Journal of Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.7546461

LINEAR MODEL IN THE PRESENCE OF MULTICOLLINEAR PREDICTORS AND AUTOCORRELATED ERRORS: INSIGHTS FROM REGULARIZED AND ROBUST REGRESSION METHODS

Nathaniel Howard[1], David Kwamena Mensah[2]*, Benjamin Arkorful[3]
[1, 2, 3] Department of Statistics, University of Cape Coast, Ghana
*Corresponding Author

Abstract
The quest for unified solution for handling multicollinearity and autocorrelation jointly has resulted in a great deal of interest in two-stage approaches within the linear modelling paradigm recently. This paper contributes to this growing interest with Two-Stage Ridge Quantile Regression (TRQR) and Two-Stage Lasso Quantile Regression (TLQR) methods. The two-stage methods fundamentally involve the application of appropriate transformation to linear regression data followed by regularization to control respectively, autocorrelation and multicollinearity. The aim is to determine if regularized and robust regression methods reduce total model error, and find which method is most effective for dealing with autocorrelation and multicollinearity problems. The utility of the methods were assessed using simulations based on models with small (2) to relatively large (8) predictors, in comparison with other regression methods using the Mean Squared Error criterion. The results indicate that the TRQR method, at 0.5 quantile level, is most suitable for handling multicollinearity and autocorrelation problems with many (8 or more) predictor variables. However, the Two-Stage Ridge Regression method performs better with few (2 or less) predictor variables. The findings show that each method is affected by sample size, number of predictors or multicollinearity level.
Keywords: Autocorrelation, Multicollinearity, Two-Stage method, Ridge regression, Lasso regression, Mean squared error

1. Introduction

The Ordinary Least Squares (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, 2012). When $\mathbf{X}^T\mathbf{X}$ is non-singular, the least squares estimate can be evaluated directly from the data by

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

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. 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.

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 violates the assumption that the design matrix $\mathbf{X}$ is of full rank, rendering OLS estimation unfeasible. The nature of multicollinearity can be classified into perfect (or exact) and imperfect (or approximate). 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 (Hashem, 2014).

The case of imperfect multicollinearity is the most common situation when the variables are highly, but not perfectly correlated. The matrix $\mathbf{X}^T\mathbf{X}$ and matrix $\mathbf{X}$ have full rank, however, it is not far from being rank-deficient. Thus, the matrix $\mathbf{X}^T\mathbf{X}$ is quasi-singular (ill-conditioned). 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, 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, $var(\hat{\boldsymbol{\beta}}_{OLS}) = \sigma^2(\mathbf{X}^T\mathbf{X})^{-1}$ (Giacalone, Panarello, & Mattera, 2018). In this scenario, regression estimates are determinate but possess large standard error implying that the coefficients cannot be estimated with great precision.

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, 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.

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. According to studies in 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 (Baye & Parker, 1984), Liu estimator (Liu, 1993), $r-d$ class estimator (Kaciranlar & Sakallioglu, 2001), Principal component two-parameter (PCTP) estimator (Chang & Yang, 2012), and $r-\left(k,d\right)$ class estimator (Ozkale, 2012).

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. 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 (Gujarati, 2003). Forecasts based on OLS in the presence of autocorrelation will be unbiased, but inefficient due to inefficient estimates of the regression parameters.

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 (Paris &

Winstein, 1954), Cochrane-Orcutt estimator (Cochrane & Orcutt, 1949), Hildreth and Lu estimator (Hildreth & Lu, 1960) and Maximum Likelihood estimator (Beach & Mackinnon, 1978).

Inevitably, both problems (autocorrelation and multicollinearity) can coexist in a linear regression model, hence proven in literature (Ayinde, Lukman & Arowolo, 2015; Bayhan & Bayhan, 1998). 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). Eledum and Zahri (2013) introduced the two-stage ridge regression (TR) method to correct for both problems. Dawoud and Kaçıranlar (2016) proposed the Two-Stage Liu (TL) method to address autocorrelation and multicollinearity issues in linear models. Arowolo, Adewale and Kayode (2016) compared the Two-Stage Principal Component regression (T-PC) and Two-Stage Partial Least Square (T-PLS). Ozbay, Kaçıranlar and Dawoud (2017) proposed Feasible Generalized Restricted Ridge regression (FGRR) method to take account of both problems. Zubair and Adenomon (2021) proposed the two-stage $K-L$ estimator to mitigate both problems.

Regularization (for example, Ridge and Lasso) in Quantile regression (QR) has been shown to improve prediction accuracy (Bager, 2018; Li & Zhu, 2008; Suhail, Chand, & Kibria, 2020). As a result, it is vital to explore regularized and robust regression methods in handling multicollinearity and autocorrelation problems, hence, the need for this research. The research seeks to develop Two-Stage Ridge Quantile regression (TRQR) and Two-Stage Lasso Quantile regression (TLQR) methods, and compare their performances with other popular methods in literature. 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.

The literature reveals a number of methods for addressing multicollinearity and autocorrelation problems. This paper exploits the regularized and robust regression methods to develop Two-Stage Ridge Quantile regression (TRQR) and Two-Stage Lasso Quantile regression (TLQR) methods in handling multicollinearity and autocorrelation problems.

The progression of the rest of the paper is organized as follows. Section 2 develops the TRQR, TLQR and other regression methods. Section 3 presents the simulation study conducted to evaluate the methods. Section 4 presents the results and, Section 5 concludes the paper.

1. Methodology

1.1 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 in matrix form as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}, \qquad (2)$$

where

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

And $\boldsymbol{\beta}$ is a $(p+1)\times 1$ vector of unknown parameters and $\mathbf{e}$ is an $n\times 1$ vector of random error terms 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 $\mathbf{X}$. The least square estimator is given by $\hat{\beta}=(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$

1.2 Two-Stage Method

The two-stage method employs variable transformation in particular to deal with autocorrelation. Various transformation approaches have been proposed by different authors. These approaches are categorized into those that utilize $\mathbf{P}^*$ matrix for transformation and those that use $\mathbf{P}$ matrix for transformation.

We adopted the $\mathbf{P}$ transformation matrix as used by other authors (Dawoud & Kaçıranlar, 2016; Eledum & Zahri, 2013; Zubair & Adenomon, 2021). 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, as given below.

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

Now, pre-multiplying both sides of Equation (2) by the $\mathbf{P}$ transformation matrix, we get an equivalent linear model

$$\mathbf{Py}=\mathbf{PX}\boldsymbol{\beta}+\mathbf{Pe} \tag{3}$$

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

$$\mathbf{y}^*=\mathbf{X}^*\boldsymbol{\beta}+\mathbf{e}^* \tag{4}$$

Model (4) satisfies the error assumption $e^*\sim N(0,\sigma^2\mathbf{I}_n)$ based on the error model assumed in Equation (1).

Therefore, the least squares estimate for the Model (4), which is called the Two-stage is

$$\hat{\beta}_{TS}=(\mathbf{X}^{*T}\mathbf{X}^*)^{-1}\mathbf{X}^{*T}\mathbf{y}^* \tag{5}$$

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

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

The matrix $\mathbf{P}$ can also 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 (3) have all of the optimal OLS properties, and the usual inferences could be true.

Now, $\mathbf{X}^{*T}\mathbf{X}^* = \mathbf{X}^T\mathbf{P}^T\mathbf{P}\mathbf{X} = \mathbf{X}^T\Omega^{-1}\mathbf{X}$ and $\mathbf{X}^{*T}\mathbf{y}^* = \mathbf{X}^T\mathbf{P}^T\mathbf{P}\mathbf{y} = \mathbf{X}^T\Omega^{-1}\mathbf{y}$
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}, the Two-stage estimator is defined as

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

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

$$\hat{\beta}_{TS} = (\mathbf{X}^T\hat{\Omega}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\hat{\Omega}^{-1}\mathbf{y} \tag{6}$$

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. Regularization is then applied to handle the multicollinearity problem.

2.3 Two-Stage Ridge Regression Method (TR)

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.

$$\hat{\beta}_{Ridge} = \min_{\beta}\left\{\sum_{i=1}^{n}(\mathbf{y}_i - \sum_{j=1}^{p}\mathbf{X}_{ij}\beta_j)^2\right\}, \; s.t. \; \sum_{j=1}^{p}\beta_j^2 \le t, \; t \ge 0 \tag{7}$$

Equivalently, the following minimization problem defines Ridge regression

$$\hat{\beta}_{Ridge} = \min_{\beta}\left\{\sum_{i=1}^{n}(\mathbf{y}_i - \sum_{j=1}^{p}\mathbf{X}_{ij}\beta_j)^2 + k\sum_{j=1}^{p}\beta_j^2\right\}, \; k \ge 0 \tag{8}$$

where $i = 1,2,\ldots,n$; $j = 1,2,\ldots,p$ and the amount of shrinkage is controlled by the regularization parameter, k. The parameter t in Equation (7) is clearly related to the parameter k in Equation (8). This means that for a given value k, there exists a value t for which the estimation Equations (7) and (8) yield the same result. Therefore, the regularized solution is given as

$$\hat{\beta}_{RR} = (\mathbf{X}^T\mathbf{X} + k\mathbf{I}_p)^{-1}\mathbf{X}^T\mathbf{y} \quad (9)$$

where $\mathbf{I}$ is the $p \times p$ identity matrix and the constant $k > 0$ is the regularization parameter.

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

$$\hat{\beta}_{TR} = \min_{\beta}\left\{\sum_{i=1}^{n}(\mathbf{y}_i^* - \sum_{j=1}^{p}\mathbf{X}_{ij}^*\beta_j)^2 + k\sum_{j=1}^{p}\beta_j^2\right\},\ k \geq 0 \quad (10)$$

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

$$\hat{\beta}_{TR} = (\mathbf{X}^{*T}\mathbf{X}^* + k\mathbf{I}_p)^{-1}\mathbf{X}^{*T}\mathbf{y}^* \quad (11)$$

Following Equations (6) and (9), the Two-stage Ridge regression estimator proposed by Eledum and Zahri (2013) takes the form

$$\hat{\beta}_{TR} = (\mathbf{X}^T\hat{\Omega}^{-1}\mathbf{X} + k\mathbf{I}_p)^{-1}\mathbf{X}^T\hat{\Omega}^{-1}\mathbf{y}\,. \quad (12)$$

1.3 Two-Stage Lasso Method (TLasso)

The Lasso method 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

$$\hat{\beta}_{Lasso} = \min_{\beta}\left\{\sum_{i=1}^{n}(\mathbf{y}_i - \sum_{j=1}^{p}\mathbf{X}_{ij}\beta_j)^2\right\},\ s.t.\ \sum_{j=1}^{p}|\beta| \leq t,\ t \geq 0 \quad (13)$$

An equivalent form of the Lasso is

$$\hat{\beta}_{Lasso} = \min_{\beta}\left\{\sum_{i=1}^{n}(\mathbf{y}_i - \sum_{j=1}^{p}\mathbf{X}_{ij}\beta_j)^2 + k\sum_{j=1}^{p}|\beta|\right\},\ k \geq 0 \quad (14)$$

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

In matrix form

$$\hat{\beta}_{Lasso} = \min_{\beta}\left\|\mathbf{y} - \mathbf{X}\beta\right\|_2^2 + k\left\|\beta\right\|_1 \quad (15)$$

The Lasso estimation is a convex optimization issue that can be addressed using a quadratic programming algorithm for a given k. Therefore, solving Equation (14) does not provide a closed form expression.

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. We replaced $\mathbf{y}$ and $\mathbf{X}$ in Equation (14) by $\mathbf{y}^*$ and $\mathbf{X}^*$, respectively. The solution can be expressed as

$$\hat{\beta}_{TLasso} = \min_{\beta} \left\{ \sum_{i=1}^{n} (\mathbf{y}_i^* - \sum_{j=1}^{p} \mathbf{X}_{ij}^* \beta_j)^2 + k \sum_{j=1}^{p} |\beta| \right\},\ k \geq 0 \tag{16}$$

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 employed $\mathbf{y}^*$ vector of response variable and $\mathbf{X}^*$ matrix of predictors.

1.4 Two-Stage Regularized and Robust Regression Methods

The robust regression method employed in this paper is the Quantile Regression method (QR). The QR method, as introduced by Koenker and Bassett (1978), is widely used to describe the distribution of a response variable given a set of predictor variables. QR allows estimating the entire distribution of the response variable's conditional quantiles.

A typical QR model is formulated as

$$Q_\tau(\mathbf{y} \mid \mathbf{X}) = \mathbf{X}^T \boldsymbol{\beta}_\tau \tag{17}$$

where $Q_\tau(\cdot|\cdot)$ is the conditional quantile function for the τth conditional quantile with $0 < \tau < 1$, τ determines the quantile level, $\mathbf{X}$ is a design matrix 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. We employed 0.25, 0.5 and 0.75 quantile levels in the analysis of the models.

Regression coefficients β_τ can be estimated by

$$\hat{\beta}_\tau = \min_{\beta_\tau} \sum_{i=1}^{n} \rho_\tau (\mathbf{y}_i - \mathbf{X}_i^T \boldsymbol{\beta}_\tau) \tag{18}$$

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.

1.4.1 Two-Stage Ridge Quantile Regression Method (TRQR)

The Ridge Quantile Regression (RQR) used 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 using

$$\hat{\beta}_{RQR} = \min_{\beta} \left\{ \sum_{i=1}^{n} \rho_\tau (\mathbf{y}_i - \mathbf{X}_i^T \beta_\tau) + k \sum_{j=1}^{p} (\beta_j)^2 \right\},\ k > 0 \tag{19}$$

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

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. In particular, replacing $\mathbf{y}$ and $\mathbf{X}$ in Equation (19) by $\mathbf{y}^*$ and $\mathbf{X}^*$, respectively, yields the TRQR estimate of β as

$$\hat{\beta}_{TRQR} = \min_{\beta} \left\{ \sum_{i=1}^{n} \rho_\tau (\mathbf{y}_i^* - \mathbf{X}_i^{*T} \beta_\tau) + k \sum_{j=1}^{p} (\beta_j)^2 \right\},\ k > 0 \tag{20}$$

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

Thus, the Two-Stage Ridge Quantile regression obtained by fitting the Ridge Quantile Regression using the data $\mathbf{y}^*$ and $\mathbf{X}^*$.

1.4.2 Two-Stage Lasso Quantile Regression Method (TLQR)

The L_1-penalty is added to the quantile loss function to formulate the Lasso Quantile Regression (LQR) method.

The LQR estimate β is given by

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

where $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, p$ and $k > 0$ is the regularization parameter controlling the amount of penalty. The second term in Equation (21) 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.

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 (21) by $\mathbf{y}^*$ and $\mathbf{X}^*$, respectively, yielding

the TLQR estimate β as

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

$i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, p$ 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.

2. Simulation Study

We evaluated the performance of the regression methods through simulation. The implementation of the methods was done in R. Simulation followed McDonald and Galarneau (1975) and Kibria (2003). The predictor variables were generated 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 regression coefficients are assumed to be unity with β_0 taken to be zero. Also, μ_i is generated from an AR (1) process as $\mu_i = \rho\mu_{i-1} + e_i$, $i = 1, 2, \ldots, n$, where e_i are independent normal pseudo random numbers and ρ is autoregressive coefficient such that $|\rho|<1$. The error terms were generated from standard normal, $e_i \sim N(0,1)$, so that , $\mu_i \sim N\left(0, \frac{\sigma_e^2}{1-\rho^2}\right)$ based on the distributional property of the autocorrelated error terms.

The design was created by varying multicollinearity levels (0.7, 0.9 and 0.99), with the number of predictor variables (p = 2 and 8), sample size (n = 25, 50, 200 and 500) and degrees of autocorrelation (ρ = 0.1 and 0.9). It is of interest 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. The simulation was replicated 2000 times in order to obtain the approximate distribution.

We generated dataset with 0.7, 0.9 and 0.99 multicollinearity levels, 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 and eight predictor variables with different sample sizes; n = 25, 50, 200 and 500.

We use the Mean Squared Error (MSE) criterion to investigate the performance of the OLS, RR, Lasso, TR, TLasso, RQR, LQR, TRQR and TLQR methods. The estimated MSE for each of the regression methods is given by

$$MSE(\hat{\beta}) = \frac{1}{2000}\sum_{r=1}^{2000}(\hat{\beta}_r - \beta)^T(\hat{\beta}_r - \beta)$$

where $r = 1, 2, \ldots, 2000$, $\hat{\beta}_r$ denotes 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.

4. Results and Discussion

This section presents the results obtained from the simulation study.

Table 1 shows the simulation results of the estimated MSEs for two predictor variables when ρ = 0.1 and across the different levels of multicollinearity (0.7, 0.9 and 0.99). The results revealed that, TR method outperforms the other methods in many of the cases, especially with large sample sizes. In addition, it can be observed that TRQR method produced better results than TLQR method, but their performances to the other methods (RR, Lasso, TR and TLasso) are not satisfactory when there is multicollinearity with less or no autocorrelated errors for two predictor variables. Therefore, the TR method produced the most efficient results in terms of lower MSE when there is multicollinearity with less or no autocorrelated errors.

Table 2 presents the estimated MSEs for the two predictor variables when ρ = 0.9. It is evident that the two-stage methods record better performance at levels of multicollinearity with increased sample sizes in general. Nevertheless, TR outperforms all the two-stage methods. Furthermore, the TR and TLasso exhibit better performance than the RR method and Lasso. In the case of the OLS, better competition is registered with Lasso for large sample size (n = 500) at 0.7 multicollinearity level. Interestingly, it can also be seen that the TRQR method at □ □ 0.5 outperforms the other methods (RQR, LQR and TLQR). But, the TLQR method is a close contender to TRQR method in terms of performance. Again, TRQR and TLQR methods have smaller MSE values than the RQR method and LQR method respectively. However, the TR method and TLasso method performs better than the TRQR method and TLQR method in the presence of multicollinearity and autocorrelation problems. Finally, it clear that the TR is effective for handling multicollinearity and autocorrelations in linear models with less number of predictors, in particular, two predictor variables models.

4.1 Simulation Results for Two Predictor Variables

Table 1: Simulation Results of MSE for 2 Predictor Variables when ρ = 0.1

γ^2	n	OLS	RR	Lasso	TR	TLasso	RQR	LQR	TRQR	TLQR
							τ = 0.5	τ = 0.5	τ = 0.5	τ = 0.5
0.7	25	0.2418	**0.1930**	0.2392	0.1999	0.2451	0.2961	0.4333	0.3048	0.5177
	50	0.1101	0.0948	0.1111	**0.0935**	0.1092	0.1405	0.1722	0.1348	0.1702
	200	0.0277	0.0262	0.0278	**0.0250**	0.0266	0.0373	0.0399	0.0359	0.0446
	500	0.0104	0.0101	0.0104	**0.0096**	0.0099	0.0146	0.0149	0.0139	0.0192
0.9	25	0.5595	**0.3086**	0.4660	0.3317	0.4851	0.6378	1.2905	0.4850	0.7907
	50	0.2524	**0.1510**	0.2422	0.1528	0.2432	0.2232	0.4127	0.2274	0.3824
	200	0.0645	0.0476	0.0646	**0.0464**	0.0631	0.0698	0.0905	0.0681	0.0877
	500	0.0239	0.0198	0.0240	**0.0192**	0.0233	0.0298	0.0344	0.0288	0.0330
0.99	25	4.8859	1.8475	**1.7691**	2.1094	1.9196	3.6056	5.5099	2.0179	3.2950
	50	2.1929	**0.7975**	0.8884	0.8368	0.8506	0.9590	1.6867	0.9758	2.0657
	200	0.5652	0.2260	0.3917	**0.2243**	0.3916	0.3186	0.6547	0.3215	0.6444
	500	0.2078	0.0845	0.1895	**0.0835**	0.1875	0.1358	0.2928	0.1318	0.2851

Table 2: Simulation Results of MSE for 2 Predictor Variables when ρ = 0.9

γ^2	n	OLS	RR	Lasso	TR	TLasso	RQR	LQR	TRQR	TLQR
							τ = 0.5	τ = 0.5	τ = 0.5	τ = 0.5
0.7	25	3.1218	2.9191	3.0372	**0.4514**	0.4850	3.4062	3.5628	0.5430	0.7035
	50	1.8836	1.7845	1.8648	**0.1233**	0.1309	2.0673	2.2119	0.1633	0.1811
	200	0.5666	0.5475	0.5665	**0.0194**	0.0200	0.6365	0.6682	0.0264	0.0273
	500	0.2312	0.2266	0.2312	**0.0070**	0.0072	0.2640	0.2714	0.0096	0.0098
0.9	25	4.0379	3.2257	3.4835	**0.5394**	0.6674	3.8431	4.3489	0.6977	1.0326
	50	2.3865	1.9389	2.1587	**0.1597**	0.2140	2.4807	2.6633	0.2386	0.2981
	200	0.7332	0.6117	0.7195	**0.0329**	0.0405	0.7563	0.9011	0.0457	0.0547
	500	0.2989	0.2585	0.2984	**0.0134**	0.0150	0.3181	0.3751	0.0188	0.0206
0.99	25	16.4996	7.8240	8.3649	**1.6548**	1.7030	8.6851	9.1177	1.7844	3.4565
	50	9.2609	4.1998	4.6072	**0.5295**	0.6952	4.5766	6.5239	0.6584	1.2261
	200	2.9996	1.3729	1.4481	**0.1292**	0.2717	1.4670	2.2833	0.1901	0.4129
	500	1.2198	0.5621	0.7481	**0.0548**	0.1202	0.6346	1.1716	0.0814	0.1675

4.2 Simulation Results for 8 Predictor Variables model

In this section the results of the 8 predictor variable linear model is presented.

Table 3 records the performance statistics of the methods for the 8 predictor variables linear model for ρ = 0.1, □ □ 0.5 at the levels of the multicollinearity and sample sizes considered in the simulation study. Clearly, the RQR and TRQR outperform the other methods in many of the cases when there is multicollinearity with less or no auto-correlated errors. However, the RQR gives better results than the TRQR. The TRQR appears to have an improved performance as sample size increases with high multicollinearity level. Also, the performance of TLQR method is not satisfactory with less or no autocorrelated errors.

Table 3: Simulation Results of MSE for 8 Predictor Variables when ρ = 0.1

γ^2	n	OLS	RR	Lasso	TR	TLasso	RQR	LQR	TRQR	TLQR
							τ = 0.5	τ = 0.5	τ = 0.5	τ = 0.5
0.7	25	1.6787	1.0687	1.4790	1.1084	1.5264	**0.9183**	2.1535	0.9943	2.2671
	50	0.6195	0.5217	0.6123	0.5300	0.6236	**0.4775**	0.8355	0.4885	0.8258
	200	0.1292	0.1249	0.1299	**0.1227**	0.1278	0.1418	0.1669	0.1389	0.1633
	500	0.0499	0.0493	0.0503	**0.0483**	0.0493	0.0590	0.0629	0.0571	0.0604
0.9	25	4.8017	1.8207	2.9806	1.9794	3.1281	**1.6880**	4.8097	1.9172	4.8175
	50	1.7925	1.1115	1.4981	1.1314	1.5334	**0.7922**	2.2668	0.8509	2.2643
	200	0.3792	0.3417	0.3787	0.3382	0.3752	0.2469	0.4771	**0.2467**	0.4684
	500	0.1441	0.1384	0.1445	0.1365	0.1425	0.1208	0.1792	**0.1169**	0.1739
0.99	25	46.7882	**6.6286**	8.7624	8.7980	9.3431	8.6848	17.7540	9.5005	18.2603
	50	18.2252	**2.8240**	6.2494	3.1046	6.3540	4.0847	12.6518	4.2099	12.7195
	200	3.7838	1.6310	2.5233	1.6145	2.5080	0.9980	4.0489	**0.9877**	3.9894
	500	1.4164	0.9746	1.1889	0.9634	1.1793	0.4058	1.7381	**0.4015**	1.6947

Table 4 reports the performance statistics of the methods for □ □ 0.5, ρ = 0.9, at the various levels of multicollinearity and autocorrelations assumed for the study. Again, it clearly seen that TRQR at □ □ 0.5 outperforms the other methods when the degree of autocorrelation is high and across the different degrees of multicollinearity, except in five cases where TR is superior. It can be seen that MSE values of TR is relatively smaller than the estimated MSE values of TRQR at □ □ 0.5 for sample sizes $50 \leq n \leq 500$ with moderate multicollinearity level (γ^2 = 0.70). Again, the TR method is best for small sample sizes ($25 \leq n \leq 50$) with high autocorrelation and severe multicollinearity (γ^2 = 0.99). The TRQR performs well as the multicollinearity level increases with increasing sample size. Also, the TRQR method at □ □ 0.5 has the least MSE values in comparison to RQR, LQR and TLQR methods for all sample sizes (n = 25, 50, 200 and 500) and across the various degrees of multicollinearity (0.7, 0.9 and 0.99). However, the TLQR appears to have an improved performance over OLS, RR, Lasso, RQR and LQR when there is multicollinearity and high autocorrelated errors. Generally, the results suggest that TRQR at □ □ 0.5 can render significant control for multicollinearity and autocorrelation problems in linear models.

Table 4: Simulation Results of MSE for 8 Predictor Variables when ρ = 0.9

γ^2	n	OLS	RR	Lasso	TR	TLasso	RQR	LQR	TRQR	TLQR
							τ = 0.5	τ = 0.5	τ = 0.5	τ = 0.5
0.7	25	7.1675	5.4677	5.6017	1.9071	2.3615	4.9676	7.6637	**1.7595**	2.3773
	50	3.8086	3.4473	3.4439	**0.5132**	0.5905	3.0516	4.7823	0.5191	0.7578

	200	1.0566	1.0361	1.0493	**0.0796**	0.0828	0.9632	1.3664	0.0919	0.1012
	500	0.4308	0.4276	0.4313	**0.0290**	0.0296	0.4346	0.5444	0.0337	0.0348
0.9	25	15.6344	7.4124	8.2774	2.5912	3.7315	7.0509	8.2266	**2.4752**	3.6621
	50	8.1538	5.6426	5.4864	0.9008	1.2613	4.3174	8.7226	**0.7439**	1.8382
	200	2.2204	2.0453	1.9599	0.2007	0.2241	1.2369	2.8869	**0.1701**	0.2719
	500	0.9061	0.8774	0.8805	0.0789	0.0824	0.5506	1.2039	**0.0760**	0.0965
0.99	25	127.970	19.349	19.329	**7.0111**	9.3557	20.017	44.122	7.488	14.117
	50	66.8472	11.2721	13.2799	**1.6489**	4.8601	12.676	34.232	2.4334	9.8676
	200	17.9437	7.9578	6.4107	0.8966	1.6259	3.7701	13.180	**0.6064**	2.4553
	500	7.3262	5.0926	3.9324	0.5441	0.7300	1.6444	6.8617	**0.2472**	0.9558

The performance comparison of only the quantiles-based methods in implementation for both the two and eight predictors linear models at the assumed levels of multicollinearity, and error autocorrelations, and quantile levels are presented in Table A1 and Table A2 in the Appendix. It can be observed that the simulated data support 0.5 quantile regression model evidenced by RQR, LQR, TRQR and TLQR exhibiting better performance at □ □ 0.5 than the other levels, □ □ 0.25 and □ □ 0.75.

5. Conclusion

In this paper, we have developed Two-Stage Ridge Quantile regression (TRQR) and Two-Stage Lasso Quantile regression (TLQR) for linear models. The applicability of the methods to both multicollinearity and autocorrelation problems is examined in comparison with other regression methods based on simulation for linear models with small (2) to relatively large (8) predicators. The simulation results showed that the OLS estimates could not perform well with regard to their MSE in the existence of autocorrelation and multicollinearity. In the presence of multicollinearity with less or no autocorrelation, both the TR method and RR method are the best methods for few predictor variables. However, the TR method performance improves with large sample sizes. Again, the study found that the TR method is the best method in handling multicollinearity and autocorrelation problems in a dataset with few predictor variables. Furthermore, sample size has a significant impact on method performance at all levels of autocorrelation and multicollinearity. The MSEs of the methods decrease with increased sample size. Generally, TR method is most suitable for addressing multicollinearity and autocorrelation problems with few predictor variables.

In existence of multicollinearity and sufficiently high degrees of autocorrelation for many predictor variables, the TRQR method at 0.5 quantile, is seen to be the best. However, in the presence of multicollinearity with less or no autocorrelation, the RQR method has minimum MSE values for a small sample size and the TRQR method at 0.5 quantile level has minimum MSE values for a large sample size. The study found that sample size and number of predictors in a model have significant impact on the performances of the regression methods for datasets which include possible multicollinearity and autocorrelation issues. Increasing the degree of multicollinearity between the predictor variables also has an adverse effect on the regression methods. Overall, the TRQR method at 0.5 quantile level is found to be suitable for addressing multicollinearity and autocorrelation problems with many predictor variables.

Acknowledgement

The authors would like to thank the University of Cape Coast and Government of Ghana for research support.

References

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., 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.

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

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.

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

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

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.

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

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.

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.

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.

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

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.

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.

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.

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.

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

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).

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.

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

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

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. (2021). Comparison of estimators efficiency for linear regressions with joint presence of autocorrelation and multicollinearity. Science World Journal, 16(2), 103-109.

Appendix

Table A1: Simulation Results of MSE for 2 Predictors linear model

ρ	γ^2	n	RQR $\tau = 0.25$	RQR $\tau = 0.5$	RQR $\tau = 0.75$	LQR $\tau = 0.25$	LQR $\tau = 0.5$	LQR $\tau = 0.75$	TRQR $\tau = 0.25$	TRQR $\tau = 0.5$	TRQR $\tau = 0.75$	TLQR $\tau = 0.25$	TLQR $\tau = 0.5$	TLQR $\tau = 0.75$
0.1	0.7	25	0.7807	0.2961	0.7767	0.8793	0.4333	0.9822	0.8148	0.3048	0.7553	0.8544	0.5177	0.9011
		50	0.6936	0.1405	0.6122	0.6619	0.1722	0.6520	0.5920	0.1348	0.5954	0.6377	0.1702	0.6355
		200	0.5113	0.0373	0.5036	0.5153	0.0399	0.5072	0.5034	0.0359	0.4956	0.5073	0.0446	0.4988
		500	0.4869	0.0146	0.4823	0.4883	0.0149	0.4833	0.4804	0.0139	0.4770	0.4818	0.0192	0.4776
	0.9	25	1.0613	0.6378	0.9779	4.1934	1.2905	5.3821	1.0836	0.4850	1.3774	2.0021	0.7907	3.0634
		50	0.7220	0.2232	0.7143	0.9645	0.4127	0.8830	0.6956	0.2274	0.7052	0.8620	0.3824	0.8592
		200	0.5496	0.0698	0.5396	0.5760	0.0905	0.5643	0.5420	0.0681	0.5320	0.5668	0.0877	0.5556
		500	0.5037	0.0298	0.4979	0.5119	0.0344	0.5050	0.4968	0.0288	0.4923	0.5051	0.0330	0.4992
	0.99	25	2.5548	3.6056	2.5836	4.1884	5.5099	4.0923	2.6346	2.0179	2.6682	2.3891	3.2950	10.8903
		50	1.5113	0.9590	1.5300	2.2501	1.6867	2.3165	1.4815	0.9758	1.5188	2.1375	2.0657	2.6623
		200	0.8251	0.3186	0.7930	1.2197	0.6547	1.1709	0.8244	0.3215	0.7846	1.1980	0.6444	1.1651
		500	0.6159	0.1358	0.5981	0.8148	0.2928	0.7809	0.6069	0.1318	0.5918	0.8005	0.2851	0.7715
0.9	0.7	25	4.7676	3.4062	4.7314	17.4131	3.5628	4.9290	1.0331	0.5430	1.0603	1.1215	0.7035	1.6039
		50	3.9214	2.0673	3.9787	4.0956	2.2119	4.1545	0.6367	0.1633	0.6301	0.6592	0.1811	0.6649
		200	2.9992	0.6365	2.8515	3.0380	0.6682	2.8937	0.5039	0.0264	0.4957	0.5065	0.0273	0.4972
		500	2.6709	0.2640	2.6125	2.6826	0.2714	2.6227	0.4831	0.0096	0.4794	0.4842	0.0098	0.4802
	0.9	25	5.2526	3.8431	5.2313	6.1786	4.3489	5.7998	1.2039	0.6977	1.2326	1.4324	1.0326	1.4468
		50	4.2068	2.4807	4.3201	4.5940	2.6633	4.6156	0.6990	0.2386	0.9177	0.8464	0.2981	1.1299
		200	3.1292	0.7563	2.9820	3.2928	0.9011	3.1461	0.5277	0.0457	0.5165	0.5403	0.0547	0.5306
		500	2.7266	0.3181	2.6703	2.7977	0.3751	2.7407	0.4934	0.0188	0.4893	0.4981	0.0206	0.4936
	0.99	25	11.8039	8.6851	13.0519	51.4255	9.1177	43.2098	2.2049	1.7844	2.4871	5.4757	3.4565	26.9616
		50	6.0713	4.5766	6.0797	8.6143	6.5239	8.6903	1.1922	0.6584	1.1804	1.8261	1.2261	1.8404
		200	3.8243	1.4670	3.6890	4.6579	2.2833	4.5712	0.6859	0.1901	0.6652	0.9598	0.4129	0.9368
		500	3.0248	0.6346	2.9683	3.6059	1.1716	3.6044	0.5569	0.0814	0.5463	0.6714	0.1675	0.6541

Table A2: Simulation Results of MSE for 8 Predictors linear model

ρ	γ^2	n	RQR $\tau = 0.25$	RQR $\tau = 0.5$	RQR $\tau = 0.75$	LQR $\tau = 0.25$	LQR $\tau = 0.5$	LQR $\tau = 0.75$	TRQR $\tau = 0.25$	TRQR $\tau = 0.5$	TRQR $\tau = 0.75$	TLQR $\tau = 0.25$	TLQR $\tau = 0.5$	TLQR $\tau = 0.75$
		25	1.4836	0.9183	1.4609	2.4872	2.1535	2.3370	1.4948	0.9943	1.4816	2.8295	2.2671	2.4334

0.1	0.7	50	1.0019	0.4775	1.0030	1.4687	0.8355	1.4803	0.9960	0.4885	0.9904	1.4439	0.8258	1.4650
0.1	0.7	200	0.6625	0.1418	0.6611	0.7337	0.1669	0.7396	0.6552	0.1389	0.6543	0.7270	0.1633	0.7281
0.1	0.7	500	0.5588	0.0590	0.5631	0.5857	0.0629	0.5896	0.5511	0.0571	0.5545	0.5789	0.0604	0.5816
0.1	0.9	25	2.3006	1.6880	2.1773	4.8624	4.8097	4.8390	2.3092	1.9172	2.4604	4.8754	4.8175	4.8510
0.1	0.9	50	1.3070	0.7922	1.3552	3.0365	2.2668	3.0898	1.3861	0.8509	1.3636	2.9947	2.2643	3.0737
0.1	0.9	200	0.7315	0.2469	0.7314	1.1098	0.4771	1.1205	0.7155	0.2467	0.7240	1.0930	0.4684	1.1106
0.1	0.9	500	0.5754	0.1208	0.5798	0.7290	0.1792	0.7326	0.5689	0.1169	0.5742	0.7173	0.1739	0.7219
0.1	0.99	25	8.8478	8.6848	8.6066	17.8612	17.7540	17.7687	9.6164	9.5005	9.4233	18.4376	18.2603	18.6518
0.1	0.99	50	4.4895	4.0847	4.6421	13.6919	12.6518	14.8426	4.5603	4.2099	4.8684	13.9417	12.7195	14.9260
0.1	0.99	200	1.5295	0.9980	1.5830	5.1397	4.0489	5.1831	1.4888	0.9877	1.4642	5.0859	3.9894	5.1096
0.1	0.99	500	0.8217	0.4058	0.8064	2.7056	1.7381	2.7153	0.8197	0.4015	0.8143	2.6693	1.6947	2.6829
0.9	0.7	25	6.1103	4.9676	6.1473	8.4275	7.6637	9.2523	2.4859	1.7595	2.5012	3.4293	2.3773	3.6602
0.9	0.7	50	4.7399	3.0516	4.9762	6.7138	4.7823	6.8760	1.1184	0.5191	1.1295	1.5233	0.7578	1.5055
0.9	0.7	200	3.3459	0.9632	3.3141	3.8581	1.3664	3.8058	0.6184	0.0919	0.6119	0.6777	0.1012	0.6678
0.9	0.7	500	2.8310	0.4346	2.8686	3.0251	0.5444	3.0638	0.5078	0.0337	0.5139	0.5203	0.0348	0.5243
0.9	0.9	25	8.1417	7.0509	7.9824	8.7725	8.2266	8.6818	3.2057	2.4752	3.1373	4.8959	3.6621	3.8607
0.9	0.9	50	5.8925	4.3174	6.3473	10.7335	8.7226	10.9916	1.3351	0.7439	1.3449	2.7640	1.8382	2.7305
0.9	0.9	200	3.6574	1.2369	3.6429	5.5532	2.8869	5.4815	0.6482	0.1701	0.6429	0.8962	0.2719	0.8898
0.9	0.9	500	2.9323	0.5506	2.9606	3.7890	1.2039	3.8206	0.5324	0.0760	0.5372	0.6017	0.0965	0.6054
0.9	0.99	25	22.1036	20.0173	21.1119	46.6178	44.1224	45.1159	8.0734	7.4880	8.1956	15.2070	14.1178	15.3034
0.9	0.99	50	13.3719	12.6760	14.6535	34.5166	34.2322	36.7291	2.7783	2.4334	3.0742	11.3101	9.8676	11.3810
0.9	0.99	200	6.5077	3.7701	6.3115	16.5426	13.1796	16.4576	0.9742	0.6064	1.0207	3.5843	2.4553	3.6355
0.9	0.99	500	4.1028	1.6444	4.0423	9.8100	6.8617	9.8207	0.6679	0.2472	0.6776	1.8210	0.9558	1.8298