealed that, TR method outperforms the other methods

in many of the cases, especially with severe multicollinearity. Increasing multicollinearity provides an overall improvement in TR over the other regression methods for two predictor variables. Moreover, the LQR and TLQR methods at quantile level 0.5 often competes favourably with the TR method.

The simulation results when $\rho = 0.9$ revealed that, the TR method has the smallest MSE values and therefore outperforms the other methods with high autocorrelation and across the different degrees of multicollinearity, except three of the cases where TLQR method at quantile level 0.5 performs best. For $25 \leq n \leq 200$, TR is best when multicollinearity is high with high autocorrelation for two predictor variables. Moreover, with large sample size (n = 500), the TLQR method at quantile level 0.5 is best. Overall, the use of the TR method for two predictor variables in a model is generally effective in handling multicollinearity and autocorrelation problems.

Simulation Results for Four Predictor Variables

The MSE values together with the regression coefficients are computed for all the regression methods. However, few of the regression coefficients of the methods are presented due to its voluminous nature and the rest in Appendix B. The Table 11 to Table 13 are some of the simulation results of regression coefficients for four predictor variables.

Table 11: **Coefficients for Four Predictors when $\gamma^2 = 0.90$, $\rho = 0.9$ and $n = 500$**

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	-0.080	-0.080	-0.080	-0.080	-0.080
β_1	0.968	0.968	0.960	0.968	0.961
β_2	1.000	0.999	0.995	0.997	0.994
β_3	0.762	0.765	0.757	0.766	0.757
β_4	1.100	1.097	1.094	1.097	1.094

Table 12: **Coefficients for Four Predictors when $\gamma^2 = 0.70$, $\rho = 0.1$ and $n = 200$**

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	-0.016	-0.016	-0.016	-0.016	-0.016
β_1	1.107	1.104	1.097	1.105	1.097
β_2	1.054	1.051	1.050	1.051	1.051
β_3	0.881	0.882	0.874	0.882	0.874
β_4	0.995	0.993	0.989	0.993	0.989

Based on 500 Monte Carlo Simulations

Table 13: **Coefficients for Four Predictors when $\gamma^2 = 0.70$, $\rho = 0.1$ and $n = 25$**

τ	Coefficient	QR	RQR	LQR	TRQR	TLQR
0.25	β_0	-0.393	-0.426	-0.421	-0.403	-0.403
	β_1	0.935	0.890	0.905	0.896	0.909
	β_2	0.980	0.895	0.938	0.910	0.941
	β_3	1.044	0.946	1.013	0.960	1.036
	β_4	0.881	0.899	0.876	0.897	0.868
0.5	β_0	0.078	0.066	0.082	0.078	0.094
	β_1	1.045	1.000	1.038	0.992	1.037
	β_2	0.990	0.940	0.970	0.945	0.977
	β_3	1.083	1.007	1.058	1.024	1.074
	β_4	0.780	0.809	0.731	0.791	0.706
0.75	β_0	0.690	0.665	0.700	0.678	0.707
	β_1	0.974	0.927	0.949	0.927	0.946
	β_2	1.120	1.017	1.068	1.018	1.070
	β_3	1.375	1.172	1.280	1.188	1.297
	β_4	0.587	0.692	0.585	0.674	0.581

Based on 500 Monte Carlo Simulations

Table 14: **Estimated MSE for Four Predictor Variables when ρ = 0.1**

γ^2	*n*	OLS	RR	Lasso	TR	TLasso
	25	0.79317	0.74632	0.79619	**0.74004**	0.78820
0.7	50	0.97184	0.94744	0.97356	**0.94468**	0.97061
	200	1.10471	1.09884	1.10532	**1.09865**	1.10512
	500	0.93315	0.93124	0.93370	**0.93121**	0.93367
	25	0.79143	0.73637	0.78102	**0.73153**	0.77422
0.9	50	0.97208	0.94300	0.97025	**0.94060**	0.96744
	200	1.10437	1.09805	1.10505	**1.09786**	1.10485
	500	0.93271	0.93073	0.93338	**0.93071**	0.93335
	25	1.51952	1.36426	1.41603	**1.35975**	1.40886
0.99	50	0.97223	0.93640	0.94745	**0.93509**	0.94617
	200	1.10411	1.09532	1.09966	**1.09520**	1.09948
	500	0.97964	0.97680	0.97946	**0.97678**	0.97943

Based on 500 Monte Carlo Simulations

Table 14 presents the simulation results of estimated MSEs for four predictor variables when ρ = 0.1. The results clearly indicate that the estimated mean squared errors for TR is always less than that of RR method. The simulation results therefore showed that the TR method outperforms OLS, RR, Lasso and TLasso methods for all the different samples sizes (*n* = 25, 50, 200 and 500) and across the different degrees of multicollinearity (0.7, 0.9 and 0.99). Again, we can see that the performance of Lasso and TLasso methods to the OLS method is not satisfactory for large sample size. The results show that the MSE for the OLS method competes with Lasso and TLasso methods with large

sample sizes ($200 \leq n \leq 500$) for four predictor variables. However, the Lasso and TLasso methods perform better than the OLS method for a very high multicollinearity ($\gamma^2 = 0.99$) across all samples size (n = 25, 50, 200 and 500). The findings from comparisons of the OLS, RR, Lasso, TR and TLasso methods show that, the TR method outperforms the other methods (OLS, RR, Lasso and TLasso), and much better than OLS method.

Table 15: **Estimated MSE for Four Predictor Variables when ρ = 0.9**

γ^2	n	OLS	RR	Lasso	TR	TLasso
0.7	25	2.68242	2.51313	2.66882	**2.49036**	2.63958
	50	3.12628	3.04559	3.12484	**3.03859**	3.11820
	200	4.49030	4.46635	4.49088	**4.46557**	4.49009
	500	5.91710	5.90495	5.91762	**5.90477**	5.91744
0.9	25	2.69813	2.48892	2.61093	**2.47101**	2.58661
	50	3.12546	3.02730	3.09662	**3.02129**	3.09004
	200	4.49126	4.46541	4.48924	**4.46465**	4.48850
	500	5.92112	5.90852	5.92130	**5.90834**	5.92111
0.99	25	1.59896	1.43701	1.50212	**1.42933**	1.49333
	50	2.72744	2.61173	2.65555	**2.60878**	2.65264
	200	4.49217	4.45411	4.46496	**4.45359**	4.46428
	500	5.92468	5.90764	5.91089	**5.90748**	5.91082

Based on 500 Monte Carlo Simulations

Table 15 presents the simulation results of the estimated MSEs for four predictor variables when ρ = 0.9. According to the MSE criterion of each regression method, the TR method has the least MSE among the other

regression methods (OLS, RR, Lasso and TLasso) for all sample sizes (n = 25, 50, 200 and 500) and across the levels of multicollinearity (0.7, 0.9 and 0.99). Moreover, the RR method on the other hand outperforms OLS, Lasso and TLasso methods in the above-mentioned categories. However, the estimated mean squared errors for TR method are always less than that of RR when the multicollinearity and autocorrelation problems exist in a dataset. The estimated MSEs of Lasso and TLasso methods decreases as the multicollinearity level increases for a decreased sample size.

The results of comparison indicated that, the Lasso and TLasso methods are better than the OLS method in existence of autocorrelation and multicollinearity problems. But with a very large sample size, OLS method is a close competitor to the Lasso and TLasso methods with respect to their MSEs. The superiority of the Lasso and TLasso methods over OLS is primarily determined by the degree of autocorrelation and the number of predictors in the data. The estimated MSEs of TLasso method gets closer to the estimated MSEs of TR method for large sample size. Therefore, the TR method seems to be an appropriate alternative to other methods in existence of autocorrelation and multicollinearity issues.

Table 16: **Estimated MSE for Four Predictor Variables when τ = 0.25**

ρ	γ^2	*n*	QR	RQR	LQR	TRQR	TLQR
		25	1.31396	1.17378	1.11539	1.15835	**1.09540**
	0.7	50	1.56791	1.42125	1.43337	1.41349	**1.41301**
		200	**1.61177**	1.62406	1.62824	1.62250	1.62459
		500	**1.33774**	1.33962	1.34153	1.33925	1.34297
		25	1.31779	1.17094	1.11706	1.15495	**1.10509**
0.1	0.9	50	1.56656	1.42478	1.43347	**1.41194**	1.41469
		200	**1.61104**	1.61154	1.63047	1.61500	1.62749
		500	1.33763	**1.33503**	1.34465	1.33507	1.34579
		25	2.39220	1.79282	1.79185	**1.79054**	1.80130
	0.99	50	1.56778	1.42501	1.42308	1.41533	**1.40905**
		200	1.61073	**1.59007**	1.62195	1.59750	1.62136
		500	1.39403	1.38851	1.40645	**1.38786**	1.40530
		25	4.47262	**3.44662**	3.52391	3.45838	3.51838
	0.7	50	4.87853	4.50648	4.55911	**4.48776**	4.53077
		200	7.27739	7.11387	7.09863	7.10118	**7.09385**
		500	9.01138	8.88897	8.92142	**8.88746**	8.91700
		25	4.49228	3.49247	3.50201	**3.49218**	3.50178
0.9	0.9	50	4.87517	4.44418	4.49152	**4.43232**	4.46777
		200	7.28705	7.11294	7.11335	**7.09828**	7.10563
		500	9.00495	8.86667	8.92679	**8.86500**	8.92235
		25	2.80273	2.28988	2.26570	2.23702	**2.22067**
	0.99	50	4.49098	4.14752	4.18015	**4.13351**	4.17253
		200	7.29835	7.07723	7.08919	**7.07073**	7.08628
		500	9.00659	8.87032	8.90629	**8.86687**	8.90063

Based on 500 Monte Carlo Simulations

In Table 16, we provide the results of estimated MSEs of the methods for four predictor variables when τ = 0.25. When ρ = 0.1, QR, RQR, TRQR and TLQR methods performed in various cases except LQR method which did not perform in any case. The QR method improved in performance with large

sample sizes. When ρ = 0.9, TRQR method performs better than the other methods except on three cases where RQR and TLQR methods produced better MSE results.

Table 17: **Estimated MSE for Four Predictor Variables when τ = 0.5**

ρ	γ^2	*n*	QR	RQR	LQR	TRQR	TLQR
		25	0.90448	0.75924	**0.73979**	0.75710	0.74255
	0.7	50	1.03008	0.95122	0.94233	0.94849	**0.93236**
		200	1.12410	**1.08938**	1.09068	1.08997	1.09077
		500	0.94090	0.92905	**0.92903**	0.92905	0.92904
		25	0.90216	0.76244	**0.73886**	0.75566	0.74049
0.1	0.9	50	1.03153	0.94813	0.93876	0.94527	**0.93450**
		200	1.12321	**1.08829**	1.08962	1.08830	1.08937
		500	0.94044	0.92854	**0.92838**	0.92849	0.92843
		25	1.66395	1.35950	1.34908	1.34965	**1.33492**
	0.99	50	1.03161	0.94637	0.93587	0.94208	**0.93063**
		200	1.12210	1.09122	1.08783	1.09116	**1.08744**
		500	0.98521	0.97408	0.97274	0.97401	**0.97268**
		25	3.06747	2.54789	2.54344	2.52256	**2.51145**
	0.7	50	3.29792	3.03599	3.01265	3.03142	**3.00107**
		200	4.55444	4.50012	4.47009	4.49726	**4.46180**
		500	5.97541	5.92110	5.91003	5.92121	**5.90906**
		25	3.09165	2.52656	2.52883	**2.50138**	2.50761
0.9	0.9	50	3.29681	3.02601	3.02121	3.01934	**3.00150**
		200	4.55552	4.48477	4.46615	4.48548	**4.46158**
		500	5.98034	5.92179	5.91414	5.92071	**5.91224**
		25	1.87440	1.52013	1.51713	1.49731	**1.48415**
	0.99	50	2.89477	2.61410	2.59416	2.61717	**2.59072**
		200	4.55689	4.46004	4.45659	4.46356	**4.44875**
		500	5.98351	5.91571	5.91309	5.91589	**5.91232**

Based on 500 Monte Carlo Simulations

Table 17 provides the simulation results of estimated MSEs of the methods for four predictor variables when $\tau = 0.5$. The two scenarios considered in the simulation studies are the cases when the degrees of the autocorrelation are at 0.1 and 0.9. Based on the MSE criterion for each regression method, we noted that all regression methods with $\tau = 0.5$ outperforms regression methods with $\tau = 0.25$ and $\tau = 0.75$ (see Table 16, Table 17 and Table 18).

The simulation results for the first scenario when $\rho = 0.1$, show that the MSEs of the TLQR methods at $\tau = 0.5$ are relatively smaller than the TRQR methods at $\tau = 0.5$, except two cases where MSE values of TRQR methods at $\tau = 0.5$ are relatively smaller than TLQR methods at $\tau = 0.5$. In fact, RQR is a close competitor with LQR in a model with less predictor variables. But the results from comparisons of the methods show that, TLQR methods at $\tau = 0.5$ are most preferred methods with very high multicollinearity level $(\gamma^2 = 0.99)$ at the different samples size (n = 25, 50, 200 and 500).

The simulation findings for the second scenario when $\rho = 0.9$ revealed that, the TRQR and TLQR methods at $\tau = 0.5$ outperforms the other methods in all the cases. It can be noticed that in all the considered situations for $\rho = 0.9$, TLQR at $\tau = 0.5$ has the least MSE and outperforms RQR, LQR and TRQR at all quantile levels. There is only one case, where TRQR at $\tau = 0.5$ outperforms TLQR, however, the difference in MSE values between the two methods appears to be small. Therefore, the TLQR method is better than the other methods (QR, RQR, LQR and TRQR) in handling multicollinearity and autocorrelation problems in a dataset for four predictor variables.

Table 18: **Estimated MSE for Four Predictor Variables when τ = 0.75**

ρ	γ^2	*n*	QR	RQR	LQR	TRQR	TLQR
		25	1.31787	**1.01934**	1.05242	1.02290	1.04090
	0.7	50	1.47537	1.34610	1.35145	**1.34541**	1.34749
		200	1.59124	**1.57675**	1.58451	1.57867	1.58432
		500	**1.32228**	1.33490	1.33903	1.33449	1.33844
		25	1.31068	0.98695	1.03265	**0.98622**	1.02691
0.1	0.9	50	1.47661	1.34286	1.35368	**1.33966**	1.34684
		200	1.59050	**1.56396**	1.58104	1.56745	1.58289
		500	**1.32130**	1.32584	1.33918	1.32709	1.33942
		25	2.56391	2.07129	2.09027	**2.04046**	2.06171
	0.99	50	1.47829	1.34381	1.33885	**1.33611**	1.33734
		200	1.59160	**1.55894**	1.57269	1.56176	1.57355
		500	1.42636	1.42203	1.43544	**1.42152**	1.43448
		25	4.24198	3.66923	**3.64140**	3.65348	3.64204
	0.7	50	4.77836	**4.36507**	4.38528	4.36631	4.36965
		200	7.16280	6.96338	6.99399	**6.95118**	6.99071
		500	8.91545	**8.81505**	8.84163	8.81812	8.83924
		25	4.28831	**3.52488**	3.56522	3.54655	3.54896
0.9	0.9	50	4.78741	**4.35174**	4.38378	4.37615	4.38269
		200	7.16071	**6.93692**	6.99307	6.93755	6.99062
		500	8.92797	**8.81098**	8.86837	8.81422	8.86685
		25	2.80838	**2.08379**	2.08880	2.08897	2.09609
	0.99	50	4.56873	4.23370	4.15337	4.17686	**4.12389**
		200	7.16285	**6.94089**	6.97935	6.94591	6.97719
		500	8.93844	**8.81946**	8.86458	8.82152	8.86687

Based on 500 Monte Carlo Simulations

Table 18 provides the simulation results of estimated MSEs of the methods for four predictor variables when τ = 0.75. When ρ = 0.1, RQR and TRQR are the best performing methods as well as QR method competing at large sample sizes (*n* = 500). When ρ = 0.9, RQR method performs better than the other

methods except on three cases where LQR, TRQR and TLQR methods produced better MSE results.

Table 19: **Comparison of Methods MSEs for Four Predictor Variables**

ρ	γ^2	n	TR	TLasso	RQR	LQR	TRQR	TLQR
					$\tau = 0.5$	$\tau = 0.5$	$\tau = 0.5$	$\tau = 0.5$
		25	0.74004	0.78820	0.75924	**0.73979**	0.75710	0.74255
	0.7	50	0.94468	0.97061	0.95122	0.94233	0.94849	**0.93236**
		200	1.09865	1.10512	**1.08938**	1.09068	1.08997	1.09077
		500	0.93121	0.93367	0.92905	**0.92903**	0.92905	0.92904
		25	**0.73153**	0.77422	0.76244	0.73886	0.75566	0.74049
0.1	0.9	50	0.94060	0.96744	0.94813	0.93876	0.94527	**0.93450**
		200	1.09786	1.10485	**1.08829**	1.08962	1.08830	1.08937
		500	0.93071	0.93335	0.92854	**0.92838**	0.92849	0.92843
		25	1.35975	1.40886	1.35950	1.34908	1.34965	**1.33492**
	0.99	50	0.93509	0.94617	0.94637	0.93587	0.94208	**0.93063**
		200	1.09520	1.09948	1.09122	1.08783	1.09116	**1.08744**
		500	0.97678	0.97943	0.97408	0.97274	0.97401	**0.97268**
		25	**2.49036**	2.63958	2.54789	2.54344	2.52256	2.51145
	0.7	50	3.03859	3.11820	3.03599	3.01265	3.03142	**3.00107**
		200	4.46557	4.49009	4.50012	4.47009	4.49726	**4.46180**
		500	**5.90477**	5.91744	5.92110	5.91003	5.92121	5.90906
		25	**2.47101**	2.58661	2.52656	2.52883	2.50138	2.50761
0.9	0.9	50	3.02129	3.09004	3.02601	3.02121	3.01934	**3.00150**
		200	4.46465	4.48850	4.48477	4.46615	4.48548	**4.46158**
		500	**5.90834**	5.92111	5.92179	5.91414	5.92071	5.91224
		25	**1.42933**	1.49333	1.52013	1.51713	1.49731	1.48415
	0.99	50	2.60878	2.65264	2.61410	2.59416	2.61717	**2.59072**
		200	4.45359	4.46428	4.46004	4.45659	4.46356	**4.44875**
		500	**5.90748**	5.91082	5.91571	5.91309	5.91589	5.91232

Based on 500 Monte Carlo Simulations

The summary of the performances of the methods in term of their MSE values over four predictor variables at various levels of multicollinearity and

sample size is given for both ρ = 0.1 and ρ = 0.9 in Table 19. When ρ = 0.1, RQR, LQR and TLQR methods at quantile level 0.5 outperforms the other methods except on one case where TR method is best when *n* = 25 at 0.9 multicollinearity level. The TLQR method produce better results than the other methods with severe multicollinearity $(\gamma^2 = 0.99)$ for the different samples size. When $\gamma^2 = 0.7$, the pattern of the findings is similar to when $\gamma^2 = 0.9$ and 0.99 for high autocorrelation $(\rho = 0.9)$. When $n = 25$ and $n = 500$, the TR method is best except when $50 \le n \le 200$. Moreover, when $50 \le n \le 200$, the TLQR method at quantile level 0.5 is generally best. Therefore, both TR method and TLQR method at quantile level 0.5 are the best methods in handling multicollinearity and autocorrelation problems in a dataset with four predictors.

Simulation Results for Eight Predictor Variables

The MSE values and regression coefficients were computed for all the regression methods. All estimated MSE values for all the regression methods are presented. The coefficients of the methods are in Appendix B. However, few of the simulation results of regression coefficients of the regression methods for eight predictor variables are presented as in Table 20 and Table 21.

Table 20: **Coefficients for Eight Predictors when γ² = 0.70, ρ = 0.1 and *n* = 500**

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	0.000	0.000	0.001	0.000	0.001
β_1	0.915	0.915	0.906	0.915	0.906
β_2	1.016	1.016	1.009	1.016	1.008
β_3	1.022	1.021	1.016	1.021	1.015
β_4	1.010	1.010	1.003	1.010	1.004
β_5	0.935	0.935	0.927	0.935	0.927
β_6	0.982	0.982	0.976	0.982	0.976
β_7	1.085	1.084	1.081	1.084	1.082
β_8	1.013	1.013	1.008	1.013	1.008

Table 21: **Coefficients for Eight Predictors when $\gamma^2 = 0.90$, $\rho = 0.1$ and $n = 500$**

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	0.000	0.000	0.001	0.000	0.001
β_1	0.856	0.858	0.844	0.858	0.844
β_2	1.031	1.030	1.022	1.030	1.022
β_3	1.040	1.039	1.032	1.039	1.032
β_4	1.021	1.020	1.014	1.020	1.014
β_5	0.890	0.892	0.881	0.892	0.881
β_6	0.971	0.972	0.966	0.972	0.966
β_7	1.150	1.146	1.149	1.146	1.150
β_8	1.026	1.025	1.021	1.024	1.021

Table 22: **Estimated MSE for Eight Predictor Variables when $\rho = 0.1$**

γ^2	n	OLS	RR	Lasso	TR	TLasso
0.7	25	1.07969	0.96810	1.07937	**0.95063**	1.05433
	50	0.94211	0.90946	0.94665	**0.90441**	0.94123
	200	1.04069	1.03468	1.04300	**1.03424**	1.04256
	500	0.96284	0.96080	0.96501	**0.96073**	0.96494
0.9	25	1.12060	0.96152	1.04285	**0.95220**	1.02310
	50	0.94199	0.89614	0.93221	**0.89190**	0.92691
	200	1.02489	1.01785	1.02740	**1.01744**	1.02697
	500	0.96294	0.96072	0.96556	**0.96065**	0.96549
0.99	25	0.94474	0.79017	0.78963	0.78676	**0.77136**
	50	0.95736	0.90131	0.89508	0.89986	**0.89293**
	200	1.02433	1.01135	1.01286	**1.01109**	1.01254
	500	0.96301	0.95912	0.96280	**0.95906**	0.96273

Based on 500 Monte Carlo Simulations

Table 22 presents the simulation results of estimated MSEs for eight predictor variables when ρ = 0.1. The results indicate that the OLS is the least effective method among the others. According to the results, we observed that, increasing the degree of multicollinearity between the predictor variables has an adverse effect on the OLS method's MSE values. In all the considered cases, the TR method is superior to the other methods. Except in two cases where TLasso excels and TR is a close competitor. We observed that with large number of predictors, TLasso has a significantly lower MSE value, corresponding to a smaller sample size.

Again, we observed that the performance of OLS method performs better than Lasso and TLasso methods as the sample size increases with multicollinearity level not being too high. The results show that the estimated MSE for the OLS method competes with Lasso and TLasso methods with large samples sizes ($200 \leq n \leq 500$) for eight predictor variables. However, Lasso and TLasso methods perform better than the OLS method for a very high multicollinearity ($\gamma^2 = 0.99$) across all sample size (n = 25, 50, 200 and 500). It can be seen that Lasso and TLasso methods performs better with increase predictor variables. The findings from comparisons of the OLS, RR, Lasso, TR and TLasso methods reveals that the TR method is superior to the other methods, and much better than OLS method.

Table 23: **Estimated MSE for Eight Predictor Variables when ρ = 0.9**

γ^2	*n*	OLS	RR	Lasso	TR	TLasso
	25	2.93603	2.62886	2.81315	**2.58344**	2.74688
0.7	50	5.40690	5.22122	5.35103	**5.19215**	5.31645
	200	5.02611	4.99686	5.02748	**4.99484**	5.02540
	500	5.83116	5.81876	5.83329	**5.81836**	5.83289
	25	3.07982	2.64225	2.77629	**2.60661**	2.71323
0.9	50	5.43221	5.17614	5.24180	**5.15134**	5.21259
	200	4.76206	4.72934	4.74726	**4.72749**	4.74553
	500	5.83009	5.81664	5.82963	**5.81625**	5.82923
	25	5.42238	4.74158	4.88592	**4.71487**	4.87319
0.99	50	3.73895	3.49185	3.49855	**3.48668**	3.49337
	200	4.44894	4.38877	4.37352	4.38777	**4.37266**
	500	4.70596	4.68657	4.67773	4.68630	**4.67749**

Based on 500 Monte Carlo Simulations

Table 23 presents the simulation results of the estimated MSEs for eight predictor variables when ρ = 0.9. The results reveal how the TR method is superior to the other regression methods, except in two cases, where TR is a close contender to TLasso in terms of performance. It is clear from the results that with large number of predictors, TLasso has a significantly lower MSE value. We observed from the results that, OLS is the worst performing method as compared to the regularized and robust methods. Moreover, the RR method on the other hand outperforms OLS, Lasso and TLasso methods in the above-mentioned categories, except in two cases where TLasso is superior. Also, the

estimated MSEs for TR method is always less than that of RR method when the multicollinearity and autocorrelation problems exist in a dataset.

The results of comparison indicated that, Lasso and TLasso methods outperforms the OLS method. However, with a very large sample size and high multicollinearity level $(\gamma^2 = 0.7)$, OLS method turns to compete with the Lasso and TLasso methods. The better performance of the Lasso and TLasso methods over OLS largely depends on the number of predictors, autocorrelation degree and multicollinearity level in the data. The MSEs of TLasso method is relatively smaller than the estimated MSEs of TR method for large sample sizes ($200 \leq n \leq 500$) with very high multicollinearity level $(\gamma^2 = 0.99)$. Therefore, the TR method seems to be an appropriate alternative to other methods in handling multicollinearity and autocorrelation issues in a dataset.

Table 24: **Estimated MSE for Eight Predictor Variables when τ = 0.25**

ρ	γ^2	*n*	QR	RQR	LQR	TRQR	TLQR
		25	2.20302	1.45091	1.48027	**1.41853**	1.46643
	0.7	50	1.57389	1.35297	**1.33512**	1.34850	1.33637
		200	1.55749	**1.53875**	1.58086	1.54456	1.57797
		500	1.36306	**1.34529**	1.35340	1.34759	1.35340
		25	2.18230	1.55069	1.50571	1.50090	**1.48489**
0.1	0.9	50	1.57837	1.32190	1.35851	**1.30959**	1.35245
		200	1.55049	**1.47417**	1.54275	1.47656	1.54529
		500	1.36380	**1.33813**	1.35105	1.33827	1.35000
		25	1.26039	**0.82645**	0.95528	1.01694	0.96769
	0.99	50	1.58784	1.31083	1.34495	**1.30515**	1.34694
		200	1.55025	**1.47375**	1.53949	1.47681	1.53437
		500	1.36354	1.33772	1.35027	**1.33767**	1.35045
		25	5.56677	**3.32772**	3.55307	3.36275	3.57860
	0.7	50	9.26491	7.97663	7.87183	7.94825	**7.78959**
		200	7.48243	7.22258	7.25300	**7.20474**	7.24774
		500	8.48460	**8.47305**	8.54428	8.48557	8.53745
		25	6.47817	4.10651	4.36974	**4.06454**	4.25326
0.9	0.9	50	9.33168	7.84606	7.84190	**7.76542**	7.77199
		200	7.12480	6.98358	**6.86553**	6.97817	6.86761
		500	8.49393	**8.43674**	8.55521	8.45071	8.54787
		25	11.09960	7.61655	**7.35555**	7.51975	7.40308
	0.99	50	6.68056	5.59788	5.53610	5.51677	**5.47184**
		200	6.62721	6.43732	6.43750	**6.42543**	6.44767
		500	6.99696	6.84654	6.87209	**6.84653**	6.87391

Based on 500 Monte Carlo Simulations

Table 24 is the results of estimated MSEs of the methods for eight predictor variables when τ = 0.25. When ρ = 0.1, RQR and TRQR methods outperforms the other methods except on two cases where LQR and TLQR methods showed relatively smaller MSEs. When ρ = 0.9, RQR, LQR, TRQR and TLQR methods

performed in various cases except QR method which did not perform in any case.

Table 25: **Estimated MSE for Eight Predictor Variables when $\tau = 0.5$**

ρ	γ^2	n	QR	RQR	LQR	TRQR	TLQR
		25	1.47640	1.05182	0.98518	1.01957	**0.97003**
	0.7	50	1.05451	0.87327	0.84135	0.86760	**0.83432**
		200	1.06446	1.01480	**1.00719**	1.01393	1.00742
		500	0.97033	0.95085	**0.94931**	0.95077	0.94953
		25	1.41001	0.97448	0.93355	0.95254	**0.91193**
0.1	0.9	50	1.05610	0.87052	0.83976	0.86653	**0.83390**
		200	1.05254	1.01827	**0.99357**	1.01732	0.99406
		500	0.97042	0.95271	**0.94935**	0.95275	0.94963
		25	1.15501	0.84273	**0.71771**	0.94022	0.71810
	0.99	50	1.06170	0.89531	0.88140	0.89759	**0.87884**
		200	1.05191	1.01698	**0.99416**	1.01626	0.99426
		500	0.97044	0.95372	**0.95003**	0.95388	0.95027
		25	3.66145	2.56612	2.46750	2.53159	**2.39661**
	0.7	50	6.06406	5.07925	5.05350	5.04006	**5.00930**
		200	5.19743	4.92285	4.91899	4.92518	**4.91559**
		500	5.89706	5.78885	5.77542	5.78781	**5.77408**
		25	3.88572	2.57627	2.50354	2.54634	**2.48789**
0.9	0.9	50	6.09679	5.06768	5.04446	5.05318	**5.01724**
		200	4.87837	4.66307	4.63826	4.66074	**4.63784**
		500	5.89621	5.78518	5.77145	5.78472	**5.77001**
		25	7.03337	4.93071	4.96186	4.89955	**4.78975**
	0.99	50	4.24534	3.62390	3.54285	3.60834	**3.51656**
		200	4.56346	4.35520	4.33865	4.35414	**4.33363**
		500	4.74864	4.66406	4.64564	4.66396	**4.64552**

Based on 500 Monte Carlo Simulations

Table 25 shows the simulation results of estimated MSEs of the methods for eight predictor variables. The two scenarios considered in the simulation studies are when $\rho = 0.1$ and $\rho = 0.9$. Based on the MSE criterion for each regression method, it can be seen that all regression methods with $\tau = 0.5$ outperforms regression methods with $\tau = 0.25$ and $\tau = 0.75$ (Comparing Table 24, Table 25 and Table 26). The simulation results when $\rho = 0.1$ in Table 20 show that, the TLQR method at quantile level 0.5 has the least MSE values and therefore outperforms the other regression methods with small sample sizes ($25 \leq n \leq 50$). However, LQR method at quantile level 0.5 performs better than the other regression methods with large samples size ($200 \leq n \leq 500$) for eight predictor variables.

The simulation results when $\rho = 0.9$ is also presented in Table 25. According to the MSE criterion of each regression method, the TLQR method at quantile level 0.5 has the least MSE values in comparison to the other regression methods (QR, RQR, LQR and TRQR) for all sample size (n = 25, 50, 200 and 500) and across the various degrees of multicollinearity (0.7, 0.9 and 0.99). Therefore, the TLQR method appears to be an appropriate alternative to other regression methods in handling multicollinearity and autocorrelation problems in a dataset.

Table 26: **Estimated MSE for Eight Predictor Variables when τ = 0.75**

ρ	γ^2	*n*	QR	RQR	LQR	TRQR	TLQR
		25	2.21612	1.39741	1.52434	**1.36849**	1.47923
	0.7	50	1.57866	1.35908	1.40185	**1.35867**	1.39690
		200	1.50637	**1.48276**	1.52982	1.48306	1.52826
		500	**1.41912**	1.42367	1.43833	1.41959	1.43222
		25	2.13994	**1.23096**	1.40921	1.24087	1.38443
0.1	0.9	50	1.57884	1.34615	1.42212	**1.34140**	1.41583
		200	1.48939	**1.42825**	1.48452	1.42969	1.48374
		500	1.41939	1.40898	1.41902	**1.40840**	1.41784
		25	1.43778	1.10097	1.01623	1.09936	**1.00166**
	0.99	50	1.56937	**1.25941**	1.30909	1.27159	1.30644
		200	1.48961	**1.41329**	1.46770	1.41844	1.46799
		500	1.41982	1.40094	1.42576	**1.40036**	1.41749
		25	5.79823	3.92680	3.98805	3.91850	**3.83065**
	0.7	50	9.19423	7.65614	7.76360	**7.64675**	7.75062
		200	7.64372	7.41537	**7.40098**	7.41816	7.40454
		500	8.71165	**8.53943**	8.66106	8.54503	8.65626
		25	5.79283	4.04398	3.81103	4.03771	**3.78167**
0.9	0.9	50	9.24314	7.65766	7.72531	**7.59990**	7.65400
		200	7.57919	7.26411	7.24616	7.25269	**7.23658**
		500	8.71108	**8.50661**	8.67923	8.51408	8.66611
		25	11.27726	7.63530	7.42371	7.52533	**7.34976**
	0.99	50	6.64791	5.41886	5.38921	5.41305	**5.38448**
		200	6.22530	5.94303	5.98940	**5.93776**	5.98178
		500	6.87508	**6.72889**	6.85626	6.73336	6.86475

Based on 500 Monte Carlo Simulations

Table 26 shows the simulation results of estimated MSEs of the methods for eight predictor variables when τ = 0.75. When ρ = 0.1, QR, RQR, TRQR and TLQR methods performed in various cases except LQR method which did not perform in any case. When ρ = 0.9, RQR performs better with large sample sizes

(n = 500). Whiles, TLQR produce better results with small sample sizes (n = 25).

Table 27: **Comparison of Methods MSEs for Eight Predictor Variables**

ρ	γ^2	n	TR	TLasso	RQR	LQR	TRQR	TLQR
					τ = 0.5	τ = 0.5	τ = 0.5	τ = 0.5
		25	**0.95063**	1.05433	1.05182	0.98518	1.01957	0.97003
	0.7	50	0.90441	0.94123	0.87327	0.84135	0.86760	**0.83432**
		200	1.03424	1.04256	1.01480	**1.00719**	1.01393	1.00742
		500	0.96073	0.96494	0.95085	**0.94931**	0.95077	0.94953
		25	0.95220	1.02310	0.97448	0.93355	0.95254	**0.91193**
0.1	0.9	50	0.89190	0.92691	0.87052	0.83976	0.86653	**0.83390**
		200	1.01744	1.02697	1.01827	**0.99357**	1.01732	0.99406
		500	0.96065	0.96549	0.95271	**0.94935**	0.95275	0.94963
		25	0.78676	0.77136	0.84273	**0.71771**	0.94022	0.71810
	0.99	50	0.89986	0.89293	0.89531	0.88140	0.89759	**0.87884**
		200	1.01109	1.01254	1.01698	**0.99416**	1.01626	0.99426
		500	0.95906	0.96273	0.95372	**0.95003**	0.95388	0.95027
		25	2.58344	2.74688	2.56612	2.46750	2.53159	**2.39661**
	0.7	50	5.19215	5.31645	5.07925	5.05350	5.04006	**5.00930**
		200	4.99484	5.02540	4.92285	4.91899	4.92518	**4.91559**
		500	5.81836	5.83289	5.78885	5.77542	5.78781	**5.77408**
		25	2.60661	2.71323	2.57627	2.50354	2.54634	**2.48789**
0.9	0.9	50	5.15134	5.21259	5.06768	5.04446	5.05318	**5.01724**
		200	4.72749	4.74553	4.66307	4.63826	4.66074	**4.63784**
		500	5.81625	5.82923	5.78518	5.77145	5.78472	**5.77001**
		25	**4.71487**	4.87319	4.93071	4.96186	4.89955	4.78975
	0.99	50	**3.48668**	3.49337	3.62390	3.54285	3.60834	3.51656
		200	4.38777	4.37266	4.35520	4.33865	4.35414	**4.33363**
		500	4.68630	4.67749	4.66406	4.64564	4.66396	**4.64552**

Based on 500 Monte Carlo Simulations

The comparison of the performances of the methods in term of their MSE values over eight predictor variables at various levels of multicollinearity and

sample sizes is given for both ρ = 0.1 and ρ = 0.9 in Table 27. The simulation findings when ρ = 0.1 indicated that, LQR and TLQR methods at quantile level 0.5 outperforms the other methods in all the cases, except one case where TR method is the best when $n = 25$ at 0.7 degree of multicollinearity. Comparing LQR and TLQR methods at quantile level 0.5, LQR performs better than TLQR with large sample sizes ($200 \leq n \leq 500$). Whiles, TLQR produce better results than LQR with small sample sizes ($25 \leq n \leq 50$).

The simulation results when ρ = 0.9 indicated that, TLQR method at quantile level 0.5 outperforms the other methods when the autocorrelation is high and across the different degrees of multicollinearity, except in two cases where TR is superior. However, for $25 \leq n \leq 50$ with high autocorrelation and severe multicollinearity, TR is best. With large sample sizes, the TLQR at quantile level 0.5 is best. The study showed that the use of the TLQR method in the cases when the predictor variables are affected by multicollinearity and autocorrelation is one of the successful ways to solve this issue.

Real Data Application

We investigate the performances of the regression methods on two real applications: The Portland cement dataset and the Historical dataset.

Analysis of Portland Cement Data

This dataset was originally used by Woods, Steinour and Starke (1932). The data includes the heat evolved in calories per gram of cement (y) as dependent variable and four ingredients as predictor variables: tricalcium aluminate (X_1), tricalcium silicate (X_2), tetracalcium aluminoferrite (X_3) and dicalcium silicate (X_4).

Table 28: **Descriptive Statistics of Variables**

Variable	N	Min	Max	Mean	Std. Deviation	Skewness	Kurtosis
y	13	72.50	115.90	95.42	15.044	-0.195	1.658
X_1	13	1.00	21.00	7.46	5.882	0.688	3.075
X_2	13	26.00	71.00	48.15	15.561	-0.047	1.677
X_3	13	4.00	23.00	11.77	6.405	0.611	1.921
X_4	13	6.00	60.00	30.00	16.738	0.330	1.986

Source: Researcher's Computation (2022)

Table 28 presents the descriptive statistics for the Portland cement data. For y and X_1, the highest and lowest means are recorded respectively. The standard deviations for each variable indicate that the deviations are bigger. We observed that the predictors appear to be right-skewed, except X_2 variable which appears to be left-skewed. The coefficient of skewness is low, this show that the data are fairly symmetrical. The variables exhibit low kurtosis and tend to have light tails. This means that there are less values in the tails compared to a normal distribution, which further signified lack of outliers. Generally, the low values of kurtosis and skewness suggest that the variables' distribution are quite close to normality.

Table 29: **Correlation Matrix of Predictor Variables of the Original Data**

Variables	X_1	X_2	X_3	X_4
X_1	1.000			
X_2	0.229	1.000		
X_3	-0.824	-0.139	1.000	
X_4	-0.245	-0.973	0.030	1.000

Source: Researcher's Computation (2022)

Table 29 is the correlation matrix which presents the correlation among the variables in the data. The smallest correlation of 0.030 was obtained between X_3 and X_4, suggesting that there may be no statistically significant association between the two variables. The strongest correlation, -0.973, was found between X_2 and X_4, suggesting that the two variables can be considered very highly correlated and have much in common. The correlation coefficient of -0.824 shows fairly strong negative relationship between X_1 and X_3, and that the two variables can be considered highly correlated. The high correlation between the variables may suggest a possible presence of multicollinearity in the data.

Table 30: **VIF and Durbin-Watson Test of the Original Data**

VIF_1	VIF_2	VIF_3	VIF_4	dl	du	DW	p-value	α	ρ
38.50	254.42	46.87	282.51	0.574	2.094	2.053	0.421	0.05	-0.09

Source: Researcher's Computation (2022)

Table 30 presents the VIF and Durbin-Watson test of the original data. The DW statistic and the p-value are 2.053 and 0.421, respectively. The results show that $dl < DW < du$, this means, the test is inconclusive. However, the p-value (p = 0.421 > 0.05) from the Durbin-Watson test concludes that there is no autocorrelation in the error term at the 5% significance level with dl = 0.574 and du = 2.094. Moreover, all the variance inflation factors exceed 10. It is evident from the VIFs that the data suffers from the problem of multicollinearity. Therefore, we obtained an estimate of rho (ρ) and use the estimate of rho (ρ = -0.09) to transform the original data.

Table 31: **Correlation Matrix of Predictor Variables of the Transformed Data**

Variables	X_{1T}	X_{2T}	X_{3T}	X_{4T}
X_{1T}	1.000			
X_{2T}	0.246	1.000		
X_{3T}	-0.786	-0.066	1.000	
X_{4T}	-0.270	-0.980	-0.015	1.000

Source: Researcher's Computation (2022)

Table 31 presents the correlation matrix of the transformed variables. It was observed that the transformed variables X_1T and X_3T are highly negatively correlated with each other with correlation coefficient of -0.786, and therefore suggests that a linear relationship exists among them. The highest correlation of -0.980 was obtained between X_2T and X_4T, suggesting that the two variables are very highly correlated. Therefore, the high correlation between the transformed variables suggests a possible existence of multicollinearity in the data.

Table 32: **VIF and Durbin-Watson Test of the Transformed Data**

VIF_1	VIF_2	VIF_3	VIF_4	dl	du	DW	p-value	α
10.06	91.18	11.05	103.30	0.574	2.094	2.218	0.619	0.05

Source: Researcher's Computation (2022)

Table 32 presents the VIF and Durbin-Watson tests of the transformed data. The transformed data was diagnosed using the VIF and the Durbin-Watson (DW) tests. The DW statistic and the p-value are 2.218 and 0.619, respectively. The results show that $DW > du$, therefore the data does not suffer from first

order autoregressive scheme. However, since the values for *VIF* of all the predictor variables are greater than 10 ($VIFs > 10$), the model still suffer from multicollinearity.

Table 33: **Regression Analysis Results for Ordinary Least Square (OLS)**

Variable	Estimate	Std. Error	*t* value	*p*-value
Intercept	62.4054	70.0710	0.891	0.3991
X_1	1.5511	0.7448	2.083	0.0708
X_2	0.5102	0.7238	0.705	0.5009
X_3	0.1019	0.7547	0.135	0.8959
X_4	-0.1441	0.7091	-0.203	0.8441
R-squared	0.9824			
Adj. R-squared	0.9736			
MSE	5.9830			
RMSE	2.4460			
F-statistic: *F* (4, 8)	111.5		*p*-value	0.0000

Source: Researcher's Computation (2022)

Table 33 presents the estimated MSE, R-squared value and regression coefficients of OLS method. According to regression analysis results conducted by using OLS method, regression coefficients of all the predictors were found to be statistically insignificant at 5% significance level. However, the *F*-statistic for the overall model is significant [*F* (4, 8) = 111.5, *p*-value = 0.0000]. The OLS regression coefficients appears to have high standard errors. This may be as a result of the presence of multicollinearity which has excessively inflates the standard error, and has cause the regression coefficients to appear insignificant. The R-squared value of 0.9824 means the predictor variables can explain 98.24% of the variation in the response variable. The MSE of the OLS method is obtained as 5.9830.

Table 34: **Regression Analysis Results for Ridge Regression (RR)**

Variable	Estimate	Std. Error	*z* value	*p*-value
Intercept	74.8025	51.5924	1.450	0.1471
X_1	1.4159	0.5528	2.561	0.0104
X_2	0.3841	0.5361	0.716	0.4737
X_3	-0.0335	0.5580	-0.060	0.9522
X_4	-0.2682	0.5242	-0.512	0.6090
R-squared	0.9823			
Adj. R-squared	0.9735			
MSE	5.8161			
RMSE	2.4117			
K	0.97701			

Source: Researcher's Computation (2022)

Table 34 presents the estimated MSE, R-squared value and regression coefficients of RR method. According to regression analysis results conducted by using RR method, the variable X_1 has a significant effect on the dependent variable according to *p*-value (0.0104) at 5% significance level. The results of the analysis show a decrease in the standard error of the RR method in comparison to the OLS method. The R-square (0.9823) reveals that the predictors explain 98.23% of the variation in the response variable. The MSE and RMSE values are 5.8161 and 2.4117 respectively, when the value of K = 0.97701.

Table 35: **Regression Analysis Results for Lasso Regression**

Variable	Estimate	Std. Error	*z* value	*p*-value
Intercept	65.3456	69.4081	0.941	0.3465
X_1	1.5192	0.7381	2.058	0.0396
X_2	0.4803	0.7170	0.670	0.5030
X_3	0.0696	0.7478	0.093	0.9259
X_4	-0.1735	0.7024	-0.247	0.8050
R-squared	0.9824			
Adj. R-squared	0.9736			
MSE	5.9775			
RMSE	2.4449			
K	0.18307			

Source: Researcher's Computation (2022)

Table 35 presents the estimated MSE, R-squared value and regression coefficients of Lasso method. The regression analysis results conducted by using Lasso method revealed that the variable X_1 has a significant effect on the response variable according to *p*-value (0.0396) at 5% level of significance. The standard errors of the Lasso method show relatively small decrease to that of the OLS method. The R-square (0.9824) shows that 98.24% variation in the response variable can be explained by the predictors. The MSE and RMSE values obtained by the Lasso method are 5.9775 and 2.4449 respectively, when the value K is 0.18307.

Table 36: **Regression Analysis Results for Two-Stage Ridge Regression (TR)**

Variable	Estimate	Std. Error	*z* value	*p*-value
Intercept	30.2745	37.5707	0.806	0.4204
X_1T	1.8909	0.4281	4.417	0.0000
X_2T	0.8406	0.3899	2.156	0.0311
X_3T	0.4363	0.4224	1.033	0.3017
X_4T	0.1814	0.3828	0.474	0.6356
R-squared	0.9843			
Adj. R-squared	0.9765			
MSE	5.7399			
RMSE	2.3958			
K	0.001			

Source: Researcher's Computation (2022)

Table 36 presents the regression analysis results for TR method. According to regression analysis results conducted by using TR method, regression coefficients of the X_1T and X_2T variables were found to be statistically significant at 5% level of significance. The TR standard errors are decreased compared to OLS. The smaller the TR method standard error, the better the regression models fit to the data. The R-square (0.9843) shows that 98.43% variation in the response variable can be explained by the predictors. The MSE and RMSE values computed are 5.7399 and 2.3958 respectively, when the value of K = 0.001.

Table 37: **Regression Analysis Results for TLasso Regression**

Variable	Estimate	Std. Error	*z* value	*p*-value
Intercept	53.1040	67.8494	0.783	0.4338
X_1T	1.6562	0.7226	2.292	0.0219
X_2T	0.6055	0.7006	0.864	0.3875
X_3T	0.1972	0.7275	0.271	0.7863
X_4T	-0.0500	0.6888	-0.073	0.9421
R-squared	0.9845			
Adj. R-squared	0.9768			
MSE	5.9630			
RMSE	2.4419			
K	0.001			

Source: Researcher's Computation (2022)

Table 37 presents the regression analysis results for TLasso method. The regression analysis results conducted by using TLasso method revealed that the variable X_1T has a significant effect on the response variable according to *p*-value (0.0219) at 5% significance level. The TLasso method standard errors are relatively smaller in comparison to OLS method. The R-square (0.9845) shows that the predictor variables can explain 98.45% of the variation in the response variable. The MSE and RMSE values obtained by the TLasso method are 5.9630 and 2.4419 respectively, when the value K is 0.001.

Table 38: **Comparison of Regression Coefficients and MSEs of the Methods**

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	62.4054	74.8025	65.3456	30.2745	53.1040
β_1	1.5511	1.4159	1.5192	1.8909	1.6562
β_2	0.5102	0.3841	0.4803	0.8406	0.6055
β_3	0.1019	-0.0335	0.0696	0.4363	0.1972
β_4	-0.1441	-0.2682	-0.1735	0.1814	-0.0500
R-squared	0.9824	0.9823	0.9824	0.9843	0.9845
Adj. R-squared	0.9736	0.9735	0.9736	0.9765	0.9768
MSE	5.9830	5.8161	5.9775	5.7399	5.9630
RMSE	2.4460	2.4117	2.4449	2.3958	2.4419

Source: Researcher's Computation (2022)

Table 38 presents the regression coefficients and MSEs of the methods. According to the regression coefficients, the regression methods results are similar to each other. The findings reveal that the TR method has the lowest MSE value (5.7399) than the other methods, however, the TLasso method has the highest R-squared value (0.9845). The TR method has the least MSE and RMSE, and is considered best. Because of the existence of multicollinearity, the OLS method produces the worst results. Therefore, the TR method is most preferred method in fitting a linear model when the multicollinearity assumption is violated.

Table 39: **Comparison of QR Methods for τ = 0.25, 0.5 and 0.75**

Coefficient	QR		
	τ = 0.25	τ = 0.5	τ = 0.75
β_0	48.0545	-13.3367	47.6909
β_1	1.4976	2.3544	1.7084
β_2	0.7000	1.2798	0.6623
β_3	-0.0025	1.0074	0.3564
β_4	0.0296	0.6006	-0.0085
R-squared	0.9647	0.9775	0.9770
Adj. R-squared	0.9471	0.9663	0.9655
MSE	11.9769	7.6260	7.8119
RMSE	3.4608	2.7615	2.7950

Source: Researcher's Computation (2022)

Table 39 presents the regression coefficient estimation and the MSE at the 0.25, 0.5 and 0.75 quantile levels for QR. The intercept term at quantile level 0.25 and 0.75 has a positive sign while that of the quantile level at 0.5 has a negative sign. This might be due to the multicollinearity effect in the data. QR method at quantile level 0.5 has the lowest MSE and the highest R-squared value. The method is considered the best in this class of methods. However, the coefficients may not be estimated with great precision due to the existence of the multicollinearity in the data.

Table 40: **Coefficients and MSEs of RQR and LQR Methods**

Coefficient	RQR			LQR		
	$\tau = 0.25$	$\tau = 0.5$	$\tau = 0.75$	$\tau = 0.25$	$\tau = 0.5$	$\tau = 0.75$
β_0	86.7400	86.6812	82.6218	62.4726	93.3169	81.6878
β_1	0.9629	0.8527	1.3554	1.2730	1.3669	1.3585
β_2	0.2972	0.3054	0.3025	0.5219	0.1161	0.3142
β_3	-0.3512	-0.2682	-0.0076	0.0000	0.0000	0.0000
β_4	-0.3433	-0.3358	-0.3546	-0.1075	-0.4567	-0.3458
R-squared	0.9614	0.9570	0.9742	0.9625	0.9721	0.9741
Adj. R-squared	0.9420	0.9355	0.9614	0.9438	0.9582	0.9611
MSE	8.0716	8.9851	5.3820	7.8268	5.8221	5.4172
RMSE	2.8411	2.9975	2.3199	2.7976	2.4129	2.3275

Source: Researcher's Computation (2022)

The results presented in Table 40 shows the regression coefficient estimation and the MSE at the 0.25, 0.5 and 0.75 quantile levels for RQR and LQR. These methods allow the estimation of the quantile regression models using the L_1 and L_2-norms. We compare the performance of the quantile levels in RQR and LQR. The results through comparisons, we found that the RQR at quantile level 0.75 is the best as it has a high R-squared value (0.9742) and the least MSE value (5.3820) according to the interpretation of the data regarding the issues under study.

Table 41: **Coefficients and MSEs of TRQR and TLQR Methods**

Coefficient	TRQR			TLQR		
	$\tau = 0.25$	$\tau = 0.5$	$\tau = 0.75$	$\tau = 0.25$	$\tau = 0.5$	$\tau = 0.75$
β_0	93.6903	57.4066	68.2472	61.1333	15.9491	-3.0915
β_1	0.8749	1.7234	1.5563	1.2003	2.0725	2.2392
β_2	0.3161	0.5969	0.5069	0.6156	0.9978	1.1830
β_3	-0.3415	0.3035	0.1936	0.0000	0.6296	0.8723
β_4	-0.3187	-0.0903	-0.1470	-0.0170	0.3267	0.5219
R-squared	0.9376	0.9800	0.9655	0.9496	0.9831	0.9737
Adj. R-squared	0.9064	0.9700	0.9482	0.9244	0.9747	0.9605
MSE	14.8235	4.7516	8.1931	11.9629	4.0073	6.2558
RMSE	3.8501	2.1798	2.8624	3.4587	2.0018	2.5012

Source: Researcher's Computation (2022)

Table 41 presents the regression coefficient estimation and the MSE at the 0.25, 0.5 and 0.75 quantile levels for TRQR and TLQR. We therefore compared the performances of the quantile levels in TRQR methods and TLQR methods. We found that the TLQR at quantile level 0.5 is the best as it has a high R-squared value (0.9831) and the smallest MSE value (4.0073). The results show that the TLQR at quantile level 0.5 produced the most efficient estimates in terms of high R-square, lower RMSE and smallest MSE.

Table 42: **Comparison of TR, TLasso, TRQR and TLQR Methods**

Coefficient	TR	TLasso	TRQR	TLQR
			$\tau = 0.5$	$\tau = 0.5$
β_0	30.2745	53.1040	57.4066	15.9491
β_1	1.8909	1.6562	1.7234	2.0725
β_2	0.8406	0.6055	0.5969	0.9978
β_3	0.4363	0.1972	0.3035	0.6296
β_4	0.1814	-0.0500	-0.0903	0.3267
R-squared	0.9843	0.9845	0.9800	0.9831
Adj. R-squared	0.9765	0.9768	0.9700	0.9747
MSE	5.7399	5.9630	4.7516	4.0073
RMSE	2.3958	2.4419	2.1798	2.0018

Source: Researcher's Computation (2022)

Table 42 presents the overall comparison of the performances of TR, TLasso, TRQR and TLQR methods. Based on the R-squared values of the methods, TLasso appears best with respect to variation the predictor variables can explain in the response variable. According to the MSE criterion of each regression method, the TLQR method at quantile level 0.5 has the least MSE in comparison to the other methods. The results from comparisons of the TR, TLasso, TRQR at quantile level 0.5 and TLQR at quantile level 0.5 methods show that the TLQR method at quantile level 0.5 is superior to the other methods. Therefore, the TLQR method in the cases when the predictor variables are affected by multicollinearity is one of the successful ways to solve this issue.

Analysis of Historical Data

This dataset was discussed by Bayhan and Bayhan (1998). The data includes the weekly quantities of shampoos sold (y) as dependent variable and two predictor variables: weekly list prices (averages from selected supermarkets) of the firm's shampoos (X_1) and weekly list prices of a certain brand of soap, substituted from shampoos (X_2).

Table 43: **Summary Statistics of Variables**

Variable	N	Min	Max	Mean	Std. Deviation	Skewness	Kurtosis
y	15	34.16	35.17	34.59	0.320	0.427	1.940
X_1	15	101.30	111.30	106.30	3.282	-0.051	1.729
X_2	15	25.30	27.80	26.59	0.804	-0.041	1.799

Source: Researcher's Computation (2022)

Table 43 presents the summary statistics of variables. For X_1 and X_2, the highest and lowest means are recorded respectively. The standard deviations for each variable indicate that there are lesser deviations. The results revealed that the predictor variables are negatively skewed. The coefficient of skewness is low, which means the data are fairly symmetrical. The variables exhibit low kurtosis and tend to have light tails. This means that there are less values in the tails than to a normal distribution, which further signified lack of outliers. Generally, the low values of kurtosis and skewness suggest that the variables' distribution are close to normality.

Table 44: **Correlation Matrix of Original Data**

Variables	X_1	X_2
X_1	1.000	
X_2	0.994	1.000

Source: Researcher's Computation (2022)

The correlation coefficient of -0.824 shows a very strong positive relationship between X_1 and X_2, and that the predictor variables can be considered very highly correlated. The very highly correlation between the variables indicates that multicollinearity might exist in the data. These are as presented in Table 44.

Table 45: **VIF and Durbin-Watson Test of Original Data**

VIF_1	VIF_2	dl	du	DW	p-value	α	ρ
79.26	79.26	0.946	1.543	0.382	0.000	0.05	0.77

Source: Researcher's Computation (2022)

Table 45 presents the VIF and Durbin-Watson tests of original data. The Durbin-Watson test was employed to check for autocorrelation and the VIF to check for multicollinearity. The DW value and the p-value are 0.382 and 0.000, respectively. Since $DW < dl$, the data suffer from positive first order autoregressive scheme and since the $VIF > 10$, the data also suffer from multicollinearity. Hence, the data suffers from both problems. We correct the autocorrelation problem by data transformation. We then obtained an estimate of ρ (0.77) and used the estimate of ρ to transform the original data.

Table 46: **Correlation Matrix of Transformed Data**

Variables	X_{1T}	X_{2T}
X_{1T}	1.000	
X_{2T}	0.999	1.000

Source: Researcher's Computation (2022)

Table 46 shows the correlation matrix after the transformation of the data. It can be seen that the transformed variables X_1T and X_2T are very highly positively correlated with each other with correlation coefficient of 0.999, and this suggesting that a linear relationship exists among them. Hence, a possible multicollinearity problem might still exist in the data.

Table 47: **VIF and Durbin-Watson Test of Transformed Data**

VIF_1	VIF_2	dl	du	DW	p-value	α
512.646	512.646	0.946	1.543	1.549	0.2385	0.05

Source: Researcher's Computation (2022)

The transformed data was diagnosed using the VIF and the Durbin-Watson (DW) tests. The DW statistic and the p-value are 1.549 and 0.2385, respectively. The results show that $DW > du$, therefore the data does not suffer from first order autoregressive scheme. This shows that the autocorrelation problem in the data is solved. But, since the value for VIF is greater than 10 $(VIF > 10)$, the model still suffers from multicollinearity. This is confirmed by the VIF and Durbin-Watson Test of transformed data in Table 47.

Table 48: **Statistical Outputs of Ordinary Least Square (OLS)**

Variable	Estimate	Std. Error	*t* value	*p*-value
Intercept	25.4341	1.3617	18.678	0.000
X_1	0.1343	0.1106	1.214	0.248
X_2	-0.1925	0.4517	-0.426	0.678
R-squared	0.8055			
Adj. R-squared	0.7730			
MSE	0.0233			
RMSE	0.1526			
F-statistic: *F* (2, 12)	24.84		*p*-value	0.00005

Source: Researcher's Computation (2022)

Table 48 presents the regression analysis of OLS method. The results indicate that the regression coefficients of the predictor variables were statistically insignificant at 5% level of significance. However, the *F*-statistic for the overall model is significant [*F* (2, 12) = 24.84, *p*-value = 0.00005]. This regression coefficients appearing to be insignificant and the overall model being significant may be due to multicollinearity among the predictors. The R-squared value of 0.8055 means the predictors can explain 80.55% of the variation in the response variable. The MSE of the OLS method is obtained as 0.0233.

Table 49: **Statistical Outputs of Ridge Regression (RR)**

Variable	Estimate	Std. Error	*z* value	*p*-value
Intercept	25.3717	1.3119	19.340	0.0000
X_1	0.1111	0.0610	1.823	0.0683
X_2	-0.0976	0.2458	-0.397	0.6913
R-squared	0.8047			
Adj. R-squared	0.7722			
MSE	0.0225			
RMSE	0.1501			
K	0.08697			

Source: Researcher's Computation (2022)

Table 49 presents the regression analysis of ridge regression (RR) method. The findings reveal that the regression coefficients of the predictors were found to be statistically insignificant at 5% level of significance. The findings of the analysis showed a decrease in standard error of the RR method in comparison to the OLS method. The R-square (0.8047) reveals that the predictor variables can explain 80.47% of the variation in the response variable. The MSE and RMSE values are 0.0225 and 0.1501 respectively, when the value of K = 0.08697.

Table 50: **Statistical Outputs of Lasso Regression**

Variable	Estimate	Std. Error	*z* value	*p*-value
Intercept	25.3138	1.2790	19.792	0.0000
X_1	0.0873	0.0120	7.247	0.0000
X_2	0.0000	0.0018	0.000	1.0000
R-squared	0.8025			
Adj. R-squared	0.7696			
MSE	0.0218			
RMSE	0.1478			
K	0.02848			

Source: Researcher's Computation (2022)

Table 50 presents the regression coefficients, R-squared value and estimated MSE of Lasso regression method. The regression analysis results conducted by using Lasso method revealed that the variable X_1 has a significant effect on the response variable according to *p*-value (0.0000) at 5% significance level. Moreover, the regression coefficient of X_2 is shrink to zero. The Lasso standard errors show a decrease compared to OLS method. The R-square (0.8025) shows that 80.25% variation in the response variable can be explained by the predictors. The MSE and RMSE values obtained by the Lasso method are 0.0218 and 0.1478 respectively, when the value of K is 0.02848.

Table 51: **Statistical Outputs of Two-Stage Ridge Regression (TR)**

Variable	Estimate	Std. Error	*z* value	*p*-value
Intercept	22.7508	2.1623	10.522	0.0000
X_{1T}	0.0514	0.0440	1.168	0.2430
X_{2T}	0.2502	0.1920	1.303	0.1920
R-squared	0.9996			
Adj. R-squared	0.9995			
MSE	0.0268			
RMSE	0.1637			
K	0.001			

Source: Researcher's Computation (2022)

Table 51 is the regression analysis results of two-stage ridge regression (TR). According to regression analysis results, regression coefficients of the X_1T and X_2T variables were found to be statistically insignificant at 5% significance level. This shows the predictor variables after the transformation of the data do not have much effect on the dependent variable. The R-square (0.9996) shows that 99.96% variation in the response variable can be explained by the predictors. The MSE and RMSE values computed are 0.0268 and 0.1637, respectively.

Table 52: **Statistical Outputs of Two-Stage Lasso Regression (TLasso)**

Variable	Estimate	Std. Error	*z* value	*p*-value
Intercept	29.0907	2.0736	14.029	0.0000
X_{1T}	0.0473	0.0327	1.448	0.1480
X_{2T}	0.0240	0.1501	0.160	0.8730
R-squared	0.9998			
Adj. R-squared	0.9998			
MSE	0.0148			
RMSE	0.1216			
K	0.001			

Source: Researcher's Computation (2022)

Table 52 is the regression analysis results of two-stage Lasso regression (TLasso). The results revealed that the predictor variables were found to be statistically insignificant at 5% significance level. The TLasso standard error is lower than that of the OLS method. The R-squared value (0.9998) shows that the predictor variables can explain 99.98% of the variation in the response variable. The MSE and RMSE values obtained by the TLasso method are 0.0148 and 0.1216, respectively.

Table 53: **Comparison of Regression Coefficients and MSE of Methods**

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	25.4341	25.3717	25.3138	22.7508	29.0907
β_1	0.1343	0.1111	0.0873	0.0514	0.0473
β_2	-0.1925	-0.0976	0.0000	0.2502	0.0240
R-squared	0.8055	0.8047	0.8025	0.9996	0.9998
Adj. R-squared	0.7730	0.7722	0.7696	0.9995	0.9998
MSE	0.0233	0.0225	0.0218	0.0268	0.0148
RMSE	0.1526	0.1501	0.1478	0.1637	0.1216

Source: Researcher's Computation (2022)

Table 53 presents the coefficients of regression and MSE of methods. The regression coefficient of Lasso and TLasso seems to be much different from the other methods. Through comparison we found that the TLasso is the best as it has a high R-squared value (0.9998) and the smallest MSE value (0.0148) according to the interpretation of the data. Therefore, the TLasso method is most prefer method in fitting a linear model when assumptions of autocorrelation and multicollinearity are violated.

Table 54: **Regression Coefficients and MSE of QR Method**

Coefficient	QR		
	$\tau = 0.25$	$\tau = 0.5$	$\tau = 0.75$
β_0	21.5975	21.9003	25.7921
β_1	-0.0115	0.0369	0.2748
β_2	0.5305	0.3274	-0.7630
R-squared	0.5572	0.6640	0.6229
Adj. R-squared	0.4833	0.6080	0.5600
MSE	0.0530	0.0402	0.0452
RMSE	0.2303	0.2006	0.2125

Source: Researcher's Computation (2022)

Table 54 presents the estimated regression coefficients and the mean squared error at the 0.25, 0.5 and 0.75 quantile levels for QR. The QR methods at various quantile levels have different coefficients sign. The coefficients having wrong signs might be due to the multicollinearity effect in the data. The QR method at quantile level 0.5 has the lowest MSE and the highest R-squared value. The R-square value of 0.6640 means the predictor variables can explain 66.40% of the variation in the response variable. The QR method at quantile level 0.5 is considered the best in this class of methods. However, this quantile level is weak in interpreting the data with the multicollinearity and autocorrelation problems.

Table 55: **Comparison of RQR and LQR Methods**

Coefficient	RQR			LQR		
	$\tau = 0.25$	$\tau = 0.5$	$\tau = 0.75$	$\tau = 0.25$	$\tau = 0.5$	$\tau = 0.75$
β_0	21.2426	22.1655	25.9392	21.9963	21.9148	26.7055
β_1	0.0016	0.0498	0.2598	0.0048	0.0374	0.0751
β_2	0.4917	0.2662	-0.7090	0.4505	0.3250	0.0000
R-squared	0.5432	0.6890	0.6542	0.6021	0.6651	0.6729
Adj. R-squared	0.4670	0.6372	0.5966	0.5358	0.6093	0.6184
MSE	0.0438	0.0298	0.0331	0.0381	0.0321	0.0313
RMSE	0.2092	0.1726	0.1820	0.1952	0.1791	0.1770

Source: Researcher's Computation (2022)

Table 55 presents the comparison of performances of RQR and LQR at the quantile levels 0.25, 0.5 and 0.75. The LQR and RQR methods uses the L_1 and L_2-penalties respectively, in estimating the QR models. It can be seen that the RQR at quantile level 0.5 is the best as it has a high R-squared value of 0.6890 and the smallest MSE value of 0.0298. However, this quantile level is weak in interpreting the data with the multicollinearity and autocorrelation problems.

Table 56: **Comparison of TRQR and TLQR Methods**

Coefficient	TRQR			TLQR		
	$\tau = 0.25$	$\tau = 0.5$	$\tau = 0.75$	$\tau = 0.25$	$\tau = 0.5$	$\tau = 0.75$
β_0	15.0360	14.3818	13.8101	15.9662	22.1828	17.4388
β_1	0.0836	0.0865	0.0948	0.0632	0.0891	0.1133
β_2	0.4506	0.4649	0.4499	0.4947	0.1369	0.2290
R-squared	0.9998	0.9999	0.9998	0.9999	0.9999	0.9998
Adj. R-squared	0.9998	0.9998	0.9998	0.9998	0.9998	0.9997
MSE	0.0089	0.0079	0.0108	0.0078	0.0078	0.0131
RMSE	0.0946	0.0891	0.1041	0.0881	0.0881	0.1145

Source: Researcher's Computation (2022)

Table 56 presents the comparison of the performances of two-stage ridge quantile regression (TRQR) and two-stage Lasso quantile regression (TLQR) at the quantile levels 0.25, 0.5 and 0.75. The MSE values (0.0078) and the R-square values (0.9999) of the TLQR at quantile level 0.25 and TLQR at quantile level 0.5 methods are obtained to be the same up to 4 decimal places, indicating the same performance. Therefore, both TLQR at quantile level 0.25 and TLQR at quantile level 0.5 are the most preferred method in fitting to the data, a linear regression model.

Table 57: **Comparison of TR, TLasso, TRQR and TLQR**

Coefficient	TR	TLasso	TRQR	TLQR	
			$\tau = 0.5$	$\tau = 0.25$	$\tau = 0.5$
β_0	22.7508	29.0907	14.3818	15.9662	22.1828
β_1	0.0514	0.0473	0.0865	0.0632	0.0891
β_2	0.2502	0.0240	0.4649	0.4947	0.1369
R-squared	0.9996	0.9998	0.9999	0.9999	0.9999
Adj. R-squared	0.9995	0.9998	0.9998	0.9998	0.9998
MSE	0.0268	0.0148	0.0079	0.0078	0.0078
RMSE	0.1637	0.1216	0.0891	0.0881	0.0881

Source: Researcher's Computation (2022)

Table 57 presents the overall comparison of the performances of TR, TLasso, TRQR and TLQR methods. According to the R-squared values of the methods, TRQR and TLQR methods appear superior with respect to the variation the predictor variables can explain in the dependent variable. According to the MSE criterion of each regression method, the TLQR method at quantile level 0.25 and 0.5 have the smallest MSE value compared to the other methods. The results from comparisons of the methods revealed that both TLQR at quantile level 0.25 and TLQR at quantile level 0.5 are superior to the other methods. Therefore, the TLQR method is the best method in handling multicollinearity and autocorrelation problems in the dataset under study.

Discussion

The findings for two predictors revealed that with low autocorrelation and high degree of multicollinearity, TR method is superior to the other methods, especially with severe multicollinearity. However, the LQR and TLQR methods at quantile level 0.5 often competes favourably with the TR method. Increasing

the degree of multicollinearity provides an overall improvement in TR over the other regression methods for small predictor variables $(p = 2)$. We found that LQR method is either best or competes favourably with TLQR method when the degree of autocorrelation is very low with large sample size. Again, the comparisons of QR, RQR, LQR, TRQR and TLQR methods revealed that, both TRQR and TLQR methods at quantile level 0.5 perform better than the other methods. However, TRQR method produce better results than TLQR method in a model with small sample sizes. For sufficient high degrees of autocorrelation $(\rho = 0.9)$ with high multicollinearity, the TR method has the smallest MSE values and therefore outperforms the other methods for small predictors (p = 2). However, with large sample size (n = 500), the TLQR method at quantile level 0.5 is best.

The results for four predictor variables revealed that, TLQR method at 0.5 quantile level produce better results than the other methods in a model with severe multicollinearity $(\gamma^2 = 0.99)$ for the different samples size with very low degree of autocorrelation $(\rho = 0.1)$. However, some distinct patterns were identified. When n = 25 and n = 500 with sufficient high degree of autocorrelation $(\rho = 0.9)$, the TR method is best, and when $50 \leq n \leq 200$, the TLQR method at quantile level 0. 5 is generally best. Generally, comparison of the regularized and robust regression methods shows a general trend. Hence, both TR method and TLQR method at quantile level 0.5 are the best methods in handling multicollinearity and autocorrelation problems based on specific cases for four predictor variables.

The results for eight predictor variables in the presence of multicollinearity and autocorrelation revealed that, TR method outperforms the other methods.

But for severe multicollinearity level and many predictor variables, the TLasso method is superior and TR method is a close competitor. It has been revealed in this study that Lasso and TLasso methods performs better with many predictor variables ($p = 8$). For no or low degree of autocorrelation, the TLQR method at quantile level 0.5 has the least MSE values and therefore outperforms the other regression methods with small sample sizes ($25 \leq n \leq 50$). However, LQR method at quantile level 0.5 performs with large samples size ($200 \leq n \leq 500$) for eight predictor variables. The results indicated that the TLQR method at quantile level 0.5 outperforms the other methods when the autocorrelation is sufficiently high and across the different degrees of multicollinearity. However, for $25 \leq n \leq 50$ with high autocorrelation and severe multicollinearity, TR is best. Overall, we found that the TLQR method with an appropriate choice of quantile level is found suitable for addressing multicollinearity and autocorrelation problems with many predictor variables.

The study now looks at the use of the real data example to assess the performance of the methods. The study shown that the TLQR method still appears to be the best method in handling multicollinearity and autocorrelation problems in our real datasets. This research notes that the L_1-penalty added to the quantile regression loss function may have influence the performance of the TLQR method. Similarly, previous studies found that regularization (for instance, ridge and Lasso) in quantile regression has been proven to be effective in improving prediction accuracy (Bager, 2018; Li & Zhu, 2008; Li, Lin, & Xi, 2010; Suhail, Chand, & Kibria, 2020; Wu & Liu, 2009).

The results indicated that, when there is high degree of autocorrelation and multicollinearity level is also high, the OLS method has largest MSE with

inaccurate estimates of regression coefficients. These results confirm the works of Oyewole and Agunbiade (2020). They found that in the existence of multicollinearity and sufficiently strong autocorrelation, the OLS estimations of regression coefficients can be substantially erroneous (Oyewole & Agunbiade, 2020). If we take the MSE values as a criterion of comparison, we found that MSE values of TR is always less than the MSE values of RR. These findings are in line with the works of Hussein and Hytham (2021). They found that the TR method outperforms the RR, and that the regularization parameter values under the TR are always less than those under the RR.

Additionally, at no or low degree of autocorrelation and sufficiently low multicollinearity level with large sample size, the OLS estimate competes consistently with Lasso and TLasso methods. We found that the estimated MSE values of Lasso and TLasso methods decreases as the degree of multicollinearity increases with a smaller sample size. Therefore, the findings indicate that the sample size and number of predictor variables in the model are important in assessing the performances of the methods. According to Kristofer, Ghazi and Kibria (2010) MSE improves when the number of predictors is increased. They also found that when sample size increases, MSE lowers, even when the correlation between predictors is high. Additionally, the degree of autocorrelation and multicollinearity affects performances of regression methods. Zubair and Adenomon (2021) stated that, to avoid erroneous inferences when estimating regression model parameters, autocorrelation and multicollinearity between variables should be taken into account.

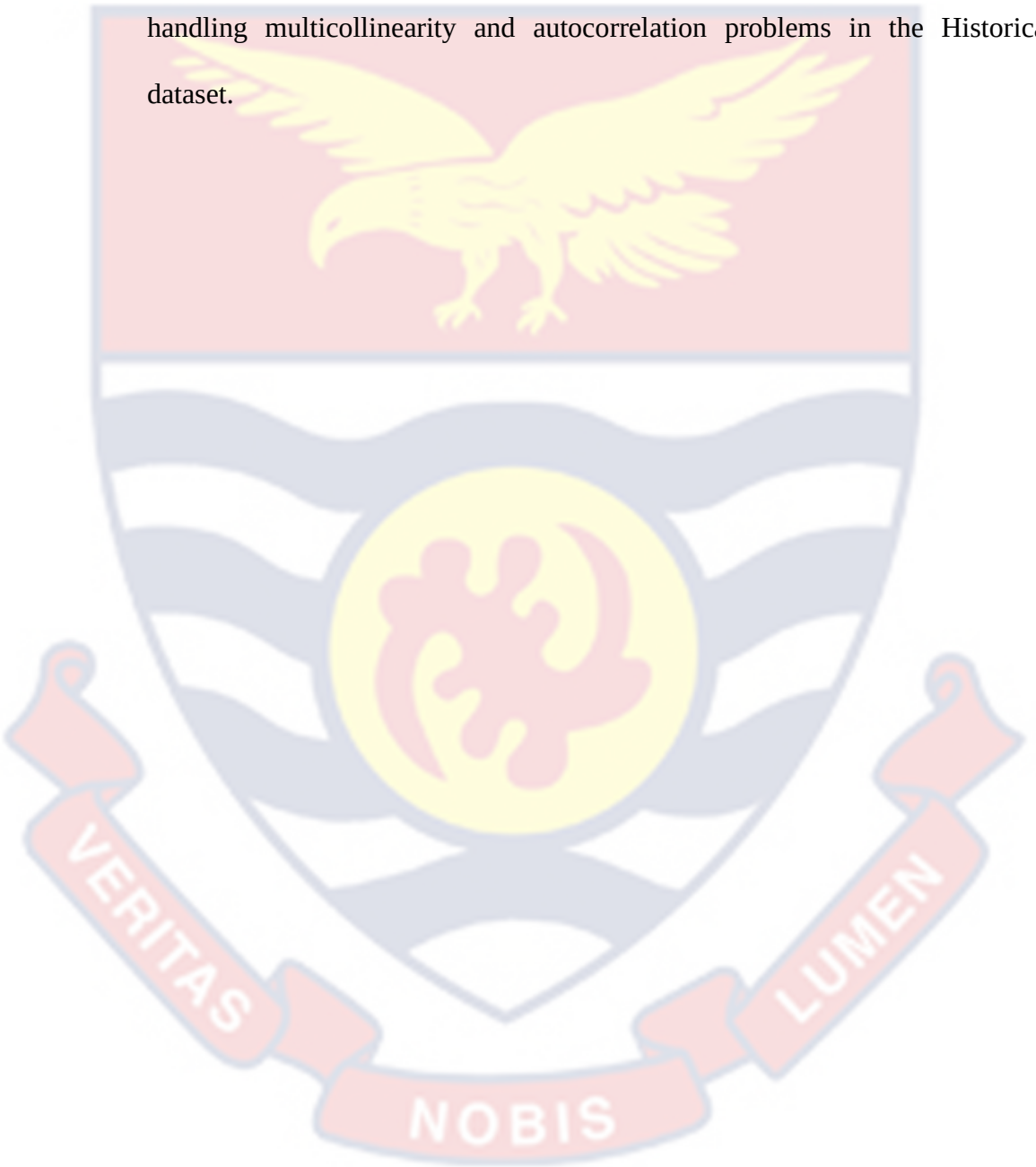
Chapter Summary

In this chapter, simulations of the regularized and robust regression methods were performed using two, four and eight predictor variables, 0.7, 0.9 and 0.99 multicollinearity levels, 0.1 and 0.9 degrees of autocorrelation, and samples of size 25, 50, 200 and 500 with 500 replications. Also, we compared the performances of the methods on two real data sets; the Portland cement data and the Historical data to determine which of the methods is the most effective in handling multicollinearity and autocorrelation problems in the datasets. The simulations and the real data analyses were performed using R software program.

The simulation results for two predictor variables show that the use of the TR method is generally effective in handling multicollinearity and autocorrelation problems for $25 \leq n \leq 200$. However, with large sample size (n = 500), the TLQR method at quantile level 0.5 was best. It was shown that, both TR method and TLQR method at quantile level 0.5 are the best methods in handling multicollinearity and autocorrelation problems in a dataset for four predictor variables. The simulation results for eight predictor variables revealed that the TLQR method at quantile level 0.5 outperforms the other methods when the autocorrelation is high and across the different degrees of multicollinearity with large sample size. However, for $25 \leq n \leq 50$ with high autocorrelation and severe multicollinearity, TR was best.

The real data was first explored to see the nature and spread of the data, and were further diagnosed using the variance inflation factor (VIF) and the Durbin-Watson (DW) tests to determine whether there is the presence of autocorrelation and/or multicollinearity in the datasets. The Portland cement data revealed the

existence of multicollinearity and the Historical data revealed the presence of both multicollinearity and autocorrelation. The study reveals that the TLQR method at quantile level 0.5 appears to be a suitable alternative in addressing the multicollinearity problem in the Portland cement dataset and also best in handling multicollinearity and autocorrelation problems in the Historical dataset.



CHAPTER FIVE

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

Overview

This chapter gives an overview of the various chapters and the conclusions that are drawn from the findings of the study. Based on the conclusions, relevant recommendations are provided.

Summary

The Ordinary Least Squares (OLS) possesses a very desirable property of being the Best Linear Unbiased Estimator (BLUE) when assumptions hold true. However, the OLS estimates can have extremely large mean squared error when there is multicollinearity. When the multicollinearity level gets higher, OLS estimates become imprecise, the model may have insignificant test, wider confidence interval, less interpretable are the parameters and the property of OLS being the BLUE does not hold anymore. Also, if the errors are correlated, then OLS can underestimate the standard error of the coefficients. When this happens, OLS estimate predictions will be inefficient. The literature reveals that the OLS estimates performance is unsatisfactory when autocorrelation and multicollinearity are present.

The topic under study seeks to examine different regularized and robust regression methods for handling multicollinearity and autocorrelation problems. Firstly, the performance of RR, Lasso, TR and TLasso methods were compared. Secondly, the performance of QR, RQR, LQR, TRQR and TLQR were also compared. Both real and simulated data are used in the study. Two real data sets were obtained for the purpose. Data was also simulated with three multicollinearity levels (0.7, 0.9 and 0.99), two degrees of autocorrelation (0.1

and 0.9), four different sample sizes (25, 50, 200 and 500) and three different number of predictor variables (2, 4 and 8). The simulation was replicated 500 times to obtain the approximate distribution. The R software program was used for the data analysis at various stages of the study. The regression methods applied for the analyses are OLS, RR, Lasso, TR, TLasso, QR, RQR, LQR, TRQR and TLQR. The study used three quantile levels (0.25, 0.5 and 0.75). The performances of these methods are compared to determine which method best addresses the problems of multicollinearity and autocorrelation.

The literature reveals a number of methods for addressing multicollinearity and autocorrelation problems, namely: generalized two stage ridge regression, feasible generalized restricted ridge regression, two-stage ridge regression, two-stage principal component regression, two-stage Liu regression, two-stage partial least square and two-stage $K-L$ method. The regularized and robust regression methods were reviewed thoroughly and their performance compared to arrive at the results of the study. These results are however different from what is in literature since it did not consider two-stage ridge quantile regression (TRQR) and two-stage Lasso quantile regression (TLQR) methods in handling multicollinearity and autocorrelation problems.

The study first presented the results from the simulated data for two predictor variables. The results from comparisons of OLS, RR, Lasso, TR and TLasso methods reveal that the TR method is superior to the other methods for both $\rho = 0.1$ and $\rho = 0.9$. Again, the comparisons of QR, RQR, LQR, TRQR and TLQR methods revealed that, both TRQR and TLQR methods at quantile level 0.5 showed superiority over the other methods. However, TRQR produce better outcomes than TLQR with small sample sizes. We compared the performances

of all the methods, and the results showed that, TR is best when multicollinearity is high with high autocorrelation for $25 \leq n \leq 200$. Moreover, with large sample size (n = 500), the TLQR method at quantile level 0.5 is best. The study further revealed that the estimated MSEs for TR is always less than that of RR when the multicollinearity and autocorrelation problems exist in a dataset. Overall, the use of the TR method for two predictor variables in a model is generally effective in handling multicollinearity and autocorrelation problems.

The simulation results for four predictor variables revealed that the TR method outperforms the OLS, RR, Lasso and TLasso methods for the samples size (25, 50, 200 and 500) and across the different degrees of multicollinearity (0.7, 0.9 and 0.99) for both $\rho = 0.1$ and $\rho = 0.9$. The results from comparisons of the methods show that, TLQR method at quantile level 0.5 is most preferred methods with very high multicollinearity level $(\gamma^2 = 0.99)$ at the different samples size (25, 50, 200 and 500) when $\rho = 0.1$. Moreover, TLQR method at quantile level 0.5 also performs better than the other methods (QR, RQR, LQR and TRQR) when $\rho = 0.9$. From the study findings, the TLQR method at 0.5 quantile level produce better results in a model with severe multicollinearity $(\gamma^2 = 0.99)$ for the different samples size when $\rho = 0.1$ than the other methods. When n = 25 and 500 with high autocorrelation $(\rho = 0.9)$, the TR method is best. However, when $50 \leq n \leq 200$, the TLQR method at quantile level 0.5 is generally best. Hence, both TR method and TLQR method at quantile level 0.5 are the best methods in handling multicollinearity and autocorrelation problems based on specific cases for four predictor variables.

The results of the study from the simulated data for eight predictor variables revealed the following findings. The results from comparisons of OLS, RR,

Lasso, TR and TLasso methods revealed that the TR method outperforms the other methods with high degree of multicollinearity. However, it is observed that with many predictor variables and very high degree of multicollinearity $(\gamma^2 = 0.99)$, TLasso has appreciable smaller MSE values and performs best. The simulation results when $\rho = 0.1$ show that the TLQR method at quantile level 0.5 outperforms the other regression methods with decreasing small sample sizes $(25 \leq n \leq 50)$. However, LQR method at quantile level 0.5 performs better than the other regression methods with large samples size $(200 \leq n \leq 500)$. The simulation results when $\rho = 0.9$ show that TLQR method at quantile level 0.5 outperforms the other methods (QR, RQR, LQR and TRQR) for the various samples size (25, 50, 200 and 500) and across the various degrees of multicollinearity (0.7, 0.9 and 0.99). Comparing the performance of all the methods, we observed that TLQR method at quantile level 0.5 outperforms the other methods with high autocorrelation $(\rho = 0.9)$. However, TR performs best with high autocorrelation and severe multicollinearity for $25 \leq n \leq 50$. Also, when $\rho = 0.1$, LQR produce better results with large sample sizes $(200 \leq n \leq 500)$. Whiles, TLQR produce better results with small sample sizes $(25 \leq n \leq 50)$.

Finally, the real data were diagnosed using the VIF and the Durbin-Watson (DW) tests to determine if there is the presence of autocorrelation and/or multicollinearity in the datasets. The tests revealed the presence of multicollinearity in the Portland cement dataset, and the presence of multicollinearity and autocorrelation in the Historical dataset. The study revealed that the TLQR method at quantile level 0.5 appears to be a suitable alternative in addressing the multicollinearity problem in the Portland cement

dataset. Also, the TLQR method still appears to be the best method in handling multicollinearity and autocorrelation problems in the Historical dataset.

Conclusions

The study found that in the existence of multicollinearity and strong autocorrelation, OLS estimations of regression coefficients can be substantially incorrect. Also, the OLS estimates could not perform well with regard to their MSE in the existence of autocorrelation and multicollinearity. With low levels of multicollinearity and large sample size, the OLS estimate competes with the Lasso and TLasso methods while the TLasso method also competes with TR when multicollinearity degree is very high with many predictor variables. It is also found from the simulation results that the MSEs for TR method is always less than that of RR method when the multicollinearity and autocorrelation problems exist in a dataset. We also observed that the TRQR and TLQR methods showed a significant improvement over the method of only quantile regression.

Whatever autocorrelation and multicollinearity degrees, the TR method has lowest MSE compared with the other methods for few predictor variables. Furthermore, sample size has a significant impact on method performance at all levels of autocorrelation and multicollinearity. However, TLQR method appears to improve in performance with large sample size. The study found that both the TR method and the TLQR method with an appropriate quantile level choice are the best methods in handling multicollinearity and autocorrelation problems in a dataset with few predictor variables depending on the sample size.

In existence of multicollinearity and sufficiently high degrees of autocorrelation for many predictor variables, regardless of the sample size, the

TLQR method is seen to be the best. However, in the presence of multicollinearity with less or no autocorrelation, the LQR method has minimum MSE compared to TLQR method for a large sample size. Also, under high degrees of autocorrelation and severe multicollinearity level, TR method is either best or competes also with TLQR method when there is a small sample size. Overall, the TLQR method with 0.5 quantile level is found to be suitable for addressing multicollinearity and autocorrelation problems with many predictor variables.

Recommendations

In developing linear models, multicollinearity and autocorrelation degrees between predictors should be considered to avoid erroneous inferences when estimating regression parameters. From the study, it has been observed that, the size of the sample has a considerable impact on effectiveness of regression methods. To effectively build a good model, it would be prudent to choose the appropriate samples with corresponding right regression method for data studies which include possible multicollinearity and autocorrelation issues.

The study has suggested the best estimation method that gives the least mean squared error under simultaneous multicollinearity and autocorrelation conditions. The performance of TR method is consistent with small sample sizes; while TLQR method is consistent with large sample sizes. Therefore, the TR method can be a good method in addressing multicollinearity and autocorrelation problems in a dataset with few predictor variables. However, for many predictor variables, TLQR method with 0.5 quantile level is what is found to be suitable in handling multicollinearity and autocorrelation problems simultaneously in regression analysis.

Finally, we would recommend for further studies that the same regularized methods and other robust options can be looked at in the existence of multicollinearity, autocorrelation and outliers. Further studies could also be conducted on assessing the performance of the regularized and robust regression methods on linear models which suffers from multicollinearity and higher order autoregressive schemes, such as second-order autoregressive process AR (2).

REFERENCES

Adepoju, A. A., & Ojo, O. O. (2018). Bayesian method for solving the problem of multicollinearity in regression. *Afrika Statistika*, *13*(3), 1823-1834.

Adlouni, S. E., Salaou, G., & St-Hilaire, A. (2018). Regularized Bayesian quantile regression. *Communication in Statistics-Simulation and Computation*, *47*(1), 277-293.

Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, *16*(2), 270-301.

Arowolo T. O., Adewale F. L., & Kayode, A. (2016). A comparative study of some method of handling multicollinearity in an autocorrelated error. *African Journal of Science and Technology (AJST) Science and Engineering Series*, *13*(2), 68-72.

Ayinde, K., Alabi, O. O., & Nwosu, U. I. (2021). Solving multicollinearity problem in linear regression model: The review suggests new idea of partitioning and extraction of the explanatory variables. *Journal of Mathematics and Statistics Studies*, *2*(1), 12-20.

Ayinde, K., Lukman, A. F., & 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.

Bager, A. S. (2018). Ridge parameter in quantile regression models: An application in biostatistics. *International Journal of Statistics and Applications*, *8*(2), 72-78.

Baye, M. R., & Parker, D. F. (1984). Combining ridge and principal component regression: a money demand illustration. *Communications in Statistics-Theory and Methods*, *13*(2), 197-205.

Bayhan, G. M., & Bayhan, M. (1998). Forecasting using autocorrelated errors and multicollinear predictor variables. *Computers & Industrial Engineering*, *34*(2), 413-421.

Beach, C. M., & MacKinnon, J. G. (1978). A maximum likelihood procedure for regression with autocorrelated errors. *Econometrica: Journal of the Econometric Society*, *46*(1), 51-58.

Belsley, D. A. (1991). Conditioning diagnostics: collinearity and weak data regression. John Wiley & Sons, New York.

Berk, R. A. (2020). *Statistical learning from a regression perspective* (3rd ed.). Cham, Switzerland: Springer Nature.

Bingham, N. H., & Fry, J. M. (2010). *Regression: Linear models in statistics*. Springer Science & Business Media.

Chang, X., & Yang, H. (2012). Combining two-parameter and principal component regression estimators. *Statistical Papers*, *53*(3), 549-562.

Chatelain, J. B., & Ralf, K. (2014). Spurious regressions and near-multicollinearity, with an application to aid, policies and growth. *Journal of Macroeconomics*, *39*, 85-96.

Cochrane, D., & Orcutt, G. H. (1949). Application of least square to relationship containing autocorrelated error terms. *Journal of the American Statistical Association*, *44*, 32-61.

Corlett, W. (1990). Multicollinearity. In: Eatwell, J., Milgate, M., Newman, P. (eds) *Econometrics*. The New Palgrave. Palgrave Macmillan, London.

Curto, J. D., & Pinto, J. C. (2011). The corrected vif (cvif). *Journal of Applied Statistics*, *38*(7), 1499-1507.

Davino, C., Romano, R., & Vistocco, D. (2022). Handling multicollinearity in quantile regression through the use of principal component regression. *METRON*, *80*(2), 153-174.

Dawoud, I., & Kaçıranlar, S. (2016). Two stage Liu regression estimator. *International Research Journal of Mathematics, Engineering and IT*, *3*(7), 7-35.

Demaris, A. (2004). *Regression with social data: Modeling continuous and limited response variables*. Hoboken, New Jersey: John Wiley & Sons, Inc.

Desviona, N., & Yanuar, F. (2020). Simulation study of autocorrelated error using Bayesian quantile regression. *Science and Technology Indonesia*, *5*(3), 70-74.

Dorugade, A. V. (2014). New ridge parameters for ridge regression. *Journal of the Association of Arab Universities for Basic and Applied Sciences*, *15*, 94-99.

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

Duzan, H., & Sharif, N. S. B. M. (2015). Ridge regression for solving the multicollinearity problem: review of methods and models. *Journal of Applied Sciences*, 15(3), 392–404.

Dwivedi, T. D., & Srivastava, V. K. (1978). On the minimum mean squared error estimators in a regression model. *Communications in Statistics-Theory and Methods*, *7*(5), 487-494.

Eledum, H. Y. A., & Alkhalifa, A. A. (2012). Generalized two stages ridge regression estimator GTR for multicollinearity and autocorrelated errors. *Canadian Journal on Science and Engineering Mathematics*, *3*(3), 79-85.

Eledum, H. Y. A., & Awadallah, H. H. (2021). A Monte Carlo study for dealing with multicollinearity and autocorrelation problems in linear regression using two stage ridge regression method. *Mathematics and Statistics*, *9*(5), 630-638.

Eledum, H. Y. A, & Zahri, M. (2013). Relaxation method for two stages ridge regression estimator. *International Journal of Pure and Applied Mathematics*, *85*(4), 653-667.

El-Fallah, M., & El-Salam, A. (2014). A note on partial least squares regression for multicollinearity (A comparative study). *International Journal of Applied Science and Technology*, *4*(1), 163-169.

Farrar, D. E., & Glauber, R. R. (1967). Multicollinearity in regression analysis: the problem revisited. *The Review of Economic and Statistics*, 92-107.

Firinguetti, L., Kibria, G., & Araya, R. (2017). Study of partial least squares and ridge regression methods. *Communications in Statistics-Simulation and Computation*, *46*(8), 6631-6644.

Flexeder, C. (2010). *Generalized lasso regularization for regression models*. Doctoral thesis, University of Munchen.

Frisch, R. (1934). *Statistical confluence analysis by means of complete regression systems*. Oslo: University Institute of Economics.

Gervini, D., & Yohai, V. J. (2002). A class of robust and fully efficient regression estimators. *The Annals of Statistics*, *30*(2), 583-616.

Giacalone, M., Panarello, D., & Mattera, R. (2018). Multicollinearity in regression: an efficiency comparison between Lp-norm and least squares estimators. *Quality & Quantity*, *52*(4), 1831-1859.

Gnat, S. (2020). Impact of the regularization of regression models on the results of the mass valuation of real estate. *Folia Oeconomica Stetinensia*, *20*(1), 163-176.

Goktas, A., & Akkus, O. (2020). Comparison of partial least squares with other prediction methods via generated data. *Journal of Statistical Computation and Simulation*, *90*(16), 3009-3024.

Greene, W. H. (2003). *Econometric analysis* (5th ed.). Upper Saddle River, NJ: Pearson Education.

Greene, W. H. (2012). *Econometric analysis* (7th ed.). Upper Saddle River, NJ: Pearson Education.

Gujarati, D. N., & Porter, D. C (2009). *Basic econometrics* (5th ed.). New York, NY: McGraw-Hill/Irwin.

Gujarati, D. N. (2003). *Basic econometrics* (4th ed.). New York, NY: McGraw-HiII/lrwin.

Hashem, H. (2014). *Regularized and robust regression methods for high dimensional data*. Doctoral dissertation, Department of Mathematical Sciences, Brunel University.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: data mining, inference, and prediction* (2nd ed.). New York: Springer.

Herawati, N., Nisa, K., Setiawan, E., Nusyirwan, N., & Tiryono, T. (2018). Regularized multiple regression methods to deal with severe multicollinearity. *International Journal of Statistics and Applications*, *8*(4), 167-172.

Hildreth, C., & Lu, J. Y. (1960). Demand relationships with autocorrelated disturbances. Agricultural Experiment Statistical Bulletin 276. East Lansing, MI: Michigan State University.

Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, *12*(1), 55-67.

Huber, P. J. (1964). Robust estimation of a location parameter. *Annals of Mathematical Statistics*, *35*(1), 73-101.

Irfan, M., Javed, M., & Raza, M. A. (2013). Comparison of shrinkage regression methods for remedy of multicollinearity problem. *Middle-East Journal of Scientific Research*, *14*(4), 570-579.

Jaya, I. G. N. M., Tantular, B., & Andriyana, Y. (2019, July). A Bayesian approach on multicollinearity problem with an Informative Prior. In *Journal of Physics: Conference Series* (Vol. 1265, No. 1, p. 012021). IOP Publishing.

Jegede, S. L., Lukman, A. F., Ayinde, K., & Odeniyi, K. A. (2022). Jackknife Kibria-Lukman M-estimator: simulation and application. *Journal of the Nigerian Society of Physical Sciences*, 251-264.

Jensen, D. R., & Ramirez, D. E. (2013). Revision: Variance inflation in regression. *Advances in Decision Sciences, 2013*, 1-15.

Johnson, S. R., Reimer, S. C., & Rothrock, T. P. (1973). Principal components and the problem of multicollinearity. *Metroeconomica, 25*(3), 306-317.

Kaciranlar, S., & Sakallioglu, S. (2001). Combining the liu estimator and the principal component regression estimator. *Communications in Statistics-Theory and Methods*, *30*(12), 2699-2705.

Kibria, B. M. G. (2003). Performance of some new ridge regression estimators. *Communications in Statistics – Simulation and Computation*, *32*, 419–435.

Klein, R. (1962). *An Introduction to Econometrics*. Prentice-Hall Pub., Englewood, Cliffs, N. J.

Koenker, R. (2004). Quantile regression for longitudinal data. *Journal of Multivariate Analysis*, *91*(1), 74-89.

Koenker, R., & Bassett, G. (1978). Regression quantiles. *Econometrica: Journal of the Econometric Society*, 33-50.

Kramer, W. (1980). Finite sample efficiency of ordinary least squares in the linear regression model with autocorrelated errors. *Journal of the American Statistical Association*, *75*(372), 1005-1009.

Lawless, J. F., & Wang, P. (1976). A simulation study of ridge and other regression estimators. *Communications in Statistics-Theory and Methods*, *5*(4):307–323.

Li, Q., Lin, N., & Xi, R. (2010). Bayesian regularized quantile regression. *Bayesian Analysis*, *5*(3), 533-556.

Li, Y., & Zhu, J. (2008). L1-norm quantile regression. *Journal of Computational and Graphical Statistics*, *17*(1), 163-185.

Liu, K. J. (1993). A new class of biased estimate in linear regression. *Communications in Statistics-Theory and Methods*, *22*(2), 393-402.

Lukman, A. F., Osowole, O. I., & Ayinde, K. (2015). Two stage robust ridge method in a linear regression model. *Journal of Modern Applied Statistical Methods*, *14*(2), 53-67.

Massy, W. F. (1965). Principal components regression in exploratory statistical research. *Journal of the American Statistical Association*, *60*(309), 234-256.

McDonald, G. C., & Galarneau, D. I. (1975). A Monte Carlo evaluation of some ridge-type estimators. *Journal of the American Statistical Association*, *70*(350), 407-416.

Mela, C. F., & Kopalle, P. K. (2002). The impact of collinearity on regression analysis: the asymmetric effect of negative and positive correlations. *Applied Economics*, *34*(6), 667-677.

Melkumova, L. E., & Shatskikh, S. Y. (2017). Comparing Ridge and LASSO estimators for data analysis. *Procedia engineering*, *201*, 746-755.

Midi, H., & Zahari, M. (2008). A simulation study on ridge regression estimators in the presence of outliers and multicollinearity. *Jurnal Teknologi*, 59-74.

Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to linear regression analysis* (5th ed.). John Wiley & Sons.

O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, *41*, 673-690.

Olaomi, J. O. (2008). Performance of the estimators of linear regression model with autocorrelated error terms which are also correlated with the geometric trended regressor. *European Journal of Scientific Research*, *20*(1), 187-196.

Onur, T. (2016). A comparative study on regression methods in the presence of multicollinearity. *Journal of Statisticians: Statistics and Actuarial Sciences*, *9*(2), 47-53.

Oyewole, O., & Agunbiade, D. A. (2020). Regression techniques in the presence of multicollinearity and autocorrelation phenomena: Monte Carlo approach. *Annals Computer Science Series*, 18, 70-77.

Ozbay, N., Kaçıranlar, S., & Dawoud, I. (2017). The feasible generalized restricted ridge regression estimator. *Journal of Statistical Computation and Simulation*, *87*(4), 753-765.

Ozkale, M. R. (2012). Combining the unrestricted estimators into a single estimator and a simulation study on the unrestricted estimators. *Journal of Statistical Computation and Simulation*, *82*(5), 653-688.

Ozkale, M. R., & Kaciranlar, S. (2007). The restricted and unrestricted two-parameter estimators. *Communication in Statistics-Theory and Methods*, *36*(15), 2707-2725.

Paris, S. J., & Winstein, C. B. (1954). Trend estimators and serial correlation. Chicago: Unpublished Cowles Commission, Discussion Paper.

Pati, K. D., Adnan, R., Rasheed, B. A., & MD. J, M. A. (2016). Estimation parameters using Bisquare weighted robust ridge regression BRLTS estimator in the presence of multicollinearity and outliers. In *AIP Conference Proceedings* (p. 060028). AIP Publishing LLC.

Pereira, J. M., Basto, M., & Da Silva, A. F. (2016). The logistic lasso and ridge regression in predicting corporate failure. *Procedia Economics and Finance*, *39*, 634-641.

Pfaffenberger, R. C., & Dielman, T. E. (1985). A comparison of robust ridge estimators. In *Proceedings of the American Statistical Association Business and Economic Statistics Section* (pp. 631-635).

Reed, C. (2011). *Bayesian parameter estimation and variable selection for quantile regression*. Doctoral dissertation, School of Information Systems, Computing and Mathematics, Brunel University.

Rousseeuw, P. J. (1984). Least median of squares regression. *Journal of the American statistical association*, *79*(388), 871-880.

Rousseeuw, P. J, & Yohai, V. (1984). Robust regression by means of S-estimators. In *Robust and nonlinear time series analysis* (pp. 256-272). Springer, New York, NY.

Samkar, H., & Alpu, O. (2010). Ridge regression based on some robust estimators. *Journal of Modern Applied Statistical Methods*, *9*(2), 495-501.

Stein, C. (1956). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. In *Proceedings of the Third Berkeley symposium on mathematical statistics and probability* (Vol. 1, No. 1, pp. 197-206).

Stuart, C. (2011). Robust regression. *Department of Mathematical Sciences, Durham University*.

Suhail, M., Chand, S., & Kibria, B. M. G. (2020). Quantile based estimation of biasing parameters in ridge regression model. *Communications in Statistics - Simulation and Computation*, *49*(10), 2732-2744.

Tanizaki, H. (2003). On regression models with autocorrelated error: small sample properties. *International Journal of Pure and Applied Mathematics*, *5*(2), 159-172.

Tareghian, R., & Rasmussen, P. F. (2013). Statistical downscaling of precipitation using quantile regression. *Journal of Hydrology*, *487*, 122-135.

Thisted, R. A. (1976). *Ridge regression, minimax estimation, and empirical Bayes methods. Division of Biostatistics* (No. 28). Stanford University, Technical Report.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*(1), 267-288.

Usman, U., Zakari, Y., & Musa, Y. (2017). A comparative study on the prediction performance methods of handling multicollinearity based on normal and uniform distributions. *Journal of Basic and Applied Research International*, *22*(3), 111-117.

van Wieringen, W. N. (2021). Lecture notes on ridge regression.

Wang, H., Li, G., & Jiang, G. (2007). Robust regression shrinkage and consistent variable selection through the LAD-Lasso. *Journal of Business & Economic Statistics*, *25*(3), 347-355.

Weisberg, S. (2005). *Applied linear regression* (3rd ed.). Hoboken, New Jersey, NJ: John Wiley & Sons.

Wold, H. (1966). Estimation of principal components and related models by iterative least squares. *Multivariate analysis*, 391-420.

Woods, H., Steinour, H. H., & Starke, H. R. (1932). Effect of composition of Portland cement on heat evolved during hardening. *Industrial & Engineering Chemistry*, *24*(11), 1207-1214.

Wu, Y., & Liu, Y. (2009). Variable selection in quantile regression. *Statistica Sinica*, *19*(2), 801-817.

Yang, H., & Chang, X. (2010). A new two-parameter estimator in linear regression. *Communication in Statistics-Theory and Methods*, *39*(6), 923-934.

Yanuar, F., Yozza, H., Firdawati, F., Rahmi, I., & Zetra, A. (2019). Applying bootstrap quantile regression for the construction of a low birth weight model. *Makara Journal of Health Research*, *23*(2), 90-95.

Yohai, V. J. (1987). High breakdown-point and high efficiency robust estimates for regression. *The Annals of statistics*, *15*(20), 642-656.

Yohai, V. J., & Zamar, R. H. (1988). High breakdown-point estimates of regression by means of the minimization of an efficient scale. *Journal of the American statistical association*, *83*(402), 406-413.

Yu, K., Alkenani, A., & Alhamzawi, R. (2012). Penalized flexible Bayesian quantile regression. *Applied Mathematics*, 3, 2155-2168.

Yuan, Y., & Yin, G. (2010). Bayesian quantile regression for longitudinal studies with nonignorable missing data. *Biometrics*, *66*(1), 105-114.

Zaikarina, H., Djuraidah, A., & Wigena, A. H. (2016). Lasso and ridge quantile regression using cross validation to estimate extreme rainfall. *Global Journal of Pure and Applied Mathematics*, *12*(3), 3305-3314.

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *67*(2), 301-320.

Zubair, M. A., & Adenomon, M. O. (2021a). Comparison of estimators efficiency for linear regressions with joint presence of autocorrelation and multicollinearity. *Science World Journal*, 16(2), 103-109.

Zubair, M. A., & Adenomon, M. O. (2021b). Robustness test of the two stage K-L estimator in models with multicollinear regressors and autocorrelated error term. *Asian Journal of Probability and Statistics*, *14*(3), 22-33.

APPENDICES

APPENDIX A

SAMPLE SIMULATION CODES IN R

```
library(MASS)          # Call the package MASS
library(lmridge)       # Call package lmridge
library(islasso)       # Call package islasso
library(quantreg)      # Call package quantreg
library(hqreg)         # Call package hqreg
library(glmnet)        # Call package glmnet
library(Matrix)        # Call package Matrix
rep <- 500             # Number of replications
N <- 500               # Population size
n <- 200               # Sample size
p <- 2                 # Number of predictors
set.seed(22)
# True values for regression coefficients
b0 <- 0                # True value for the intercept
b1 <- 1                # True value for Beta1
b2 <- 1                # True value for Beta2
# Empty vectors for OLS
beta0<- c()            # Empty vector for storing the simulated intercept
beta1<- c()            # Empty vector for storing the simulated Beta1
beta2<- c()            # Empty vector for storing the simulated Beta2
mseols<- c()           # Empty vector for storing the MSE of OLS
# Empty vectors for RIDGE
betar0<- c()           # Empty vector for storing the simulated intercept
betar1<- c()           # Empty vector for storing the simulated Beta1
betar2<- c()           # Empty vector for storing the simulated Beta2
mser<- c()             # Empty vector for storing the MSE of Ridge
hk<- c()               # Empty vector for Horl and Konard's method
lw<- c()               # Empty vector for Lawless and Wang's method
# Empty vectors for LASSO
betal0<- c()           # Empty vector for storing the simulated intercept
betal1<- c()           # Empty vector for storing the simulated Beta1
betal2<- c()           # Empty vector for storing the simulated Beta2
mselr<- c()            # Empty vector for storing the MSE of Lasso
# Empty vectors for TWO STAGE RIDGE
betatr0<- c()          # Empty vector for storing the simulated intercept
betatr1<- c()          # Empty vector for storing the simulated Beta1
betatr2<- c()          # Empty vector for storing the simulated Beta2
msetr<- c()            # Empty vector for storing the MSE of TR
hk<- c()               # Empty vector for Horl and Konard's method
lw<- c()               # Empty vector for Lawless and Wang's method
# Empty vectors for TWO STAGE LASSO
betatlasso0<- c()      # Empty vector for storing the simulated intercept
```

```
betatlasso1<- c()       # Empty vector for storing the simulated Beta1
betatlasso2<- c()       # Empty vector for storing the simulated Beta2
msetlasso<- c()         # Empty vector for storing the MSE of TLasso
# Empty vectors for QR (0.25)
betaQl0<- c()           # Empty vector for storing the simulated intercept
betaQl1<- c()           # Empty vector for storing the simulated Beta1
betaQl2<- c()           # Empty vector for storing the simulated Beta2
mseQl<- c()             # Empty vector for storing the MSE of QR(0.25)
# Empty vectors for QR (0.50)
betaQm0<- c()           # Empty vector for storing the simulated intercept
betaQm1<- c()           # Empty vector for storing the simulated Beta1
betaQm2<- c()           # Empty vector for storing the simulated Beta2
mseQm<- c()             # Empty vector for storing the MSE of QR(0.50)
# Empty vectors for QR (0.75)
betaQu0<- c()           # Empty vector for storing the simulated intercept
betaQu1<- c()           # Empty vector for storing the simulated Beta1
betaQu2<- c()           # Empty vector for storing the simulated Beta2
mseQu<- c()             # Empty vector for storing the MSE of QR(0.75)
# Empty vectors for RQR (0.25)
betaRQl0<- c()          # Empty vector for storing the simulated intercept
betaRQl1<- c()          # Empty vector for storing the simulated Beta1
betaRQl2<- c()          # Empty vector for storing the simulated Beta2
mseRQl<- c()            # Empty vector for storing the MSE of RQR(0.25)
# Empty vectors for RQR (0.50)
betaRQm0<- c()          # Empty vector for storing the simulated intercept
betaRQm1<- c()          # Empty vector for storing the simulated Beta1
betaRQm2<- c()          # Empty vector for storing the simulated Beta2
mseRQm<- c()            # Empty vector for storing the MSE of RQR(0.50)
# Empty vectors for RQR (0.75)
betaRQu0<- c()          # Empty vector for storing the simulated intercept
betaRQu1<- c()          # Empty vector for storing the simulated Beta1
betaRQu2<- c()          # Empty vector for storing the simulated Beta2
mseRQu<- c()            # Empty vector for storing the MSE of RQR(0.75)
# Empty vectors for LQR (0.25)
betaLQl0<- c()          # Empty vector for storing the simulated intercept
betaLQl1<- c()          # Empty vector for storing the simulated Beta1
betaLQl2<- c()          # Empty vector for storing the simulated Beta2
mseLQl<- c()            # Empty vector for storing the MSE of LQR(0.25)
# Empty vectors for LQR (0.50)
betaLQm0<- c()          # Empty vector for storing the simulated intercept
betaLQm1<- c()          # Empty vector for storing the simulated Beta1
betaLQm2<- c()          # Empty vector for storing the simulated Beta2
mseLQm<- c()            # Empty vector for storing the MSE of LQR(0.50)
# Empty vectors for LQR (0.75)
betaLQu0<- c()          # Empty vector for storing the simulated intercept
betaLQu1<- c()          # Empty vector for storing the simulated Beta1
```

```
betaLQu2<- c()        # Empty vector for storing the simulated Beta2
mseLQu<- c()          # Empty vector for storing the MSE of LQR(0.75)
# Empty vectors for TWO STAGE RQR (0.25)
betaTRQl0<- c()       # Empty vector for storing the simulated intercept
betaTRQl1<- c()       # Empty vector for storing the simulated Beta1
betaTRQl2<- c()       # Empty vector for storing the simulated Beta2
mseTRQl<- c()         # Empty vector for storing the MSE of TRQR(0.25)
# Empty vectors for TWO STAGE RQR (0.50)
betaTRQm0<- c()       # Empty vector for storing the simulated intercept
betaTRQm1<- c()       # Empty vector for storing the simulated Beta1
betaTRQm2<- c()       # Empty vector for storing the simulated Beta2
mseTRQm<- c()         # Empty vector for storing the MSE of TRQR(0.50)
# Empty vectors for TWO STAGE RQR (0.75)
betaTRQu0<- c()       # Empty vector for storing the simulated intercept
betaTRQu1<- c()       # Empty vector for storing the simulated Beta1
betaTRQu2<- c()       # Empty vector for storing the simulated Beta2
mseTRQu<- c()         # Empty vector for storing the MSE of TRQR(0.75)
# Empty vectors for TWO STAGE LQR (0.25)
betaTLQl0<- c()       # Empty vector for storing the simulated intercept
betaTLQl1<- c()       # Empty vector for storing the simulated Beta1
betaTLQl2<- c()       # Empty vector for storing the simulated Beta2
mseTLQl<- c()         # Empty vector for storing the MSE of TLQR(0.25)
# Empty vectors for TWO STAGE LQR (0.50)
betaTLQm0<- c()       # Empty vector for storing the simulated intercept
betaTLQm1<- c()       # Empty vector for storing the simulated Beta1
betaTLQm2<- c()       # Empty vector for storing the simulated Beta2
mseTLQm<- c()         # Empty vector for storing the MSE of TLQR(0.50)
# Empty vectors for TWO STAGE LQR (0.75)
betaTLQu0<- c()       # Empty vector for storing the simulated intercept
betaTLQu1<- c()       # Empty vector for storing the simulated Beta1
betaTLQu2<- c()       # Empty vector for storing the simulated Beta2
mseTLQu<- c()         # Empty vector for storing the MSE of TLQR(0.75)

# Multicollinear predictors
gamma<- sqrt(0.90)    # r=0.90
u1<- rnorm(N)
u2<- rnorm(N)
u3<- rnorm(N)
x1<- sqrt(1-gamma^2)*u1+gamma*u3
x2<- sqrt(1-gamma^2)*u2+gamma*u3
xmatrix<- cbind.data.frame(x1,x2) # Dataframe of X variables
# Autocorrelated errors
rho<- 0.9             # Set value of rho
var_vt<- 1            # Variance vt
ett<- c()
pett<- c()
```

```
vt<- rnorm(N,0,var_vt)   # Generating vt
# Generating the first et
sdv<- sqrt(var_vt)
sigmaeps<- sdv/sqrt(1-rho^2)
et1<- rnorm(1,0,sigmaeps)
# Computing the first pe_(t-1)=et1-vt1
pet1<- et1-vt[1]
ets<- et1
pets<- pet1
# If t>1 pe_(t-1)=pet and et=pe_(t-1)+vt
i<- 2
while(i<=N){
  pets<- rho*ets
  ets<- pets+vt[i]
  pett<- append(pett[],pets)
  ett<- append(ett[],ets)
  i<- i+1
}
eps<- c(et1,ett)
y <- b0+b1*x1+b2*x2+eps
mydata<- cbind.data.frame(y, xmatrix)        # create population data frame
x<- data.matrix(xmatrix[,c('x1','x2')])

# ORDINARY LEAST SQUARES METHOD
for (i in 1:rep){#start the loop
  df<- data.frame(mydata[sample(N,n),])      # select a random sample
  ols<- lm(y~x1+x2,data=df)                  # Estimate OLS model
  summary(ols)
  ols$coef
  olsanova<- anova(ols)
  # Store Regression Coefficients
  beta0<- append(beta0,ols$coefficients[1])
  beta1<- append(beta1,ols$coefficients[2])
  beta2<- append(beta2,ols$coefficients[3])
  # Store MSE
  mseols<- append(mseols,anova(ols)$'Mean Sq'[3])

# RIDGE REGRESSION METHOD
  model.ridge<- lm.ridge(y~ x1+x2, data=df, lambda=seq(0.0,1,0.001))
  whichIsBest <- which.min(model.ridge$GCV)      # Determine min GCV and
k
  coefr<- coef(model.ridge)[whichIsBest,]    # Coefficient corresponding to
best k
  # Store regression coefficients
  betar0<- append(betar0,coefr[1])
  betar1<- append(betar1,coefr[2])
```

```
  betar2<- append(betar2,coefr[3])
  # Store methods of k
  hk<- append(hk,model.ridge$kHKB)
  lw<- append(lw,model.ridge$kLW)
  # MSE for Ridge
  yhat_r<- coefr[1] + as.matrix(df[,2:3])%*%coefr[2:3]
  resid.ridge <- df[,1] - yhat_r
  x_r<- as.matrix(df[,2:3])
  lambda.ridge <- seq(0,1,0.001)[which.min(model.ridge$GCV)]    # Best k
  d <- svd(as.matrix(x_r[,1:2]))$d
  dfree_r <- n - sum(d^2/(lambda.ridge+d^2))
  MSE.r <- sum(resid.ridge^2)/dfree_r
  mser<- append(mser,MSE.r)

# LASSO REGRESSION METHOD
  model.lasso <- islasso(y ~ x1+x2, data=df, family = gaussian, alpha=1)
  coefl<- coef(model.lasso)
  # Store regression coefficients
  betal0<- append(betal0,coefl[1])
  betal1<- append(betal1,coefl[2])
  betal2<- append(betal2,coefl[3])
  # MSE for Lasso
  yhat_l<- coefl[1] + as.matrix(df[,2:3])%*%coefl[2:3]
  resid.lasso <- df[,1] - yhat_l
  dfree_l <- model.lasso$df.residual
  MSE.lr <- sum(resid.lasso^2)/dfree_l
  mselr<- append(mselr,MSE.lr)

# TWO STAGE RIDGE REGRESSION METHOD
  model.ts<- lm(y~x1+x2,data=df)
  m<- length(residuals(model.ts))
  cor(residuals(model.ts)[-1], residuals(model.ts)[-m])
  Phi<- cor(residuals(model.ts)[-1], residuals(model.ts)[-m])
  X<- model.matrix(model.ts)
  Sigma<- diag(m)
  Sigma<- (Phi)^abs(row(Sigma)-col(Sigma))
  S<- chol(Sigma)
  S_inv<- solve(t(S))
  SX_matr<- S_inv%*%X
  SX<- as.matrix(SX_matr[,2:3])
  SY<- S_inv%*%(df$y)
  df2<- cbind.data.frame(SY, SX)
  model.tr<- lm.ridge(SY~SX, data=df2 ,lambda=seq(0.0,1,0.001))
  whichIsBest <- which.min(model.tr$GCV) # Determine min GCV and k
  coeftr<- coef(model.tr)[whichIsBest,]        # Coefficient corresponding to
best k
```

```
  # Store regression coefficients
  betatr0<- append(betatr0,coeftr[1])
  betatr1<- append(betatr1,coeftr[2])
  betatr2<- append(betatr2,coeftr[3])
  # Store methods of k
  hk<- append(hk,model.tr$kHKB)
  lw<- append(lw,model.tr$kLW)
  # MSE for Two Stage Ridge
  yhat_tr<- coeftr[1] + as.matrix(df2[,2:3])%*%coeftr[2:3]
  resid.tr <- df2[,1] - yhat_tr
  lambda.tr <- seq(0,1,0.001)[which.min(model.tr$GCV)]          # Best k
  d <- svd(as.matrix(SX[,1:2]))$d
  dfree_tr <- n - sum(d^2/(lambda.tr+d^2))
  MSE.tr <- sum(resid.tr^2)/dfree_tr
  msetr<- append(msetr,MSE.tr)

# TWO STAGE LASSO REGRESSION METHOD
  model.ts<- lm(y~x1+x2,data=df)
  m<- length(residuals(model.ts))
  cor(residuals(model.ts)[-1], residuals(model.ts)[-m])
  Phi<- cor(residuals(model.ts)[-1], residuals(model.ts)[-m])
  X<- model.matrix(model.ts)
  Sigma<- diag(m)
  Sigma<- (Phi)^abs(row(Sigma)-col(Sigma))
  S<- chol(Sigma)
  S_inv<- solve(t(S))
  SX<- S_inv%*%X
  SX_mat<- as.matrix(SX[,2:3])
  SY<- S_inv%*%(df$y)
  df3<- cbind.data.frame(SY, SX_mat)
  model.tlasso <- islasso(SY ~ SX_mat, data=df3, family = gaussian, alpha=1)
  coeftlasso<- coef(model.tlasso)
  # Store regression coefficients
  betatlasso0<- append(betatlasso0,coeftlasso[1])
  betatlasso1<- append(betatlasso1,coeftlasso[2])
  betatlasso2<- append(betatlasso2,coeftlasso[3])
  # MSE for Two Stage Lasso
  yhat_tlasso<- coeftlasso[1] + as.matrix(df3[,2:3])%*%coeftlasso[2:3]
  resid.tlasso <- df3[,1] - yhat_tlasso
  dfree_tlasso <- model.tlasso$df.residual
  MSE.tlasso <- sum(resid.tlasso^2)/dfree_tlasso
  msetlasso<- append(msetlasso,MSE.tlasso)

# QUANTILE REGRESSION METHOD (0.25)
  model.Ql<- rq(y~x1+x2, tau=0.25, data=df)        # Estimate QR (0.25)
model
```

```
  coefQl<- coef(model.Ql)
  # Store Regression Coefficients
  betaQl0<- append(betaQl0,coefQl[1])
  betaQl1<- append(betaQl1,coefQl[2])
  betaQl2<- append(betaQl2,coefQl[3])
  # MSE for QR(0.25)
  yhat_Ql<- coefQl[1] + as.matrix(df[,2:3])%*%coefQl[2:3]
  resid.Ql <- df[,1] - yhat_Ql
  df_Ql <- summary(model.Ql)$rdf
  MSE.Ql <- sum(resid.Ql^2)/df_Ql
  mseQl<- append(mseQl,MSE.Ql)

# QUANTILE REGRESSION METHOD (0.50)
  model.Qm<- rq(y~x1+x2, tau=0.5, data=df)          #   Estimate   QR   (0.50)
model
  coefQm<- coef(model.Qm)
  # Store Regression Coefficients
  betaQm0<- append(betaQm0,coefQm[1])
  betaQm1<- append(betaQm1,coefQm[2])
  betaQm2<- append(betaQm2,coefQm[3])
  # MSE for QR(0.50)
  yhat_Qm<- coefQm[1] + as.matrix(df[,2:3])%*%coefQm[2:3]
  resid.Qm <- df[,1] - yhat_Qm
  df_Qm <- summary(model.Qm)$rdf
  MSE.Qm <- sum(resid.Qm^2)/df_Qm
  mseQm<- append(mseQm,MSE.Qm)

# QUANTILE REGRESSION METHOD (0.75)
  model.Qu<- rq(y~x1+x2, tau=0.75, data=df)         #   Estimate   QR   (0.75)
model
  coefQu<- coef(model.Qu)
  # Store Regression Coefficients
  betaQu0<- append(betaQu0,coefQu[1])
  betaQu1<- append(betaQu1,coefQu[2])
  betaQu2<- append(betaQu2,coefQu[3])
  # MSE for QR(0.75)
  yhat_Qu<- coefQu[1] + as.matrix(df[,2:3])%*%coefQu[2:3]
  resid.Qu <- df[,1] - yhat_Qu
  df_Qu <- summary(model.Qu)$rdf
  MSE.Qu <- sum(resid.Qu^2)/df_Qu
  mseQu<- append(mseQu,MSE.Qu)

# RIDGE QUANTILE REGRESSION METHOD (0.25)
  X<- as.matrix(df[,2:3])
  y_r<- df$y
  grid.r<- 10^seq(-3, 5, length = 100)
```

```
model.RQl<- cv.hqreg(X, y_r, tau=0.25, method="quantile", alpha=0, lambda=grid.r)                    # Estimate RQR (0.25) model
coefRQl<- coef(model.RQl,lambda="lambda.min")
# Store Regression Coefficients
betaRQl0<- append(betaRQl0,coefRQl[1])
betaRQl1<- append(betaRQl1,coefRQl[2])
betaRQl2<- append(betaRQl2,coefRQl[3])
# MSE for RQR(0.25)
yhat_RQl<- coefRQl[1] + as.matrix(df[,2:3])%*%coefRQl[2:3]
resid.RQl <- df[,1] - yhat_RQl
MSE.RQl <- mean(resid.RQl^2)
mseRQl<- append(mseRQl,MSE.RQl)

# RIDGE QUANTILE REGRESSION METHOD (0.50)
X<- as.matrix(df[,2:3])
y_r<- df$y
grid.r<- 10^seq(-3, 5, length = 100)
model.RQm<- cv.hqreg(X, y_r, tau=0.5, method="quantile", alpha=0, lambda=grid.r)                    # Estimate RQR (0.50) model
coefRQm<- coef(model.RQm, lambda="lambda.min")
# Store Regression Coefficients
betaRQm0<- append(betaRQm0,coefRQm[1])
betaRQm1<- append(betaRQm1,coefRQm[2])
betaRQm2<- append(betaRQm2,coefRQm[3])
# MSE for RQR(0.50)
yhat_RQm<- coefRQm[1] + as.matrix(df[,2:3])%*%coefRQm[2:3]
resid.RQm <- df[,1] - yhat_RQm
MSE.RQm <- mean(resid.RQm^2)
mseRQm<- append(mseRQm,MSE.RQm)

# RIDGE QUANTILE REGRESSION METHOD (0.75)
X<- as.matrix(df[,2:3])
y_r<- df$y
grid.r<- 10^seq(-3, 5, length = 100)
model.RQu<- cv.hqreg(X, y_r, tau=0.75, method="quantile", alpha=0, lambda=grid.r)                    # Estimate RQR (0.75) model
coefRQu<- coef(model.RQu, lambda="lambda.min")
# Store Regression Coefficients
betaRQu0<- append(betaRQu0,coefRQu[1])
betaRQu1<- append(betaRQu1,coefRQu[2])
betaRQu2<- append(betaRQu2,coefRQu[3])
# MSE for RQR(0.75)
yhat_RQu<- coefRQu[1] + as.matrix(df[,2:3])%*%coefRQu[2:3]
resid.RQu <- df[,1] - yhat_RQu
MSE.RQu <- mean(resid.RQu^2)
mseRQu<- append(mseRQu,MSE.RQu)
```

```
# LASSO QUANTILE REGRESSION METHOD (0.25)
 x_l<- as.matrix(df[,2:3])
 y_l<- df$y
 grid.l<- 10^seq(-3, 5, length = 100)
 model.LQl<- cv.hqreg(x_l, y_l, tau=0.25, method="quantile", alpha=1,
lambda=grid.l)                  # Estimate LQR (0.25) model
 coefLQl<- coef(model.LQl, lambda="lambda.min")
 # Store Regression Coefficients
 betaLQl0<- append(betaLQl0,coefLQl[1])
 betaLQl1<- append(betaLQl1,coefLQl[2])
 betaLQl2<- append(betaLQl2,coefLQl[3])
 # MSE for LQR(0.25)
 yhat_LQl<- coefLQl[1] + as.matrix(df[,2:3])%*%coefLQl[2:3]
 resid.LQl <- df[,1] - yhat_LQl
 MSE.LQl <- mean(resid.LQl^2)
 mseLQl<- append(mseLQl,MSE.LQl)

# LASSO QUANTILE REGRESSION METHOD (0.50)
 x_l<- as.matrix(df[,2:3])
 y_l<- df$y
 grid.l<- 10^seq(-3, 5, length = 100)
 model.LQm<- cv.hqreg(x_l, y_l, tau=0.5, method="quantile", alpha=1,
lambda=grid.l)                  # Estimate LQR (0.50) model
 coefLQm<- coef(model.LQm, lambda="lambda.min")
 # Store Regression Coefficients
 betaLQm0<- append(betaLQm0,coefLQm[1])
 betaLQm1<- append(betaLQm1,coefLQm[2])
 betaLQm2<- append(betaLQm2,coefLQm[3])
 # MSE for LQR(0.50)
 yhat_LQm<- coefLQm[1] + as.matrix(df[,2:3])%*%coefLQm[2:3]
 resid.LQm <- df[,1] - yhat_LQm
 MSE.LQm <- mean(resid.LQm^2)
 mseLQm<- append(mseLQm,MSE.LQm)

# LASSO QUANTILE REGRESSION METHOD (0.75)
 x_l<- as.matrix(df[,2:3])
 y_l<- df$y
 grid.l<- 10^seq(-3, 5, length = 100)
 model.LQu<- cv.hqreg(x_l, y_l, tau=0.75, method="quantile", alpha=1,
lambda=grid.l)                  # Estimate LQR (0.75) model
 coefLQu<- coef(model.LQu, lambda="lambda.min")
 # Store Regression Coefficients
 betaLQu0<- append(betaLQu0,coefLQu[1])
 betaLQu1<- append(betaLQu1,coefLQu[2])
 betaLQu2<- append(betaLQu2,coefLQu[3])
```

```
# MSE for LQR(0.75)
yhat_LQu<- coefLQu[1] + as.matrix(df[,2:3])%*%coefLQu[2:3]
resid.LQu <- df[,1] - yhat_LQu
MSE.LQu <- mean(resid.LQu^2)
mseLQu<- append(mseLQu,MSE.LQu)

# TWO STAGE RIDGE QUANTILE REGRESSION METHOD (0.25)
model.ts<- lm(y~x1+x2, data=df)
m<- length(residuals(model.ts))
cor(residuals(model.ts)[-1], residuals(model.ts)[-m])
Phi<- cor(residuals(model.ts)[-1], residuals(model.ts)[-m])
X<- model.matrix(model.ts)
Sigma<- diag(m)
Sigma<- (Phi)^abs(row(Sigma)-col(Sigma))
S<- chol(Sigma)
S_inv<- solve(t(S))
SX_mat<- S_inv%*%X
SX<- as.matrix(SX_mat[,2:3])
SY<- S_inv%*%(df$y)
df2<- cbind.data.frame(SY, SX)
grid.r<- 10^seq(-3, 5, length = 100)
model.TRQl<- cv.hqreg(SX, SY, tau=0.25, method="quantile", alpha=0,
lambda=grid.r)                    # Estimate TRQR (0.25) model
coefTRQl<- coef(model.TRQl, lambda="lambda.min")
# Store Regression Coefficients
betaTRQl0<- append(betaTRQl0,coefTRQl[1])
betaTRQl1<- append(betaTRQl1,coefTRQl[2])
betaTRQl2<- append(betaTRQl2,coefTRQl[3])
# MSE for TRQR(0.25)
yhat_TRQl<- coefTRQl[1] + as.matrix(df2[,2:3])%*%coefTRQl[2:3]
resid.TRQl <- df2[,1] - yhat_TRQl
MSE.TRQl <- mean(resid.TRQl^2)
mseTRQl<- append(mseTRQl,MSE.TRQl)

# TWO STAGE RIDGE QUANTILE REGRESSION METHOD (0.50)
model.ts<-lm(y~x1+x2, data=df)
m<- length(residuals(model.ts))
cor(residuals(model.ts)[-1], residuals(model.ts)[-m])
Phi<- cor(residuals(model.ts)[-1], residuals(model.ts)[-m])
X<- model.matrix(model.ts)
Sigma<- diag(m)
Sigma<- (Phi)^abs(row(Sigma)-col(Sigma))
S<- chol(Sigma)
S_inv<- solve(t(S))
SX_mat<- S_inv%*%X
SX<- as.matrix(SX_mat[,2:3])
```

```
SY<- S_inv%*%(df$y)
df2<- cbind.data.frame(SY, SX)
grid.r<- 10^seq(-3, 5, length = 100)
model.TRQm<- cv.hqreg(SX, SY, tau=0.5, method="quantile", alpha=0,
lambda=grid.r)                    # Estimate TRQR (0.50) model
coefTRQm<- coef(model.TRQm, lambda="lambda.min")
# Store Regression Coefficients
betaTRQm0<- append(betaTRQm0,coefTRQm[1])
betaTRQm1<- append(betaTRQm1,coefTRQm[2])
betaTRQm2<- append(betaTRQm2,coefTRQm[3])
# MSE for TRQR(0.50)
yhat_TRQm<- coefTRQm[1] + as.matrix(df2[,2:3])%*%coefTRQm[2:3]
resid.TRQm <- df2[,1] - yhat_TRQm
MSE.TRQm <- mean(resid.TRQm^2)
mseTRQm<- append(mseTRQm,MSE.TRQm)

# TWO STAGE RIDGE QUANTILE REGRESSION METHOD (0.75)
model.ts<-lm(y~x1+x2, data=df)
m<- length(residuals(model.ts))
cor(residuals(model.ts)[-1], residuals(model.ts)[-m])
Phi<- cor(residuals(model.ts)[-1], residuals(model.ts)[-m])
X<- model.matrix(model.ts)
Sigma<- diag(m)
Sigma<- (Phi)^abs(row(Sigma)-col(Sigma))
S<- chol(Sigma)
S_inv<- solve(t(S))
SX_mat<- S_inv%*%X
SX<- as.matrix(SX_mat[,2:3])
SY<- S_inv%*%(df$y)
df2<- cbind.data.frame(SY, SX)
grid.r<- 10^seq(-3, 5, length = 100)
model.TRQu<- cv.hqreg(SX, SY, tau=0.75, method="quantile", alpha=0,
lambda=grid.r)                    # Estimate TRQR (0.75) model
coefTRQu<- coef(model.TRQu, lambda="lambda.min")
# Store Regression Coefficients
betaTRQu0<- append(betaTRQu0,coefTRQu[1])
betaTRQu1<- append(betaTRQu1,coefTRQu[2])
betaTRQu2<- append(betaTRQu2,coefTRQu[3])
# MSE for TRQR(0.75)
yhat_TRQu<- coefTRQu[1] + as.matrix(df2[,2:3])%*%coefTRQu[2:3]
resid.TRQu <- df2[,1] - yhat_TRQu
MSE.TRQu <- mean(resid.TRQu^2)
mseTRQu<- append(mseTRQu,MSE.TRQu)

# TWO STAGE LASSO QUANTILE REGRESSION METHOD (0.25)
model.ts<- lm(y~x1+x2, data=df)
```

```
m<- length(residuals(model.ts))
cor(residuals(model.ts)[-1], residuals(model.ts)[-m])
Phi<- cor(residuals(model.ts)[-1], residuals(model.ts)[-m])
X<- model.matrix(model.ts)
Sigma<- diag(m)
Sigma<- (Phi)^abs(row(Sigma)-col(Sigma))
S<- chol(Sigma)
S_inv<- solve(t(S))
SX_mat<- S_inv%*%X
SX<- as.matrix(SX_mat[,2:3])
SY<- S_inv%*%(df$y)
df2<- cbind.data.frame(SY, SX)
grid.l<- 10^seq(-3, 5, length = 100)
model.TLQl<- cv.hqreg(SX, SY, tau=0.25, method="quantile", alpha=1,
lambda=grid.l)                    # Estimate TLQR (0.25) model
coefTLQl<- coef(model.TLQl, lambda="lambda.min")
# Store Regression Coefficients
betaTLQl0<- append(betaTLQl0,coefTLQl[1])
betaTLQl1<- append(betaTLQl1,coefTLQl[2])
betaTLQl2<- append(betaTLQl2,coefTLQl[3])
# MSE for TLQR(0.25)
yhat_TLQl<- coefTLQl[1] + as.matrix(df2[,2:3])%*%coefTLQl[2:3]
resid.TLQl <- df2[,1] - yhat_TLQl
MSE.TLQl <- mean(resid.TLQl^2)
mseTLQl<- append(mseTLQl,MSE.TLQl)

# TWO STAGE LASSO QUANTILE REGRESSION METHOD (0.50)
model.ts<- lm(y~x1+x2, data=df)
m<- length(residuals(model.ts))
cor(residuals(model.ts)[-1], residuals(model.ts)[-m])
Phi<- cor(residuals(model.ts)[-1], residuals(model.ts)[-m])
X<- model.matrix(model.ts)
Sigma<- diag(m)
Sigma<- (Phi)^abs(row(Sigma)-col(Sigma))
S<- chol(Sigma)
S_inv<- solve(t(S))
SX_mat<- S_inv%*%X
SX<- as.matrix(SX_mat[,2:3])
SY<- S_inv%*%(df$y)
df2<- cbind.data.frame(SY, SX)
grid.l<- 10^seq(-3, 5, length = 100)
model.TLQm<- cv.hqreg(SX, SY, tau=0.5, method="quantile", alpha=1,
lambda=grid.l)                    # Estimate TLQR (0.50) model
coefTLQm<- coef(model.TLQm, lambda="lambda.min")
# Store Regression Coefficients
betaTLQm0<- append(betaTLQm0,coefTLQm[1])
```

```
  betaTLQm1<- append(betaTLQm1,coefTLQm[2])
  betaTLQm2<- append(betaTLQm2,coefTLQm[3])
  # MSE for TLQR(0.50)
  yhat_TLQm<-coefTLQm[1] + as.matrix(df2[,2:3])%*%coefTLQm[2:3]
  resid.TLQm <- df2[,1] - yhat_TLQm
  MSE.TLQm <- mean(resid.TLQm^2)
  mseTLQm<- append(mseTLQm,MSE.TLQm)

# TWO STAGE LASSO QUANTILE REGRESSION METHOD (0.75)
  model.ts<- lm(y~x1+x2, data=df)
  m<- length(residuals(model.ts))
  cor(residuals(model.ts)[-1], residuals(model.ts)[-m])
  Phi<- cor(residuals(model.ts)[-1], residuals(model.ts)[-m])
  X<- model.matrix(model.ts)
  Sigma<- diag(m)
  Sigma<- (Phi)^abs(row(Sigma)-col(Sigma))
  S<- chol(Sigma)
  S_inv<- solve(t(S))
  SX_mat<- S_inv%*%X
  SX<- as.matrix(SX_mat[,2:3])
  SY<- S_inv%*%(df$y)
  df2<- cbind.data.frame(SY, SX)
  grid.l<- 10^seq(-3, 5, length = 100)
  model.TLQu<- cv.hqreg(SX, SY, tau=0.75, method="quantile", alpha=1,
 lambda=grid.l)                          # Estimate TLQR (0.75) model
  coefTLQu<- coef(model.TLQu, lambda="lambda.min")
  # Store Regression Coefficients
  betaTLQu0<- append(betaTLQu0,coefTLQu[1])
  betaTLQu1<- append(betaTLQu1,coefTLQu[2])
  betaTLQu2<- append(betaTLQu2,coefTLQu[3])
  # MSE for TLQR(0.75)
  yhat_TLQu<- coefTLQu[1] + as.matrix(df2[,2:3])%*%coefTLQu[2:3]
  resid.TLQu <- df2[,1] - yhat_TLQu
  MSE.TLQu <- mean(resid.TLQu^2)
  mseTLQu<- append(mseTLQu,MSE.TLQu)
}

# Ordinary least squares output
B0<- mean(beta0)
B1<- mean(beta1)
B2<- mean(beta2)
mse<- mean(mseols)
# Ridge regression output
Br0<- mean(betar0)
Br1<- mean(betar1)
Br2<- mean(betar2)
```

```
khk<- mean(hk)
klw<- mean(lw)
mse_r<- mean(mser)
# Lasso regression output
Bl0<- mean(betal0)
Bl1<- mean(betal1)
Bl2<- mean(betal2)
mse_lr<- mean(mselr)
# Two Stage Ridge regression output
Btr0<- mean(betatr0)
Btr1<- mean(betatr1)
Btr2<- mean(betatr2)
khk<- mean(hk)
klw<- mean(lw)
mse_tr<- mean(msetr)
# Two Stage Lasso regression output
Btlasso0<- mean(betatlasso0)
Btlasso1<- mean(betatlasso1)
Btlasso2<- mean(betatlasso2)
mse_tlasso<- mean(msetlasso)
# QR (0.25) output
BQl0<- mean(betaQl0)
BQl1<- mean(betaQl1)
BQl2<- mean(betaQl2)
mse_Ql<- mean(mseQl)
# QR (0.50) output
BQm0<- mean(betaQm0)
BQm1<- mean(betaQm1)
BQm2<- mean(betaQm2)
mse_Qm<- mean(mseQm)
# QR (0.75) output
BQu0<- mean(betaQu0)
BQu1<- mean(betaQu1)
BQu2<- mean(betaQu2)
mse_Qu<- mean(mseQu)
# RQR (0.25) output
BRQl0<- mean(betaRQl0)
BRQl1<- mean(betaRQl1)
BRQl2<- mean(betaRQl2)
mse_RQl<- mean(mseRQl)
# RQR (0.50) output
BRQm0<- mean(betaRQm0)
BRQm1<- mean(betaRQm1)
BRQm2<- mean(betaRQm2)
mse_RQm<- mean(mseRQm)
# RQR (0.75) output
```

```
BRQu0<- mean(betaRQu0)
BRQu1<- mean(betaRQu1)
BRQu2<- mean(betaRQu2)
mse_RQu<- mean(mseRQu)
# LQR (0.25) output
BLQl0<- mean(betaLQl0)
BLQl1<- mean(betaLQl1)
BLQl2<- mean(betaLQl2)
mse_LQl<- mean(mseLQl)
# LQR (0.50) output
BLQm0<- mean(betaLQm0)
BLQm1<- mean(betaLQm1)
BLQm2<- mean(betaLQm2)
mse_LQm<- mean(mseLQm)
# LQR (0.75) output
BLQu0<- mean(betaLQu0)
BLQu1<- mean(betaLQu1)
BLQu2<- mean(betaLQu2)
mse_LQu<- mean(mseLQu)
# Two Stage RQR (0.25) output
BTRQl0<- mean(betaTRQl0)
BTRQl1<- mean(betaTRQl1)
BTRQl2<- mean(betaTRQl2)
mse_TRQl<- mean(mseTRQl)
# Two Stage RQR (0.50) output
BTRQm0<- mean(betaTRQm0)
BTRQm1<- mean(betaTRQm1)
BTRQm2<- mean(betaTRQm2)
mse_TRQm<- mean(mseTRQm)
# Two Stage RQR (0.75) output
BTRQu0<- mean(betaTRQu0)
BTRQu1<- mean(betaTRQu1)
BTRQu2<- mean(betaTRQu2)
mse_TRQu<- mean(mseTRQu)
# Two Stage LQR (0.25) output
BTLQl0<- mean(betaTLQl0)
BTLQl1<- mean(betaTLQl1)
BTLQl2<- mean(betaTLQl2)
mse_TLQl<- mean(mseTLQl)
# Two Stage LQR (0.50) output
BTLQm0<- mean(betaTLQm0)
BTLQm1<- mean(betaTLQm1)
BTLQm2<- mean(betaTLQm2)
mse_TLQm<- mean(mseTLQm)
# Two Stage LQR (0.75) output
BTLQu0<- mean(betaTLQu0)
```

```
BTLQu1<- mean(betaTLQu1)
BTLQu2<- mean(betaTLQu2)
mse_TLQu<- mean(mseTLQu)

cat("::::::::::::: OLS
Outputs::::::::::::::::
","\n","intercept=","\t",B0,"\n","B1=","\t",B1,"\n","B2=","\t",B2,"\n","MSE=",
"\t",mse,"\n","
:::::::::::::::: Ridge Regression
Outputs::::::::::::::::
","\n","intercept=","\t",Br0,"\n","B1=","\t",Br1,"\n","B2=","\t",Br2,"\n","MSE
_R=","\t",mse_r,"\n","KHK=","\t",khk,"\n","KLW=","\t",klw,"\n","
:::::::::::::::: Lasso Regression
Outputs::::::::::::::::
","\n","intercept=","\t",Bl0,"\n","B1=","\t",Bl1,"\n","B2=","\t",Bl2,"\n","MSE
_L=","\t",mse_lr,"\n","
:::::::::::::::: Two Stage Ridge Regression
Outputs::::::::::::::::
","\n","intercept=","\t",Btr0,"\n","B1=","\t",Btr1,"\n","B2=","\t",Btr2,"\n","MS
E_TR=","\t",mse_tr,"\n","KHK=","\t",khk,"\n","KLW=","\t",klw,"\n","
:::::::::::::::: Two Stage Lasso Regression
Outputs::::::::::::::::
","\n","intercept=","\t",Btlasso0,"\n","B1=","\t",Btlasso1,"\n","B2=","\t",Btlass
o2,"\n","MSE_TLASSO=","\t",mse_tlasso,"\n","
:::::::::::::::: QR (0.25)
Outputs::::::::::::::::
","\n","intercept=","\t",BQl0,"\n","B1=","\t",BQl1,"\n","B2=","\t",BQl2,"\n","
MSE_QL=","\t",mse_Ql,"\n","
:::::::::::::::: QR (0.50)
Outputs::::::::::::::::
","\n","intercept=","\t",BQm0,"\n","B1=","\t",BQm1,"\n","B2=","\t",BQm2,"\
n","MSE_QM=","\t",mse_Qm,"\n","
:::::::::::::::: QR (0.75)
Outputs::::::::::::::::
","\n","intercept=","\t",BQu0,"\n","B1=","\t",BQu1,"\n","B2=","\t",BQu2,"\n",
"MSE_QU=","\t",mse_Qu,"\n","
:::::::::::::::: RQR (0.25)
Outputs::::::::::::::::
","\n","intercept=","\t",BRQl0,"\n","B1=","\t",BRQl1,"\n","B2=","\t",BRQl2,"
\n","MSE_RQL=","\t",mse_RQl,"\n","
:::::::::::::::: RQR (0.50)
Outputs::::::::::::::::
","\n","intercept=","\t",BRQm0,"\n","B1=","\t",BRQm1,"\n","B2=","\t",BRQ
m2,"\n","MSE_RQM=","\t",mse_RQm,"\n","
:::::::::::::::: RQR (0.75)
```

```
Outputs::::::::::::::::
","\n","intercept=","\t",BRQu0,"\n","B1=","\t",BRQu1,"\n","B2=","\t",BRQu2
,"\n","MSE_RQU=","\t",mse_RQu,"\n","
:::::::::::::::: LQR (0.25)
Outputs::::::::::::::::
","\n","intercept=","\t",BLQl0,"\n","B1=","\t",BLQl1,"\n","B2=","\t",BLQl2,"\
n","MSE_LQL=","\t",mse_LQl,"\n","
:::::::::::::::: LQR (0.50)
Outputs::::::::::::::::
","\n","intercept=","\t",BLQm0,"\n","B1=","\t",BLQm1,"\n","B2=","\t",BLQm
2,"\n","MSE_LQM=","\t",mse_LQm,"\n","
:::::::::::::::: LQR (0.75)
Outputs::::::::::::::::
","\n","intercept=","\t",BLQu0,"\n","B1=","\t",BLQu1,"\n","B2=","\t",BLQu2,
"\n","MSE_LQU=","\t",mse_LQu,"\n","
:::::::::::::::: Two Stage RQR (0.25)
Outputs::::::::::::::::
","\n","intercept=","\t",BTRQl0,"\n","B1=","\t",BTRQl1,"\n","B2=","\t",BTR
Ql2,"\n","MSE_TRQL=","\t",mse_TRQl,"\n","
:::::::::::::::: Two Stage RQR (0.50)
Outputs::::::::::::::::
","\n","intercept=","\t",BTRQm0,"\n","B1=","\t",BTRQm1,"\n","B2=","\t",BT
RQm2,"\n","MSE_TRQM=","\t",mse_TRQm,"\n","
:::::::::::::::: Two Stage RQR (0.75)
Outputs::::::::::::::::
","\n","intercept=","\t",BTRQu0,"\n","B1=","\t",BTRQu1,"\n","B2=","\t",BTR
Qu2,"\n","MSE_TRQU=","\t",mse_TRQu,"\n","
:::::::::::::::: Two Stage LQR (0.25)
Outputs::::::::::::::::
","\n","intercept=","\t",BTLQl0,"\n","B1=","\t",BTLQl1,"\n","B2=","\t",BTL
Ql2,"\n","MSE_TLQL=","\t",mse_TLQl,"\n","
:::::::::::::::: Two Stage LQR (0.50)
Outputs::::::::::::::::
","\n","intercept=","\t",BTLQm0,"\n","B1=","\t",BTLQm1,"\n","B2=","\t",BT
LQm2,"\n","MSE_TLQM=","\t",mse_TLQm,"\n","
:::::::::::::::: Two Stage LQR (0.75)
Outputs::::::::::::::::
","\n","intercept=","\t",BTLQu0,"\n","B1=","\t",BTLQu1,"\n","B2=","\t",BTL
Qu2,"\n","MSE_TLQU=","\t",mse_TLQu)


# All Estimators MSE
mse<-
cbind(mse,mse_r,mse_lr,mse_tr,mse_tlasso,mse_Ql,mse_Qm,mse_Qu,mse_R
Ql,mse_RQm,mse_RQu,mse_LQl,mse_LQm,mse_LQu,mse_TRQl,mse_TRQ
m,mse_TRQu,mse_TLQl,mse_TLQm,mse_TLQu)
print(mse)
```

APPENDIX B

SIMULATION REGRESSION COEFFICIENTS

Table B1: Simulation results of regression coefficients when γ^2=0.7, ρ=0.1 and n=25

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	-0.018	-0.024	-0.022	-0.025	-0.024
β_1	0.898	0.880	0.872	0.879	0.869
β_2	1.017	0.987	0.999	0.990	1.001

Table B2: Estimated regression coefficients when γ^2=0.7, ρ=0.1 and n=25

τ	Coefficient	QR	RQR	LQR	TRQR	TLQR
0.25	β_0	-0.744	-0.823	-0.802	-0.789	-0.771
	β_1	0.592	0.538	0.524	0.568	0.549
	β_2	1.044	0.882	0.942	0.888	0.963
0.5	β_0	-0.031	-0.082	-0.069	-0.088	-0.068
	β_1	0.720	0.674	0.630	0.682	0.650
	β_2	1.127	0.990	1.059	0.976	1.052
0.75	β_0	0.637	0.573	0.601	0.573	0.590
	β_1	1.039	0.886	0.915	0.886	0.911
	β_2	0.974	0.877	0.911	0.879	0.922

Table B3: Simulation results of regression coefficients when γ^2=0.7, ρ=0.1 and n=50

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	-0.059	-0.060	-0.061	-0.062	-0.063
β_1	0.916	0.910	0.905	0.909	0.903
β_2	1.105	1.087	1.089	1.087	1.090

Table B4: Estimated regression coefficients when γ^2=0.7, ρ=0.1 and n=50

τ	Coefficient	QR	RQR	LQR	TRQR	TLQR
0.25	β_0	-0.692	-0.723	-0.710	-0.713	-0.710
	β_1	0.868	0.862	0.849	0.854	0.848
	β_2	1.095	0.992	1.027	1.002	1.028
0.5	β_0	-0.032	-0.041	-0.044	-0.048	-0.047
	β_1	0.903	0.878	0.870	0.874	0.865
	β_2	1.138	1.064	1.090	1.067	1.100
0.75	β_0	0.588	0.581	0.587	0.580	0.585
	β_1	0.904	0.875	0.872	0.881	0.878
	β_2	1.219	1.144	1.173	1.134	1.160

Table B5: Simulation results of regression coefficients when γ^2=0.7, ρ=0.1 and n=200

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	0.002	0.002	0.002	0.002	0.002
β_1	1.141	1.136	1.135	1.136	1.135
β_2	0.868	0.867	0.859	0.867	0.859

Table B6: Estimated regression coefficients when γ^2=0.7, ρ=0.1 and n=200

τ	Coefficient	QR	RQR	LQR	TRQR	TLQR
0.25	β_0	-0.711	-0.715	-0.716	-0.714	-0.715
	β_1	1.126	1.096	1.109	1.096	1.110
	β_2	0.897	0.896	0.894	0.896	0.893
0.5	β_0	-0.012	-0.011	-0.010	-0.011	-0.010
	β_1	1.127	1.092	1.109	1.092	1.110
	β_2	0.858	0.849	0.843	0.851	0.843
0.75	β_0	0.696	0.693	0.694	0.693	0.693
	β_1	1.190	1.160	1.171	1.160	1.170
	β_2	0.810	0.814	0.808	0.815	0.808

Table B7: Simulation results of regression coefficients when γ^2=0.9, ρ=0.1 and n=25

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	-0.016	-0.022	-0.019	-0.023	-0.021
β_1	0.862	0.864	0.840	0.861	0.840
β_2	1.064	1.024	1.048	1.028	1.050

Table B8: Estimated regression coefficients when γ^2=0.9, ρ=0.1 and n=25

τ	Coefficient	QR	RQR	LQR	TRQR	TLQR
0.25	β_0	-0.743	-0.794	-0.789	-0.779	-0.767
	β_1	0.454	0.497	0.391	0.545	0.450
	β_2	1.222	1.003	1.141	0.995	1.122
0.5	β_0	-0.024	-0.071	-0.063	-0.077	-0.065
	β_1	0.579	0.643	0.550	0.651	0.570
	β_2	1.297	1.088	1.206	1.070	1.182
0.75	β_0	0.641	0.584	0.622	0.581	0.606
	β_1	1.062	0.912	0.967	0.892	0.958
	β_2	0.951	0.885	0.871	0.887	0.892

Table B9: Simulation results of regression coefficients when γ^2=0.9, ρ=0.1 and n=50

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	0.116	0.116	0.116	0.119	0.119
β_1	0.965	0.956	0.956	0.957	0.958
β_2	0.954	0.944	0.943	0.944	0.943

Table B10: Estimated regression coefficients when γ^2=0.9, ρ=0.1 and n=50

τ	Coefficient	QR	RQR	LQR	TRQR	TLQR
0.25	β_0	-0.567	-0.573	-0.574	-0.567	-0.571
	β_1	1.065	0.994	1.026	0.999	1.026
	β_2	0.873	0.863	0.847	0.848	0.850
0.5	β_0	0.089	0.087	0.097	0.096	0.100
	β_1	1.065	0.987	1.029	0.979	1.018
	β_2	0.917	0.888	0.871	0.890	0.877
0.75	β_0	0.787	0.807	0.809	0.809	0.808
	β_1	0.993	0.929	0.960	0.928	0.953
	β_2	0.921	0.865	0.861	0.870	0.868

Table B11: Simulation results of regression coefficients when γ^2=0.9, ρ=0.1 and n=200

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	0.003	0.002	0.003	0.002	0.003
β_1	1.236	1.223	1.231	1.223	1.231
β_2	0.757	0.766	0.747	0.766	0.747

Table B12: Estimated regression coefficients when γ^2=0.9, ρ=0.1 and n=200

τ	Coefficient	QR	RQR	LQR	TRQR	TLQR
0.25	β_0	-0.711	-0.711	-0.715	-0.712	-0.714
	β_1	1.191	1.133	1.175	1.131	1.176
	β_2	0.813	0.833	0.810	0.833	0.808
0.5	β_0	-0.011	-0.014	-0.011	-0.015	-0.011
	β_1	1.222	1.153	1.204	1.151	1.205
	β_2	0.754	0.779	0.741	0.783	0.742
0.75	β_0	0.695	0.693	0.694	0.693	0.694
	β_1	1.329	1.272	1.305	1.267	1.303
	β_2	0.672	0.703	0.674	0.707	0.675

Table B13: Simulation results of regression coefficients when γ^2=0.9, ρ=0.1 and n=500

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	-0.005	-0.005	-0.006	-0.005	-0.006
β_1	1.000	0.999	0.994	1.000	0.994
β_2	1.016	1.015	1.010	1.015	1.010

Table B14: Estimated regression coefficients when γ^2=0.9, ρ=0.1 and n=500

τ	Coefficient	QR	RQR	LQR	TRQR	TLQR
0.25	β_0	-0.705	-0.709	-0.711	-0.708	-0.711
	β_1	1.022	1.007	1.010	1.007	1.010
	β_2	0.994	0.997	1.003	0.998	1.004
0.5	β_0	-0.027	-0.023	-0.022	-0.022	-0.021
	β_1	0.988	0.980	0.989	0.981	0.989
	β_2	0.979	0.973	0.977	0.972	0.976
0.75	β_0	0.661	0.673	0.674	0.673	0.675
	β_1	1.176	1.129	1.139	1.124	1.136
	β_2	0.834	0.869	0.863	0.874	0.866

Table B15: Simulation results of regression coefficients when γ^2=0.99, ρ=0.1 and n=25

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	-0.014	-0.020	-0.018	-0.021	-0.019
β_1	0.655	0.763	0.725	0.756	0.723
β_2	1.279	1.142	1.177	1.153	1.185

Table B16: Estimated regression coefficients when γ^2=0.99, ρ=0.1 and n=25

τ	Coefficient	QR	RQR	LQR	TRQR	TLQR
0.25	β_0	-0.736	-0.772	-0.806	-0.768	-0.767
	β_1	-0.367	0.183	0.165	0.326	0.044
	β_2	2.067	1.395	1.401	1.270	1.563
0.5	β_0	-0.021	-0.055	-0.051	-0.056	-0.053
	β_1	-0.165	0.422	0.204	0.420	0.383
	β_2	2.057	1.377	1.596	1.358	1.421
0.75	β_0	0.641	0.584	0.602	0.586	0.601
	β_1	1.159	0.866	0.930	0.859	1.006
	β_2	0.846	0.957	0.926	0.961	0.855

Table B17: Simulation results of regression coefficients when γ^2=0.99, ρ=0.1 and n=50

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	0.116	0.117	0.117	0.120	0.120
β_1	0.975	0.958	0.978	0.960	0.970
β_2	0.944	0.944	0.921	0.944	0.930

Table B18: Estimated regression coefficients when γ^2=0.99, ρ=0.1 and n=50

τ	Coefficient	QR	RQR	LQR	TRQR	TLQR
0.25	β_0	-0.566	-0.572	-0.571	-0.565	-0.567
	β_1	1.277	1.072	1.178	1.066	1.153
	β_2	0.651	0.785	0.687	0.793	0.710
0.5	β_0	0.088	0.084	0.090	0.093	0.096
	β_1	1.213	1.041	1.077	1.012	1.074
	β_2	0.766	0.856	0.822	0.875	0.827
0.75	β_0	0.790	0.805	0.806	0.802	0.801
	β_1	1.058	0.949	1.042	0.975	1.031
	β_2	0.858	0.873	0.787	0.859	0.800

Table B19: Estimated regression coefficients when γ^2=0.99, ρ=0.1 and n=200

τ	Coefficient	QR	RQR	LQR	TRQR	TLQR
0.25	β_0	-0.668	-0.664	-0.671	-0.664	-0.670
	β_1	1.451	1.161	1.349	1.169	1.346
	β_2	0.581	0.828	0.653	0.822	0.659
0.5	β_0	0.080	0.083	0.077	0.081	0.076
	β_1	0.971	0.967	0.931	0.965	0.919
	β_2	1.025	0.992	1.048	0.994	1.059
0.75	β_0	0.691	0.709	0.709	0.711	0.710
	β_1	0.546	0.703	0.556	0.700	0.557
	β_2	1.354	1.167	1.326	1.171	1.323

Table B20: Simulation results of regression coefficients when γ^2=0.99, ρ=0.1 and n=500

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	-0.005	-0.005	-0.005	-0.005	-0.005
β_1	0.985	0.989	0.980	0.989	0.981
β_2	1.039	1.033	1.030	1.033	1.030

Table B21: Estimated regression coefficients when γ^2=0.99, ρ=0.1 and n=500

τ	Coefficient	QR	RQR	LQR	TRQR	TLQR
0.25	β_0	-0.706	-0.708	-0.713	-0.707	-0.712
	β_1	1.052	1.018	1.019	1.016	1.018
	β_2	0.963	0.982	0.996	0.984	0.997
0.5	β_0	-0.027	-0.023	-0.021	-0.022	-0.021
	β_1	1.006	0.992	1.020	0.996	1.021
	β_2	0.973	0.965	0.958	0.963	0.956
0.75	β_0	0.662	0.672	0.675	0.672	0.675
	β_1	1.546	1.291	1.434	1.278	1.425
	β_2	0.474	0.719	0.580	0.731	0.589

Table B22: Simulation results of regression coefficients when γ^2=0.7, ρ=0.9 and n=25

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	0.200	0.193	0.180	0.201	0.190
β_1	0.719	0.714	0.643	0.709	0.644
β_2	1.207	1.163	1.127	1.163	1.131

Table B23: Estimated regression coefficients when γ^2=0.7, ρ=0.9 and n=25

τ	Coefficient	QR	RQR	LQR	TRQR	TLQR
0.25	β_0	-1.071	-1.270	-1.311	-1.235	-1.252
	β_1	0.585	0.480	0.409	0.501	0.441
	β_2	1.345	0.867	0.902	0.895	0.956
0.5	β_0	0.351	0.327	0.341	0.316	0.314
	β_1	0.865	0.835	0.800	0.804	0.756
	β_2	1.208	1.019	1.103	1.030	1.110
0.75	β_0	1.469	1.533	1.508	1.529	1.498
	β_1	0.800	0.690	0.666	0.669	0.650
	β_2	0.985	0.780	0.839	0.798	0.870

Table B24: Simulation results of regression coefficients when γ^2=0.7, ρ=0.9 and n=50

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	-0.683	-0.684	-0.687	-0.709	-0.711
β_1	0.991	0.977	0.959	0.975	0.954
β_2	0.927	0.919	0.894	0.916	0.896

Table B25: Estimated regression coefficients when γ^2=0.7, ρ=0.9 and n=50

τ	Coefficient	QR	RQR	LQR	TRQR	TLQR
0.25	β_0	-2.274	-2.319	-2.298	-2.332	-2.319
	β_1	1.245	0.977	1.063	0.958	1.042
	β_2	0.710	0.635	0.600	0.654	0.616
0.5	β_0	-0.405	-0.391	-0.402	-0.450	-0.450
	β_1	1.113	0.895	0.942	0.870	0.941
	β_2	0.863	0.776	0.787	0.806	0.799
0.75	β_0	1.043	1.067	1.082	1.045	1.052
	β_1	0.747	0.628	0.636	0.647	0.678
	β_2	1.017	0.777	0.817	0.790	0.809

Table B26: Simulation results of regression coefficients when γ^2=0.7, ρ=0.9 and n=200

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	-0.046	-0.046	-0.046	-0.047	-0.046
β_1	1.217	1.210	1.210	1.210	1.210
β_2	0.622	0.624	0.612	0.624	0.612

Table B27: Estimated regression coefficients when γ^2=0.7, ρ=0.9 and n=200

τ	Coefficient	QR	RQR	LQR	TRQR	TLQR
0.25	β_0	-1.785	-1.781	-1.781	-1.775	-1.776
	β_1	1.272	1.130	1.215	1.133	1.209
	β_2	0.651	0.654	0.608	0.656	0.605
0.5	β_0	0.021	0.042	0.037	0.035	0.029
	β_1	1.108	0.975	1.040	0.980	1.047
	β_2	0.743	0.695	0.671	0.698	0.674
0.75	β_0	1.516	1.518	1.531	1.520	1.533
	β_1	1.238	1.143	1.214	1.136	1.210
	β_2	0.498	0.524	0.470	0.528	0.471

Table B28: Simulation results of regression coefficients when γ^2=0.7, ρ=0.9 and n=500

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	-0.007	-0.007	-0.008	-0.007	-0.008
β_1	1.057	1.056	1.050	1.055	1.050
β_2	1.009	1.008	1.002	1.008	1.002

Table B29: Simulation results of regression coefficients when γ^2=0.9, ρ=0.9 and n=25

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	0.197	0.189	0.178	0.197	0.186
β_1	0.518	0.583	0.516	0.575	0.521
β_2	1.381	1.279	1.246	1.284	1.248

Table B30: Estimated regression coefficients when γ^2=0.9, ρ=0.9 and n=25

τ	Coefficient	QR	RQR	LQR	TRQR	TLQR
0.25	β_0	-1.070	-1.275	-1.250	-1.231	-1.199
	β_1	0.279	0.439	0.367	0.641	0.374
	β_2	1.599	0.965	1.080	0.896	1.148
0.5	β_0	0.348	0.332	0.342	0.331	0.317
	β_1	0.706	0.796	0.726	0.727	0.690
	β_2	1.339	1.067	1.175	1.114	1.168
0.75	β_0	1.471	1.558	1.547	1.537	1.508
	β_1	0.751	0.719	0.734	0.701	0.678
	β_2	1.074	0.846	0.903	0.870	0.918

Table B31: Simulation results of regression coefficients when γ^2=0.9, ρ=0.9 and n=50

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	-0.681	-0.681	-0.684	-0.707	-0.708
β_1	1.019	0.996	0.987	0.994	0.976
β_2	0.916	0.919	0.883	0.916	0.894

Table B32: Estimated regression coefficients when γ^2=0.9, ρ=0.9 and n=50

τ	Coefficient	QR	RQR	LQR	TRQR	TLQR
0.25	β_0	-2.267	-2.311	-2.306	-2.329	-2.318
	β_1	1.430	1.091	1.171	1.066	1.158
	β_2	0.519	0.582	0.517	0.586	0.530
0.5	β_0	-0.407	-0.394	-0.396	-0.452	-0.449
	β_1	1.226	0.961	1.046	0.944	1.021
	β_2	0.798	0.804	0.778	0.812	0.786
0.75	β_0	1.042	1.075	1.076	1.056	1.047
	β_1	0.653	0.639	0.604	0.662	0.630
	β_2	1.181	0.858	0.934	0.871	0.947

Table B33: Simulation results of regression coefficients when γ^2=0.9, ρ=0.9 and n=200

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	-0.056	-0.056	-0.057	-0.057	-0.058
β_1	1.237	1.223	1.219	1.223	1.218
β_2	0.713	0.721	0.714	0.720	0.713

Table B34: Estimated regression coefficients when γ^2=0.9, ρ=0.9 and n=200

τ	Coefficient	QR	RQR	LQR	TRQR	TLQR
0.25	β_0	-1.762	-1.762	-1.760	-1.763	-1.761
	β_1	1.204	1.058	1.176	1.055	1.173
	β_2	0.727	0.764	0.700	0.759	0.703
0.5	β_0	-0.144	-0.142	-0.135	-0.137	-0.135
	β_1	1.066	0.932	1.010	0.942	1.019
	β_2	0.793	0.775	0.745	0.769	0.740
0.75	β_0	1.828	1.791	1.795	1.779	1.785
	β_1	1.207	1.058	1.172	1.061	1.170
	β_2	0.765	0.795	0.734	0.796	0.733

Table B35: Simulation results of regression coefficients when γ^2=0.9, ρ=0.9 and n=500

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	-0.006	-0.006	-0.007	-0.006	-0.007
β_1	1.085	1.083	1.078	1.082	1.077
β_2	1.003	1.002	0.996	1.003	0.996

Table B36: Estimated regression coefficients when γ^2=0.9, ρ=0.9 and n=500

τ	Coefficient	QR	RQR	LQR	TRQR	TLQR
0.25	β_0	-1.486	-1.495	-1.497	-1.496	-1.497
	β_1	0.914	0.916	0.905	0.913	0.904
	β_2	1.151	1.087	1.135	1.090	1.135
0.5	β_0	0.021	0.027	0.023	0.025	0.022
	β_1	1.044	1.013	1.033	1.014	1.033
	β_2	1.014	0.987	1.003	0.987	1.004
0.75	β_0	1.337	1.349	1.358	1.354	1.362
	β_1	1.209	1.134	1.192	1.135	1.190
	β_2	0.851	0.871	0.848	0.874	0.850

Table B37: Simulation results of regression coefficients when γ^2=0.99, ρ=0.9 and n=25

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	-0.758	-0.757	-0.756	-0.798	-0.797
β_1	0.694	0.934	0.869	0.902	0.863
β_2	1.516	1.237	1.307	1.266	1.306

Table B38: Estimated regression coefficients when γ^2=0.99, ρ=0.9 and n=25

τ	Coefficient	QR	RQR	LQR	TRQR	TLQR
0.25	β_0	-2.003	-1.899	-1.931	-1.942	-1.960
	β_1	-0.311	0.645	0.302	0.592	0.336
	β_2	2.696	1.539	1.900	1.583	1.852
0.5	β_0	-0.779	-0.733	-0.746	-0.741	-0.772
	β_1	1.611	1.240	1.347	1.180	1.286
	β_2	0.624	0.740	0.649	0.783	0.745
0.75	β_0	0.552	0.716	0.694	0.609	0.578
	β_1	1.757	1.042	1.253	1.033	1.226
	β_2	0.279	0.743	0.568	0.780	0.634

Table B39: Simulation results of regression coefficients when γ^2=0.99, ρ=0.9 and n=50

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	-0.680	-0.679	-0.679	-0.704	-0.704
β_1	1.122	1.022	1.090	1.017	1.013
β_2	0.826	0.909	0.817	0.911	0.895

Table B40: Estimated regression coefficients when γ^2=0.99, ρ=0.9 and n=50

τ	Coefficient	QR	RQR	LQR	TRQR	TLQR
0.25	β_0	-2.262	-2.325	-2.309	-2.341	-2.330
	β_1	2.455	1.311	1.558	1.259	1.501
	β_2	-0.517	0.389	0.181	0.444	0.247
0.5	β_0	-0.410	-0.394	-0.395	-0.445	-0.452
	β_1	1.747	1.030	1.182	1.002	1.265
	β_2	0.316	0.805	0.693	0.836	0.621
0.75	β_0	1.045	1.068	1.064	1.049	1.046
	β_1	0.079	0.570	0.386	0.591	0.463
	β_2	1.801	1.019	1.245	1.034	1.195

Table B41: Simulation results of regression coefficients when γ^2=0.99, ρ=0.9 and n=200

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	-0.056	-0.057	-0.058	-0.058	-0.059
β_1	1.813	1.565	1.515	1.567	1.522
β_2	0.143	0.384	0.423	0.381	0.414

Table B42: Estimated regression coefficients when γ^2=0.99, ρ=0.9 and n=200

τ	Coefficient	QR	RQR	LQR	TRQR	TLQR
0.25	β_0	-1.762	-1.766	-1.764	-1.766	-1.764
	β_1	1.751	1.157	1.393	1.157	1.375
	β_2	0.186	0.692	0.484	0.686	0.496
0.5	β_0	-0.145	-0.146	-0.142	-0.143	-0.141
	β_1	1.369	0.989	1.132	1.008	1.151
	β_2	0.492	0.755	0.627	0.744	0.625
0.75	β_0	1.833	1.794	1.798	1.783	1.788
	β_1	1.681	1.182	1.381	1.178	1.379
	β_2	0.304	0.723	0.537	0.724	0.538

Table B43: Simulation results of regression coefficients when γ^2=0.99, ρ=0.9 and n=500

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	-0.006	-0.006	-0.007	-0.006	-0.007
β_1	1.177	1.156	1.154	1.154	1.154
β_2	0.925	0.944	0.934	0.945	0.935

Table B44: Estimated regression coefficients when γ^2=0.99, ρ=0.9 and n=500

τ	Coefficient	QR	RQR	LQR	TRQR	TLQR
0.25	β_0	-1.487	-1.492	-1.496	-1.493	-1.496
	β_1	0.663	0.856	0.701	0.853	0.700
	β_2	1.410	1.172	1.350	1.174	1.351
0.5	β_0	0.019	0.027	0.023	0.025	0.022
	β_1	1.081	1.044	1.079	1.039	1.079
	β_2	0.993	0.985	0.972	0.990	0.974
0.75	β_0	1.340	1.359	1.360	1.364	1.365
	β_1	1.590	1.263	1.522	1.271	1.518
	β_2	0.484	0.771	0.527	0.769	0.533

Table B45: Regression coefficients for four predictors when γ^2=0.7, ρ=0.1 and n=25

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	0.174	0.173	0.177	0.182	0.187
β_1	0.943	0.926	0.927	0.929	0.933
β_2	1.079	1.047	1.063	1.047	1.064
β_3	1.249	1.195	1.227	1.199	1.230
β_4	0.652	0.708	0.643	0.703	0.638

Table B46: Regression coefficients for four predictors when γ^2=0.7, ρ=0.1 and n=50

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	0.112	0.116	0.120	0.119	0.124
β_1	0.890	0.889	0.875	0.892	0.878
β_2	1.063	1.054	1.056	1.051	1.052
β_3	1.133	1.117	1.125	1.120	1.129
β_4	0.950	0.950	0.937	0.948	0.936

Table B47: Regression coefficients for four predictors when γ^2=0.7, ρ=0.1 and n=500

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	-0.012	-0.012	-0.011	-0.012	-0.012
β_1	1.027	1.026	1.020	1.027	1.020
β_2	0.983	0.982	0.977	0.982	0.977
β_3	0.972	0.971	0.965	0.971	0.965
β_4	1.041	1.041	1.035	1.041	1.035

Table B48: Regression coefficients for four predictors when γ^2=0.9, ρ=0.1 and n=25

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	0.168	0.160	0.165	0.169	0.175
β_1	0.919	0.905	0.892	0.909	0.895
β_2	1.152	1.072	1.122	1.073	1.120
β_3	1.439	1.287	1.381	1.296	1.386
β_4	0.409	0.617	0.471	0.607	0.470

Table B49: Regression coefficients for four predictors when γ^2=0.9, ρ=0.1 and n=50

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	0.112	0.114	0.120	0.117	0.124
β_1	0.798	0.824	0.785	0.827	0.789
β_2	1.102	1.077	1.095	1.073	1.088
β_3	1.221	1.183	1.213	1.187	1.219
β_4	0.906	0.923	0.897	0.921	0.894

Table B50: Regression coefficients for four predictors when γ^2=0.9, ρ=0.1 and n=200

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	-0.016	-0.016	-0.016	-0.016	-0.015
β_1	1.181	1.173	1.168	1.173	1.169
β_2	1.085	1.079	1.085	1.080	1.086
β_3	0.789	0.798	0.782	0.798	0.782
β_4	0.985	0.984	0.979	0.984	0.979

Table B51: Regression coefficients for four predictors when γ^2=0.9, ρ=0.1 and n=500

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	-0.012	-0.012	-0.012	-0.012	-0.012
β_1	1.045	1.044	1.037	1.045	1.037
β_2	0.967	0.967	0.962	0.966	0.962
β_3	0.949	0.950	0.942	0.950	0.942
β_4	1.069	1.067	1.063	1.067	1.063

Table B52: Regression coefficients for four predictors when γ^2=0.99, ρ=0.1 and n=25

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	-0.197	-0.191	-0.195	-0.200	-0.203
β_1	0.805	0.951	0.888	0.947	0.863
β_2	1.170	1.013	1.091	1.020	1.078
β_3	0.791	0.905	0.777	0.884	0.800
β_4	1.091	0.956	1.065	0.973	1.083

Table B53: Regression coefficients for four predictors when γ^2=0.99, ρ=0.1 and n=50

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	0.113	0.106	0.115	0.109	0.119
β_1	0.339	0.698	0.508	0.687	0.511
β_2	1.311	1.105	1.158	1.112	1.136
β_3	1.683	1.340	1.539	1.358	1.554
β_4	0.688	0.866	0.785	0.852	0.790

Table B54: Regression coefficients for four predictors when γ^2=0.99, ρ=0.1 and n=200

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	-0.015	-0.015	-0.015	-0.015	-0.015
β_1	1.554	1.381	1.450	1.382	1.452
β_2	1.246	1.163	1.199	1.164	1.201
β_3	0.312	0.544	0.454	0.543	0.453
β_4	0.931	0.949	0.915	0.948	0.913

Table B55: Regression coefficients for four predictors when γ^2=0.99, ρ=0.1 and n=500

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	-0.052	-0.052	-0.051	-0.052	-0.052
β_1	0.819	0.845	0.821	0.843	0.819
β_2	1.051	1.043	1.032	1.046	1.034
β_3	1.284	1.237	1.269	1.235	1.267
β_4	0.810	0.837	0.817	0.839	0.819

Table B56: Regression coefficients for four predictors when γ^2=0.7, ρ=0.9 and n=25

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	0.817	0.822	0.822	0.884	0.884
β_1	1.165	1.116	1.118	1.122	1.133
β_2	0.786	0.791	0.758	0.784	0.753
β_3	0.847	0.841	0.805	0.846	0.813
β_4	0.885	0.890	0.859	0.887	0.861

Table B57: Regression coefficients for four predictors when γ^2=0.7, ρ=0.9 and n=50

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	1.167	1.166	1.164	1.198	1.197
β_1	0.956	0.952	0.933	0.951	0.933
β_2	0.740	0.744	0.738	0.745	0.738
β_3	1.087	1.074	1.064	1.070	1.059
β_4	1.139	1.125	1.118	1.127	1.120

Table B58: Regression coefficients for four predictors when γ^2=0.7, ρ=0.9 and n=200

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	-0.159	-0.159	-0.158	-0.160	-0.160
β_1	1.001	1.000	0.991	0.999	0.990
β_2	1.097	1.093	1.093	1.093	1.094
β_3	0.816	0.818	0.809	0.819	0.811
β_4	1.000	0.998	0.994	0.998	0.994

Table B59: Regression coefficients for four predictors when γ^2=0.7, ρ=0.9 and n=500

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	-0.078	-0.078	-0.077	-0.078	-0.078
β_1	0.956	0.956	0.949	0.956	0.949
β_2	0.975	0.975	0.970	0.974	0.969
β_3	0.838	0.838	0.832	0.839	0.832
β_4	1.031	1.030	1.024	1.030	1.025

Table B60: Regression coefficients for four predictors when γ^2=0.9, ρ=0.9 and n=25

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	0.797	0.805	0.796	0.866	0.857
β_1	1.352	1.197	1.255	1.209	1.269
β_2	0.709	0.772	0.717	0.760	0.701
β_3	0.794	0.830	0.768	0.836	0.765
β_4	0.878	0.897	0.874	0.892	0.893

Table B61: Regression coefficients for four predictors when γ^2=0.9, ρ=0.9 and n=50

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	1.167	1.167	1.166	1.199	1.198
β_1	0.941	0.942	0.905	0.940	0.910
β_2	0.564	0.614	0.620	0.616	0.615
β_3	1.169	1.132	1.117	1.125	1.112
β_4	1.258	1.219	1.213	1.223	1.221

Table B62: Regression coefficients for four predictors when γ^2=0.9, ρ=0.9 and n=200

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	-0.160	-0.161	-0.160	-0.162	-0.162
β_1	1.022	1.021	1.008	1.019	1.006
β_2	1.185	1.172	1.179	1.173	1.180
β_3	0.700	0.713	0.702	0.715	0.704
β_4	1.018	1.014	1.010	1.015	1.011

Table B63: Regression coefficients for four predictors when γ^2=0.99, ρ=0.9 and n=25

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	-0.025	-0.032	-0.026	-0.033	-0.028
β_1	0.908	0.934	0.962	0.933	0.935
β_2	1.820	1.247	1.395	1.273	1.398
β_3	0.493	0.773	0.679	0.751	0.702
β_4	0.654	0.891	0.774	0.898	0.786

Table B64: Regression coefficients for four predictors when γ^2=0.99, ρ=0.9 and n=50

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	0.057	0.051	0.054	0.055	0.057
β_1	1.796	1.338	1.387	1.347	1.391
β_2	1.610	1.301	1.366	1.313	1.366
β_3	0.695	0.864	0.810	0.861	0.828
β_4	0.016	0.606	0.518	0.589	0.496

Table B65: Regression coefficients for four predictors when γ^2=0.99, ρ=0.9 and n=200

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	-0.161	-0.162	-0.162	-0.163	-0.163
β_1	1.116	1.083	1.033	1.078	1.022
β_2	1.623	1.398	1.432	1.399	1.433
β_3	0.093	0.398	0.419	0.402	0.427
β_4	1.099	1.048	1.025	1.049	1.028

Table B66: Regression coefficients for four predictors when γ^2=0.99, ρ=0.9 and n=500

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	-0.082	-0.081	-0.081	-0.082	-0.081
β_1	0.997	0.994	0.973	0.995	0.973
β_2	1.093	1.072	1.055	1.068	1.050
β_3	0.339	0.436	0.470	0.438	0.478
β_4	1.419	1.344	1.327	1.345	1.325

Table B67: Coefficients for eight predictors when γ^2=0.7, ρ=0.1 and n=25

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	-0.065	-0.061	-0.095	-0.062	-0.096
β_1	0.931	0.949	0.865	0.953	0.872
β_2	1.222	1.177	1.247	1.179	1.246
β_3	0.899	0.888	0.933	0.886	0.930
β_4	1.000	1.018	0.949	1.013	0.948
β_5	0.993	0.992	0.928	0.992	0.928
β_6	1.279	1.243	1.300	1.248	1.305
β_7	0.830	0.855	0.818	0.855	0.819
β_8	1.087	1.083	1.050	1.082	1.051

Table B68: Coefficients for eight predictors when γ^2=0.7, ρ=0.1 and n=50

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	0.120	0.122	0.124	0.127	0.128
β_1	1.018	1.011	1.019	1.012	1.020
β_2	1.004	1.003	0.994	1.001	0.992
β_3	0.984	0.983	0.962	0.981	0.961
β_4	0.745	0.755	0.742	0.757	0.744
β_5	1.204	1.192	1.185	1.193	1.185
β_6	1.042	1.032	1.049	1.034	1.050
β_7	0.986	0.985	0.967	0.982	0.963
β_8	1.062	1.056	1.040	1.057	1.042

Table B69: Coefficients for eight predictors when γ^2=0.7, ρ=0.1 and n=200

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	0.017	0.017	0.015	0.017	0.016
β_1	1.013	1.012	1.010	1.012	1.009
β_2	0.965	0.965	0.953	0.964	0.953
β_3	1.038	1.036	1.033	1.037	1.034
β_4	0.878	0.880	0.870	0.880	0.871
β_5	1.013	1.013	1.005	1.013	1.005
β_6	1.045	1.043	1.036	1.042	1.035
β_7	1.256	1.251	1.247	1.251	1.248
β_8	0.813	0.815	0.810	0.815	0.810

Table B70: Coefficients for eight predictors when γ^2=0.9, ρ=0.1 and n=25

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	-0.063	-0.066	-0.069	-0.071	-0.075
β_1	0.522	0.777	0.676	0.776	0.672
β_2	0.862	0.858	0.871	0.861	0.872
β_3	1.611	1.325	1.470	1.331	1.479
β_4	0.731	0.841	0.774	0.841	0.777
β_5	1.091	0.995	1.015	0.990	1.005
β_6	1.090	1.069	1.125	1.062	1.114
β_7	1.107	1.085	1.062	1.083	1.066
β_8	1.104	1.115	0.991	1.122	1.005

Table B71: Coefficients for eight predictors when γ^2=0.9, ρ=0.1 and n=50

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	0.120	0.123	0.125	0.128	0.129
β_1	1.029	1.012	1.034	1.016	1.038
β_2	1.000	1.012	0.991	1.009	0.988
β_3	0.961	0.968	0.938	0.965	0.937
β_4	0.554	0.617	0.584	0.621	0.586
β_5	1.344	1.291	1.310	1.292	1.309
β_6	1.071	1.047	1.085	1.049	1.087
β_7	0.969	0.970	0.943	0.965	0.937
β_8	1.101	1.090	1.066	1.091	1.069

Table B72: Coefficients for eight predictors when γ^2=0.9, ρ=0.1 and n=200

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	0.053	0.053	0.052	0.053	0.052
β_1	1.217	1.207	1.203	1.207	1.202
β_2	1.036	1.033	1.022	1.035	1.024
β_3	0.869	0.875	0.857	0.876	0.859
β_4	0.918	0.922	0.912	0.922	0.912
β_5	1.142	1.133	1.134	1.132	1.133
β_6	1.133	1.125	1.137	1.126	1.138
β_7	0.722	0.733	0.725	0.733	0.724
β_8	0.882	0.886	0.872	0.884	0.870

Table B73: Coefficients for eight predictors when γ^2=0.99, ρ=0.1 and n=25

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	-0.045	0.009	-0.018	0.007	-0.023
β_1	-1.480	0.921	0.039	0.985	0.051
β_2	0.236	0.990	0.886	0.912	0.372
β_3	5.997	1.354	2.965	1.421	3.420
β_4	-2.251	0.664	0.000	0.673	0.001
β_5	2.152	0.896	0.772	0.897	0.824
β_6	0.792	1.163	1.196	1.143	1.270
β_7	1.127	1.155	1.296	1.124	1.181
β_8	1.991	1.236	1.199	1.284	1.287

Table B74: Coefficients for eight predictors when γ^2=0.99, ρ=0.1 and n=50

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	0.010	0.003	0.007	0.003	0.008
β_1	0.620	0.829	0.710	0.829	0.709
β_2	1.379	1.163	1.230	1.179	1.240
β_3	0.721	0.865	0.692	0.865	0.700
β_4	1.811	1.349	1.548	1.362	1.550
β_5	1.097	1.041	1.060	1.037	1.060
β_6	0.319	0.726	0.554	0.706	0.543
β_7	1.879	1.377	1.643	1.384	1.635
β_8	0.223	0.684	0.561	0.671	0.561

Table B75: Coefficients for eight predictors when γ^2=0.99, ρ=0.1 and n=200

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	0.053	0.051	0.050	0.051	0.050
β_1	1.709	1.483	1.579	1.483	1.578
β_2	1.132	1.070	1.049	1.074	1.054
β_3	0.600	0.742	0.644	0.744	0.647
β_4	0.766	0.848	0.796	0.848	0.796
β_5	1.471	1.298	1.368	1.295	1.364
β_6	1.443	1.301	1.388	1.303	1.391
β_7	0.149	0.419	0.357	0.418	0.357
β_8	0.649	0.751	0.681	0.747	0.677

Table B76: Coefficients for eight predictors when γ^2=0.99, ρ=0.1 and n=500

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	0.000	0.000	0.001	0.000	0.001
β_1	0.549	0.625	0.572	0.625	0.572
β_2	1.103	1.085	1.080	1.084	1.079
β_3	1.130	1.112	1.105	1.111	1.104
β_4	1.070	1.056	1.052	1.057	1.053
β_5	0.657	0.717	0.673	0.718	0.675
β_6	0.914	0.930	0.916	0.930	0.915
β_7	1.477	1.392	1.463	1.394	1.465
β_8	1.086	1.067	1.073	1.066	1.071

Table B77: Coefficients for eight predictors when γ^2=0.7, ρ=0.9 and n=25

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	0.523	0.547	0.520	0.587	0.563
β_1	0.533	0.610	0.602	0.612	0.599
β_2	1.073	1.111	1.017	1.121	1.030
β_3	0.764	0.781	0.699	0.784	0.707
β_4	0.833	0.828	0.814	0.826	0.810
β_5	0.856	0.860	0.861	0.854	0.856
β_6	1.347	1.243	1.291	1.240	1.293
β_7	1.305	1.244	1.292	1.236	1.284
β_8	0.783	0.789	0.741	0.789	0.744

Table B78: Coefficients for eight predictors when γ^2=0.7, ρ=0.9 and n=50

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	-0.169	-0.173	-0.177	-0.181	-0.186
β_1	1.236	1.216	1.205	1.221	1.215
β_2	0.906	0.917	0.899	0.915	0.897
β_3	1.008	1.018	0.986	1.018	0.987
β_4	0.803	0.814	0.791	0.811	0.783
β_5	1.006	1.006	0.980	1.009	0.984
β_6	1.382	1.358	1.322	1.361	1.329
β_7	1.516	1.474	1.474	1.473	1.470
β_8	0.713	0.735	0.740	0.732	0.740

Table B79: Coefficients for eight predictors when γ^2=0.7, ρ=0.9 and n=200

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	0.374	0.374	0.371	0.376	0.373
β_1	0.937	0.937	0.932	0.938	0.933
β_2	1.177	1.173	1.169	1.173	1.169
β_3	1.213	1.208	1.206	1.208	1.206
β_4	0.955	0.956	0.944	0.955	0.944
β_5	0.731	0.735	0.722	0.736	0.724
β_6	1.139	1.136	1.135	1.135	1.134
β_7	1.051	1.050	1.046	1.052	1.047
β_8	0.915	0.916	0.903	0.915	0.903

Table B80: Coefficients for eight predictors when γ^2=0.7, ρ=0.9 and n=500

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	0.139	0.138	0.138	0.140	0.139
β_1	0.955	0.954	0.948	0.956	0.950
β_2	1.142	1.141	1.134	1.140	1.133
β_3	0.876	0.876	0.868	0.876	0.867
β_4	1.086	1.085	1.079	1.085	1.079
β_5	0.940	0.940	0.932	0.940	0.932
β_6	1.046	1.046	1.038	1.046	1.039
β_7	1.115	1.114	1.112	1.114	1.112
β_8	0.767	0.768	0.760	0.769	0.761

Table B81: Coefficients for eight predictors when γ^2=0.9, ρ=0.9 and n=25

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	0.911	0.917	0.901	0.971	0.953
β_1	1.724	1.363	1.475	1.366	1.467
β_2	1.099	1.014	0.983	1.023	1.004
β_3	0.217	0.683	0.512	0.686	0.525
β_4	0.222	0.488	0.354	0.479	0.360
β_5	0.220	0.604	0.451	0.617	0.467
β_6	2.662	2.131	2.410	2.119	2.402
β_7	1.597	1.315	1.324	1.316	1.323
β_8	0.595	0.756	0.638	0.746	0.633

Table B82: Coefficients for eight predictors when γ^2=0.9, ρ=0.9 and n=50

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	-0.171	-0.182	-0.186	-0.191	-0.194
β_1	1.343	1.268	1.253	1.275	1.266
β_2	0.778	0.839	0.826	0.835	0.827
β_3	0.947	1.005	0.960	1.006	0.962
β_4	0.605	0.670	0.659	0.663	0.642
β_5	0.954	0.972	0.935	0.977	0.943
β_6	1.595	1.500	1.448	1.503	1.455
β_7	1.829	1.654	1.680	1.653	1.673
β_8	0.424	0.543	0.581	0.540	0.581

Table B83: Coefficients for eight predictors when γ^2=0.9, ρ=0.9 and n=200

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	0.060	0.060	0.058	0.061	0.059
β_1	0.597	0.615	0.617	0.617	0.619
β_2	0.827	0.836	0.813	0.835	0.812
β_3	1.101	1.094	1.087	1.096	1.089
β_4	0.626	0.645	0.639	0.647	0.640
β_5	1.300	1.282	1.270	1.280	1.268
β_6	1.547	1.519	1.511	1.520	1.512
β_7	1.472	1.448	1.441	1.447	1.439
β_8	0.439	0.464	0.476	0.463	0.475

Table B84: Coefficients for eight predictors when γ^2=0.9, ρ=0.9 and n=500

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	0.139	0.139	0.138	0.140	0.139
β_1	0.928	0.929	0.922	0.931	0.925
β_2	1.253	1.248	1.239	1.247	1.238
β_3	0.791	0.795	0.787	0.793	0.786
β_4	1.157	1.154	1.147	1.154	1.147
β_5	0.901	0.903	0.892	0.903	0.893
β_6	1.086	1.084	1.074	1.085	1.075
β_7	1.207	1.201	1.203	1.201	1.203
β_8	0.603	0.610	0.608	0.610	0.608

Table B85: Coefficients for eight predictors when γ^2=0.99, ρ=0.9 and n=25

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	-0.492	-0.443	-0.447	-0.470	-0.466
β_1	0.837	1.001	0.939	0.988	0.883
β_2	0.262	0.682	0.621	0.659	0.646
β_3	0.924	0.981	0.984	1.002	1.082
β_4	0.838	0.938	0.883	0.919	0.897
β_5	1.480	1.130	1.175	1.199	1.248
β_6	2.588	1.422	1.427	1.415	1.379
β_7	2.423	1.586	1.571	1.626	1.524
β_8	-1.353	0.235	0.290	0.171	0.253

Table B86: Coefficients for eight predictors when γ^2=0.99, ρ=0.9 and n=50

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	-0.066	-0.053	-0.056	-0.053	-0.055
β_1	1.286	1.060	1.037	1.075	1.052
β_2	0.364	0.729	0.666	0.708	0.649
β_3	0.359	0.679	0.621	0.674	0.610
β_4	2.716	1.658	1.918	1.684	1.926
β_5	2.379	1.552	1.723	1.574	1.714
β_6	-0.224	0.655	0.545	0.631	0.543
β_7	0.608	0.846	0.747	0.851	0.772
β_8	0.459	0.752	0.634	0.733	0.627

Table B87: Coefficients for eight predictors when γ^2=0.99, ρ=0.9 and n=200

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	0.802	0.802	0.802	0.811	0.810
β_1	2.230	1.811	1.847	1.812	1.846
β_2	1.687	1.449	1.453	1.444	1.446
β_3	0.737	0.812	0.794	0.817	0.796
β_4	1.345	1.201	1.178	1.200	1.176
β_5	0.276	0.522	0.526	0.522	0.526
β_6	0.337	0.593	0.577	0.594	0.583
β_7	1.191	1.128	1.103	1.132	1.105
β_8	0.235	0.518	0.515	0.514	0.515

Table B88: Coefficients for eight predictors when γ^2=0.99, ρ=0.9 and n=500

Coefficient	OLS	RR	Lasso	TR	TLasso
β_0	-0.046	-0.046	-0.046	-0.046	-0.047
β_1	1.257	1.227	1.184	1.227	1.184
β_2	0.392	0.494	0.521	0.491	0.518
β_3	0.124	0.280	0.377	0.280	0.378
β_4	1.580	1.472	1.451	1.470	1.448
β_5	0.667	0.723	0.697	0.724	0.699
β_6	1.988	1.808	1.796	1.809	1.796
β_7	1.243	1.212	1.169	1.215	1.172
β_8	0.848	0.883	0.855	0.883	0.855