3-24-2015

Comparison of Some Improved Estimators for Linear Regression Model under Different Conditions

Smit Shah
Florida International University, sshah060@fiu.edu

DOI: 10.25148/etd.FI15032146

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

COMPARISON OF SOME IMPROVED ESTIMATORS FOR LINEAR REGRESSION MODEL UNDER DIFFERENT CONDITIONS

A thesis submitted in partial fulfillment of

the requirements for the degree of

MASTER OF SCIENCE

in

STATISTICS

by

Smit Nailesh Shah

2015

To: Dean Michael R. Heithaus
College of Arts and Sciences

This thesis, written by Smit Nailesh Shah, and entitled Comparison of some Improved Estimators for Linear Regression Model under Different Conditions, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this thesis and recommend that it be approved.

Wensong Wu

Florence George, Co-Major Professor

B. M. Golam Kibria, Co-Major Professor

Date of Defense: March 24, 2015

The thesis of Smit Nailesh Shah is approved.

Dean Michael R. Heithaus
College of Arts and Sciences

Dean Lakshmi N. Reddi
University Graduate School

Florida International University, 2015

ABSTRACT OF THE THESIS

COMPARISON OF SOME IMPROVED ESTIMATORS FOR LINEAR REGRESSION MODEL UNDER DIFFERENT CONDITIONS

by

Smit Nailesh Shah

Florida International University, 2015

Miami, Florida

Professor B. M. Golam Kibria, Co-Major Professor

Professor Florence George, Co-Major Professor

Multiple linear regression model plays a key role in statistical inference and it has extensive applications in business, environmental, physical and social sciences. Multicollinearity has been a considerable problem in multiple regression analysis. When the regressor variables are multicollinear, it becomes difficult to make precise statistical inferences about the regression coefficients. There are some statistical methods that can be used, which are discussed in this thesis are ridge regression, Liu, two parameter biased and LASSO estimators. Firstly, an analytical comparison on the basis of risk was made among ridge, Liu and LASSO estimators under orthonormal regression model. I found that LASSO dominates least squares, ridge and Liu estimators over a significant portion of the parameter space for large dimension. Secondly, a simulation study was conducted to compare performance of ridge, Liu and two parameter biased estimator by their mean squared error criterion. I found that two parameter biased estimator performs better than its corresponding ridge regression estimator. Overall, Liu estimator performs better than both ridge and two parameter biased estimator.

TABLE OF CONTENTS

I. INTRODUCTION

Regression is a statistical technique for determining relationship between variables, this relationship is formulated by a statistical equation. This statistical equation allows us to predict the values of dependent variable on the basis of fixed values of one or more independent variables(or regressors or predictors), which is called regression equation or prediction model and the technique is called regression analysis. Along with the dependent variable and known independent variables, a regression equation also contains unknown regression coefficients. The main goal of a regression analysis is to appropriately estimate the values of regression coefficients and fit a good model. Regression analysis is used in almost all fields including psychology, economics, engineering, management, biology and sociology (for examples, see Mansson and Kibria (2012), Liu (2003)). Sir Francis Galton first introduced regression analysis in 1880s in his studies of hereditary and eugenics. A regression equation with a degree of one is called linear regression equation. The simplest form of linear regression is with one dependent and only one independent variable and it is called the simple linear regression model. Usually, the dependent variable is explained by more than one variable, and we use multiple linear regression model. The standard multiple linear regression model is expressed as

$$y = \mathbf{X}\beta + \varepsilon, \tag{1.1}$$

where y is a nx1 vector of response variable, $\mathbf{X}$ is a design matrix of order nxp, β is a px1 vector of regression coefficients and ε is a nx1 vector of random error, which is normally distributed with mean vector 0 and variance $\sigma^2\mathbf{I}_n$. Here $\mathbf{I}_n$ is identity matrix of order n. The least square estimator (LSE) of β is a linear function of y and is defined as

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'y \tag{1.2}$$

and the covariance matrix of $\hat{\beta}$ is obtained as

$$cov\left(\hat{\beta}\right) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}. \tag{1.3}$$

It is noted that the least squares estimator is unbiased and has a minimum variance. Naturally, we deal with data where the variables may or may not be independent, thus making the $\mathbf{X}'\mathbf{X}$ matrix ill-conditioned (that is, near linear dependency among various columns of $\mathbf{X}'\mathbf{X}$). We see from equations (1.2) and (1.3) that the LSE and its variance-covariance matrix heavily depend on the property of $\mathbf{X}'\mathbf{X}$ matrix. The dependence of the columns of **X** matrix leads to the problem of multicollinearity and produce a number of errors in estimating β which affects the reliability of the statistical inference.

To overcome this multicollinearity problem, Hoerl and Kennard (1970) introduced a new kind of estimator, the ridge regression estimator, where they proposed to add a small positive number to the diagonal elements of the $\mathbf{X}'\mathbf{X}$ matrix. The ridge regression estimator proposed by Hoerl and Kennard is given by

$$\hat{\beta}_k = (\mathbf{X}'\mathbf{X} + k\,\mathbf{I}_{\boldsymbol{p}})^{-1}\mathbf{X}'y, \qquad k \geq 0. \tag{1.4}$$

For a small positive value of *k*, this estimator provides a smaller mean squared error (MSE) compared to the LSE. The constant *k* is called the ridge or biased parameter. Literature reveals a lot of discussion related to estimating a good estimator of *k*, which is to be estimated from the real data. The estimation of *k* are discussed by Hoerl and Kennard (1970), Golub et al. (1979), Kibria (2003), Saleh (2006), Muniz and Kibria (2009), Dorugade (2013), Aslam (2014), Hefnawy and Farag (2014), and very recently Kibria and Banik (2015) among others.

Motivated by the interpretation of the ridge estimator, Liu (1993), to combat the multicollinearity problem proposed a new class of biased estimate, the Liu estimator, defined as

$$\hat{\beta}_d = \left(\mathbf{X}'\mathbf{X} + \mathbf{I}_{\boldsymbol{p}}\right)^{-1}\left(\mathbf{X}'y + d\,\hat{\beta}\right), \qquad 0 < d < 1. \qquad (1.5)$$

For any value of d, this estimator provides a smaller mean squared error compared to the least square estimator. The constant *d* is called the shrinkage parameter. The advantage of the Liu estimator over the ridge estimator, which is a complex function of *k*, is that $\hat{\beta}_d$ is a linear function of *d* and so it is convenient.

Hoerl and Kennard (1970) suggested that the appropriate range of *k* is between 0 to 1, but in application the chosen *k* may not be large enough to correct the ill conditioning problem, especially when $\mathbf{X}'\mathbf{X}$ is severely ill conditioned. In this case, the small *k* may not be able to reduce the condition number of $\mathbf{X}'\mathbf{X} + k\,\mathbf{I}_{\boldsymbol{p}}$ to proper extent, thus the resulting ridge regression may still remain unstable. This reason of instability motivated Liu (2003) to propose a new two parameter biased estimator which is defined as

$$\hat{\beta}_{k,d} = \left(\mathbf{X}'\mathbf{X} + k\,\mathbf{I}_{\boldsymbol{p}}\right)^{-1}(\mathbf{X}'y - d\,\tilde{\beta}), \qquad k > 0,\ -\infty < d < \infty. \qquad (1.6)$$

where $\tilde{\beta}$ can be any estimator of *β*. $\hat{\beta}_{k,d}$ is generalization of $\hat{\beta}_d = \left(\mathbf{X}'\mathbf{X} + \mathbf{I}_{\boldsymbol{p}}\right)^{-1}(\mathbf{X}'y + d\,\hat{\beta})$ when $\tilde{\beta} = \hat{\beta}_{LS}$, which is the Liu estimator.

When $\tilde{\beta} = \hat{\beta}_R$,$\hat{\beta}_{k,d} = (\mathbf{X}'\mathbf{X} + k\,\mathbf{I}_{\boldsymbol{p}})^{-1}\mathbf{X}'y - d(\mathbf{X}'\mathbf{X} + k\,\mathbf{I}_{\boldsymbol{p}})^{-2}\mathbf{X}'y$, the estimator can fully address the ill conditioning problem. For any *k* > 0, we can always find a value of *d* so that the mean squared error provided by this estimator is less than or equal to that provided by ridge estimator.

I will briefly discuss about the above three estimators in the latter part of the thesis, where I compare them under multicollinear regression model and error assuming a normal distribution. The comparison will be made using optimum value of *d* proposed by Liu (2003) and few suggested *k*s from the literature.

The least square estimator, ridge regression estimator and Liu estimator were not considered satisfactory because, least square estimates have large variance hence less prediction accuracy. Also, with large number of predictors we would like to determine smaller subsets that has the strongest effects and thus produce easily interpretable models. On the other hand ridge regression and Liu estimators are continuous process that shrink coefficients and thus are more stable; however the problem of interpreting model with large predictors still remain unsolved as they do not set any of the coefficients to 0.Tibshirani (1996) proposed a new technique, called the LASSO, for 'least absolute shrinkage and selection operator'. It minimizes the residual sum of squares subject to the sum of the absolute value of the coefficients being less than a constant. Because of this nature of the constraint it shrinks some coefficients and tends to set others to exactly 0, thus retaining "selection" a good feature of subset selection method and "shrinking of coefficients" a good feature of ridge regression and Liu estimator. Because of these good features the LASSO gives interpretable models.

Suppose $x^i = (x_{i1}, . . ., x_{ip})'$, i = 1,2,...,n are the predictor variables and y_i are responses. I assume that the x_{ij} are standardized so that $\sum_i x_{ij} / n = 0$, $\sum_i x_{ij}^2 / n = 1$.

Letting $\hat{\beta} = (\hat{\beta}_1, \dots, \hat{\beta}_p)'$, the LASSO estimate $(\hat{\alpha}, \hat{\beta})$ is obtained as follow,

$$(\hat{\alpha}, \hat{\beta}) = \arg\min \left\{\sum_{i=1}^{n}\left(y_i - \alpha - \sum_j \beta_j x_{ij}\right)^2\right\}, \quad \text{subject to } \sum_j |\beta_j| \leq t. \quad (1.7)$$

Here $t \geq 0$ is a tuning parameter. For all t, the solution for α is $\hat{\alpha} = \bar{y}$. I assume without loss of generality that $\bar{y} = 0$ and hence omit α.

The purpose of this research is to investigate the least square estimator, ridge regression estimator, Liu estimator and LASSO estimator and make an analytical comparison amongst them. This analytical comparison will be made under orthonormal regression model and based on the smallest mean squared error or risk and efficiency over least square estimator.

The organization of the thesis is as follows: The risk functions of the proposed estimators under the orthonormal model is given in Chapter II. Chapter III contains details of analysis of risks and efficiencies of the estimators with the tables and graphs. In Chapter IV, I reviewed some estimators of *k* and *d* and use Monte Carlo simulation to evaluate the performance of all estimators. Finally some concluding remarks are given in Chapter V.

II. STATISTICAL METHODOLOGY

To make an analytical comparison, I have expressed all risk functions under the orthogonal regression model in this chapter. It is noted that we are restricted to compare the performance of the estimators under the orthonormal regression model as the risk of LASSO is available under the orthonormal regression model.

2.1. Regression models in orthogonal form and their MSEs

From (1.1) we have the multiple linear regression model as,

$$y = \mathbf{X}\beta + \varepsilon.$$

Suppose, there exists an orthogonal matrix **Q** whose columns constitute the eigen vectors of **X′X**, then **Q′X′XQ** = **Λ** = diag(λ_1, λ_2,..., λ_p), where $\lambda_1 \geq \lambda_2 \geq . . . \geq \lambda_p > 0$ are ordered eigenvalues of **X′X**. Thus the canonical form of (1.1) is

$$y = \mathbf{X}^{*}\alpha + \varepsilon, \tag{2.1}$$

where $\mathbf{X}^*$ = **XQ** and α = **Q′**β. Here the least square estimate is given as

$$\hat{\alpha}_{ls} = \mathbf{\Lambda}^{-1}\mathbf{Q}'\mathrm{y}. \tag{2.2}$$

The ridge regression approach replaces **X′X** with **X′X** + k**I**, which is same as replacing λ_i with λ_i+k. Then the generalized ridge regression estimators of α are given as

$$\hat{\alpha}_k = (\mathbf{\Lambda} + \mathbf{K}\,\mathbf{I})^{-1}\mathbf{Q}'\mathrm{y}, \tag{2.3}$$

where, **K** = diag(k_1, k_2,…, k_p), $k_i > 0$. The relationship between both models is as $\hat{\beta}_k = \mathbf{Q}\hat{\alpha}_k$. Now, MSE ($\hat{\beta}_k$) = MSE ($\hat{\alpha}_k$). MSE($\hat{\alpha}_k$) is obtained as,

$$\text{MSE}\ (\hat{\alpha}_k) = \sigma^2 \sum_{i=1}^{p} \frac{\lambda_i}{(\lambda_i+k_i)^2} + \sum_{i=1}^{p} \frac{k_i^2\lambda_i^2}{(\lambda_i+k_i)^2}, \quad k > 0. \tag{2.4}$$

The Liu estimator of α is given as

$$\hat{\alpha}_d = (\mathbf{\Lambda} + \mathbf{I})^{-1}(\mathbf{Q}'\mathrm{y} + d\,\hat{\alpha}). \quad (2.5)$$

The relationship between the estimators under the linear regression model and orthogonal model is as follows:

$$\hat{\beta}_d = \mathbf{Q}\hat{\alpha}_d.$$

The MSE of $\hat{\alpha}_d$ is obtained as,

$$\text{MSE } (\hat{\alpha}_d) = \sigma^2 \sum_{i=1}^{p} \frac{(\lambda_i + d)^2}{\lambda_i(\lambda_i+1)^2} + (d-1)^2 \sum_{i=1}^{p} \frac{\alpha_i^2}{(\lambda_i+1)^2}, \quad d > 0. \quad (2.6)$$

Equations (2.4) and (2.6) provides the mean squared error (MSE) of the ridge estimator and Liu estimator respectively. The MSEs are combination of their corresponding variance of the estimator and the bias in the estimator. In (2.4) the first term on right side is the sum of variances of the parameters in $\hat{\beta}_k$ and the second term is the square of the bias in $\hat{\beta}_k$. Similarly in (2.6) the first term on right side is the sum of variances of the parameters in $\hat{\beta}_d$ and the second term is the square of the bias in $\hat{\beta}_d$.

For LASSO estimator, let us consider the canonical form with full least square estimate, orthogonal regressors and normal errors with known variance. Let **X** be *nxp* design matrix with *ij*th entry x_{ij} and $\mathbf{X}'\mathbf{X} = \mathbf{I}_p$.

The LASSO estimator equals,

$$\hat{\beta}^L = (t(\hat{\beta}_1), t(\hat{\beta}_2), \ldots, t(\hat{\beta}_p))', \quad (2.7)$$

where t(x) = sign(x) (|x| - λ)$_+$ which is exactly same as soft shrinkage proposals of Donoho and Johnstone (1994). Here, λ is the tuning parameter.

For any estimator $\bar{\beta}$ of β, one may define the normalized mean squared error or risk as,

$$\mathrm{R}(\bar{\beta}, \beta) = \frac{E(\bar{\beta}-\beta)'(\bar{\beta}-\beta)}{p\sigma_n^2}\,. \tag{2.8}$$

Now taking the LASSO estimator, For $\lambda = \sigma_n\sqrt{2\ln(p)}$, its risk satisfies the bound

$$\mathrm{R}(\hat{\beta}^L, \beta) \;\leq\; \sigma^2(1 + 2\ \mathrm{ln(p)})\left(\frac{1}{p} + c_n^L(\beta)\right), \tag{2.9}$$

where, $c_n^L(\beta) = \frac{1}{p}\sum_{j=1}^{p} min\left(\frac{\beta_j^2}{\sigma_n^2}, 1\right)$.

For the case when some coefficients are non-zero and some are zero. In particular, suppose $q < p$ coefficients satisfy ${\beta_j}^2 \geq \sigma_n^2$ and remaining equal zero. Then

$c_n^L(\beta) = q/p$ so the bound in (2.9) is

$$\mathrm{R}(\hat{\beta}^L, \beta) \;\leq\; \sigma^2(1 + 2\ \mathrm{ln(p)})\left(\frac{1+q}{p}\right). \tag{2.10}$$

which approaches zero as $p \to \infty$ with q fixed. More details on this see Donoho and Johnstone (1994).

2.2. Risk functions of Estimators

Let $(\hat{\theta} - \theta)^2$ be the quadratic loss function or squared error loss function, then $\mathrm{E}(\hat{\theta} - \theta)^2$ is termed as the risk function of the estimator, which in fact is the mean square error (MSE) of estimator $\hat{\theta}$ of a parameter θ. In this section I present the risk functions of ridge regression estimator and Liu estimator.

The risk function of LSE can be obtained as,

We know, MSE ($\hat{\beta}$) = E(($\hat{\beta} - \beta$)($\hat{\beta} - \beta$)') = $\sigma^2(\mathbf{X}'\mathbf{X})^{-1}$ from (1.3)

Risk ($\hat{\beta}$) = $\sigma^2\ tr((\mathbf{X}'\mathbf{X})^{-1})$

Let $\mathbf{X'X} = \mathbf{I}_p$

Risk $(\hat{\beta}) = \sigma^2\ tr(\mathbf{I}_p)$

$= \sigma^2 p.$ (2.11)

2.2.1. Risk function of ridge regression estimator

From (1.4) we have the ridge regression estimator as,

$\hat{\beta}_k = (\mathbf{X'X} + k\,\mathbf{I}_p)^{-1}\mathbf{X}'y$

$= (\mathbf{I}_p + k(\mathbf{X'X})^{-1})^{-1}(\mathbf{X'X})^{-1}\mathbf{X}'y$

Let $\mathbf{W} = (\mathbf{I}_p + k(\mathbf{X'X})^{-1})^{-1}$

$\hat{\beta}_k = \mathbf{W}\hat{\beta}$

MSE$(\hat{\beta}_k) = E\ (\hat{\beta}_k - \beta)(\hat{\beta}_k - \beta)'$

$\hat{\beta}_k - \beta = \mathbf{W}\hat{\beta} - \mathbf{W}\beta + \mathbf{W}\beta - \beta$

$= \mathbf{W}(\hat{\beta} - \beta) + (\mathbf{W} - \mathbf{I}_p)\beta$

MSE$(\hat{\beta}_k) = E\left\{\left(\mathbf{W}(\hat{\beta} - \beta) + (\mathbf{W} - \mathbf{I}_p)\beta\right)\left(\mathbf{W}(\hat{\beta} - \beta) + (\mathbf{W} - \mathbf{I}_p)\beta\right)'\right\}$

$= \mathbf{W}\left(E(\hat{\beta} - \beta)(\hat{\beta} - \beta)'\right)\mathbf{W}' + E(\mathbf{W} - \mathbf{I}_p)\boldsymbol{\beta}\boldsymbol{\beta}'(\mathbf{W} - \mathbf{I}_p)'$

$= \sigma^2\mathbf{W}(\mathbf{X'X})^{-1}\mathbf{W}' + (\mathbf{W} - \mathbf{I}_p)\boldsymbol{\beta}\boldsymbol{\beta}'(\mathbf{W} - \mathbf{I}_p)'$

Risk$(\hat{\beta}_k) = \sigma^2\ tr(\mathbf{W}(\mathbf{X'X})^{-\mathbf{1}}\mathbf{W}') + \beta'(\mathbf{W} - \mathbf{I}_p)^2\beta$

Let $\mathbf{X'X} = \mathbf{I}_p \Rightarrow \mathbf{W} = (\mathbf{I}_p + k\,\mathbf{I}_p)^{-1} = \frac{1}{(1+k)}\mathbf{I}_p$

$$= \sigma^2\, tr\left(\frac{1}{(1+k)}\mathbf{I_p I_p I_p}\frac{1}{(1+k)}\right) + \beta'\left(\frac{1}{(1+k)}\mathbf{I_p} - \mathbf{I_p}\right)^2 \beta$$

$$= \frac{\sigma^2\, p}{(1+k)^2} + \frac{k^2 \beta' \beta}{(1+k)^2}$$

taking, $\Delta^2 = \beta'\beta / \sigma^2$

$$= \frac{\sigma^2}{(1+k)^2}[p + k^2\Delta^2], \qquad k > 0,\ \Delta^2 \geq 0. \tag{2.12}$$

where, Δ^2 is defined as the divergence parameter. It is the sum of squares of the normalized coefficients.

2.2.2. Risk function of Liu estimator

From (1.5) we have Liu estimator as,

$$\hat{\beta}_d = \left(\mathbf{X'X} + \mathbf{I_p}\right)^{-1}(\mathbf{X'}y + d\,\hat{\beta})$$

$$= \left(\mathbf{X'X} + \mathbf{I_p}\right)^{-1}\left(\mathbf{X'X} + d\,\mathbf{I_p}\right)\hat{\beta}$$

Let **F** = $\left(\mathbf{X'X} + \mathbf{I_p}\right)^{-1}\left(\mathbf{X'X} + d\,\mathbf{I_p}\right)$, then

$$\hat{\beta}_d = \mathbf{F}\hat{\beta}$$

MSE($\hat{\beta}_d$) = $E\,(\hat{\beta}_d - \beta)(\hat{\beta}_d - \beta)'$

$$\hat{\beta}_d - \beta = \mathbf{F}\hat{\beta} - \mathbf{F}\beta + \mathbf{F}\beta - \beta$$

$$= \mathbf{F}\left(\hat{\beta} - \beta\right) + \left(\mathbf{F} - \mathbf{I_p}\right)\beta$$

MSE($\hat{\beta}_d$) = $E\left\{\left(\mathbf{F}\left(\hat{\beta} - \beta\right) + \left(\mathbf{F} - \mathbf{I_p}\right)\beta\right)\left(\mathbf{F}\left(\hat{\beta} - \beta\right) + \left(\mathbf{F} - \mathbf{I_p}\right)\beta\right)'\right\}$

$$= \mathbf{F}\left(E\left(\hat{\beta} - \beta\right)\left(\hat{\beta} - \beta\right)'\right)\mathbf{F}' + E\left(\mathbf{F} - \mathbf{I_p}\right)\boldsymbol{\beta\beta}'\left(\mathbf{F} - \mathbf{I_p}\right)'$$

$$= \sigma^2 \mathbf{F}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{F}' + \left(\mathbf{F} - \mathbf{I}_p\right)\boldsymbol{\beta}\boldsymbol{\beta}'(\mathbf{F} - \mathbf{I}_p)'$$

Risk($\hat{\beta}_d$) $= \sigma^2\, tr(\mathbf{F}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{F}') + \beta'\left(\mathbf{F} - \mathbf{I}_p\right)^2 \beta$

Let **X'X** = $\mathbf{I}_p \quad \Rightarrow$ **F** = $\left(\mathbf{I}_p + \mathbf{I}_p\right)^{-1}\left(\mathbf{I}_p + d\, \mathbf{I}_p\right) \quad = \frac{(1+d)}{2}\mathbf{I}_p$

$$= \sigma^2\, tr\left(\frac{(1+d)}{2}\mathbf{I}_p\mathbf{I}_p\mathbf{I}_p\frac{(1+d)}{2}\right) + \beta'\left(\frac{(1+d)}{2}\mathbf{I}_p - \mathbf{I}_p\right)^2 \beta$$

$$= \frac{\sigma^2(1+d)^2 p}{4} + \frac{(d-1)^2\beta'\beta}{4}$$

taking, $\Delta^2 = \beta'\beta / \sigma^2$

$$\text{Risk}(\hat{\beta}_d) = \frac{\sigma^2}{4}\left[(1+d)^2 p + (d-1)^2\Delta^2\right], \quad d > 0,\ \Delta^2 \geq 0. \qquad (2.13)$$

where, Δ^2 is defined as the divergence parameter.

III. ANALYSIS OF DOMINANCE PROPERTIES OF THE ESTIMATORS

In this chapter, I consider the risks and relative efficiencies comparison of various estimators using the risk functions from (2.10), (2.11), (2.12) and (2.13). The relative efficiencies of each estimator is compared with LSE, which is simply the ratio of risk of LSE to risk of corresponding estimator. I computed risks and provided them as tabular form in Tables 3.1-3.8 (for fixed *p* and different values of Δ^2) and graphically presented in Figures 3.1-3.8 for different *p*. The efficiencies are provided as tabular form in Tables 3.9-3.12 for *p* = 3, 5, 7 and 10 with graphical presentation in Figures 3.9-3.12.

3.1. Comparison of LASSO with Least Square Estimator.

The risk of LASSO will be less than that of LSE of β when,

$$R(\hat{\beta}^{L}) - R(\hat{\beta}) < 0$$

$$\sigma^2(1 + 2\ln(p))\left(\frac{1+q}{p}\right) - p\sigma^2 < 0$$

$$q < \frac{p^2}{(1+2\ln(p))} - 1 \tag{3.1}$$

Thus for all *q* satisfying (3.1), the risk of LASSO will be less than that of LSE.

3.2. Comparison of LASSO with Ridge regression Estimator.

For fixed *k* and *q*, the risk of LASSO to be less than that of ridge regression estimator when,

$$R(\hat{\beta}^{L}) - R(\hat{\beta}_k) < 0$$

$$\sigma^2(1 + 2\ln(p))\left(\frac{1+q}{p}\right) - \frac{\sigma^2}{(1+k)^2}[p + k^2\Delta^2] < 0$$

$$\Delta^2 > \frac{(1+2\ln(p))(1+q)(1+k)^2 - p^2}{pk^2}. \tag{3.2}$$

For all Δ^2 satisfying (3.2) risk of LASSO will be less than that of ridge regression estimator. Otherwise, ridge regression estimator will have smaller risk than that of LASSO.

3.3. Comparison of LASSO with Liu Estimator.

For fixed *k* and *q,* the risk of LASSO to be less than that of Liu estimator when,

$$R(\hat{\beta}^L) - R(\hat{\beta}_d) < 0$$

$$\sigma^2(1 + 2\ln(p))\left(\frac{1+q}{p}\right) - \frac{\sigma^2}{4}[(1+d)^2 p + (d-1)^2\Delta^2] < 0$$

$$\Delta^2 > \frac{4(1+2\ln(p))(1+q) - ((1+d)p)^2}{(d-1)^2} \quad (3.3)$$

Otherwise, Liu will dominate LASSO estimator.

3.4. Comparison of Liu with Ridge regression Estimator.

The risk of Liu estimator to be less than that of ridge regression estimator when,

$$R(\hat{\beta}_d) - R(\hat{\beta}_k) < 0$$

$$\frac{\sigma^2}{4}[(1+d)^2 p + (d-1)^2\Delta^2] - \frac{\sigma^2}{(1+k)^2}[p + k^2\Delta^2] < 0$$

When *k=d*

$$\Delta^2 < \frac{p[4-(1+k)^4]}{[(k^2-1)^2 - 4k^2]} \quad (3.4)$$

Otherwise ridge will dominate Liu estimator.

3.5. Comparison of Liu with Least Square Estimator.

The risk of Liu estimator will be less than that of the least square estimator when,

$$R(\hat{\beta}_d) - R(\hat{\beta}) < 0$$

$$\frac{\sigma^2}{4}[(1+d)^2 p \ + (d-1)^2\Delta^2] - p\sigma^2 < 0$$

$$\Delta^2 < \frac{p[4-(1+d)^2]}{(d-1)^2} \qquad (3.5)$$

For all values of Δ^2 satisfying (3.5), Liu estimator dominates the least square estimator.

3.6. Comparison of Ridge regression with Least Square Estimator.

The risk of ridge regression estimator will be less than that of the least square estimator when,

$$R(\hat{\beta}_k) - R(\hat{\beta}) < 0$$

$$\frac{\sigma^2}{(1+k)^2}[p + \ k^2\Delta^2] - \ p\sigma^2 < 0$$

$$\Delta^2 < \frac{p(2+k)}{k} \qquad (3.6)$$

Otherwise, LSE dominates the ridge regression estimator.

The risks and relative efficiencies of LASSO, ridge regression, Liu and LS estimators for different values of *k, d* and *q* and for *p* = 3, 4, 5, 6, 7, 8, 9, 10 are presented in Tables 3.1 -3.12 respectively. These tables are in support of the comparison among all the estimators. See also the Figures 3.1-3.12 in this respect. In these figures three different values of *k* and *d* (0.1, 0.5 and 0.9) are considered and the risk line corresponding to the particular value of estimator is denoted by estimator followed by its value (e.g. k0.1 for when k = 0.1 and d0.5 for when d = 0.5). We know $q < p$, but q = 0 provides a model with no explanatory variables thus we do not consider the value 0 for q. In the figures, for a value of q the risk line is denoted by la followed by the value of q (e.g. la3 for when q = 3 and la for LASSO).

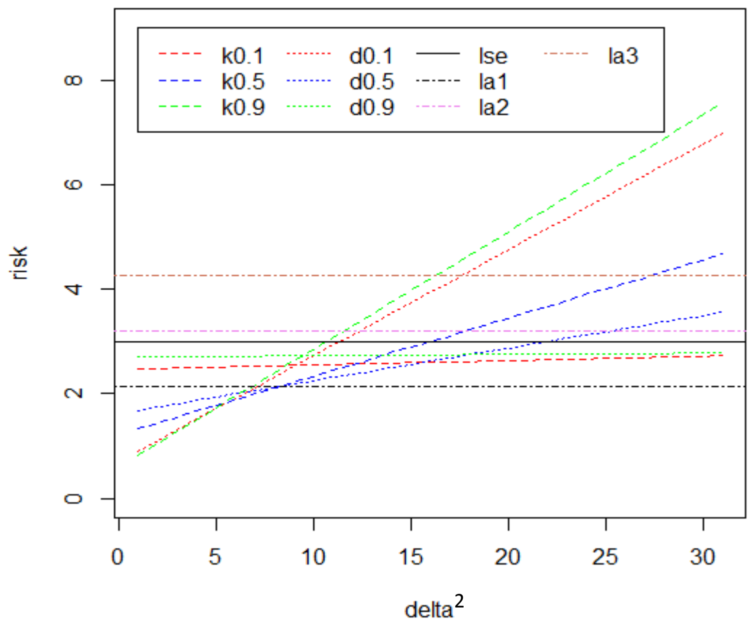

Figure 3.1: Risks of all estimators as a function of Δ^2 for p=3

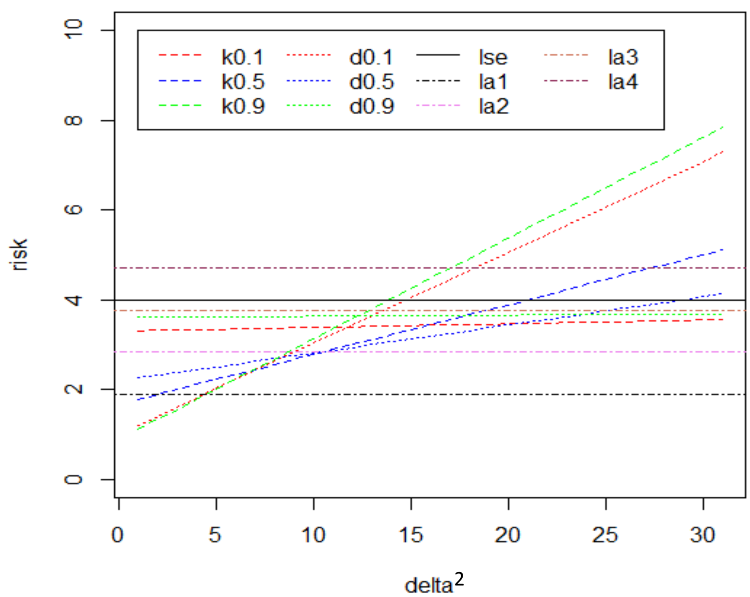

Figure 3.2: Risks of all estimators as a function of Δ^2 for p=4

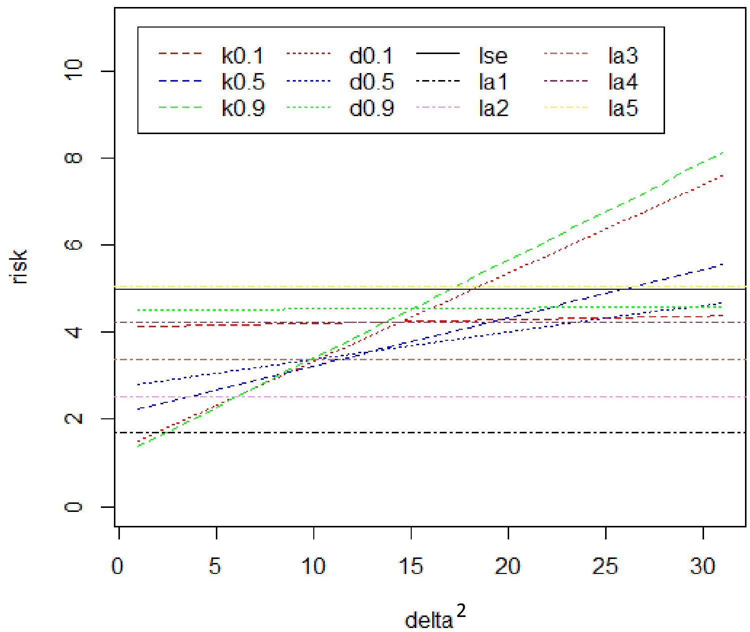

Figure 3.3: Risks of all estimators as a function of Δ^2 for p=5

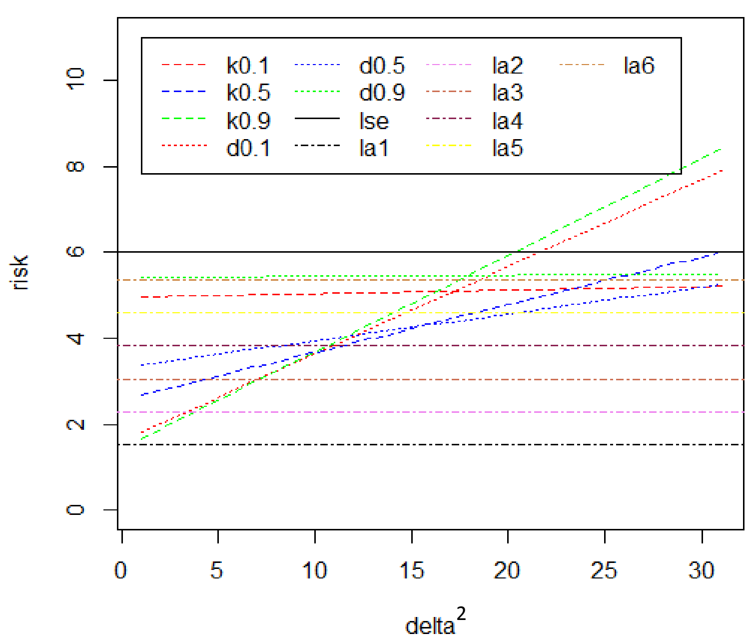

Figure 3.4: Risks of all estimators as a function of Δ^2 for p=6

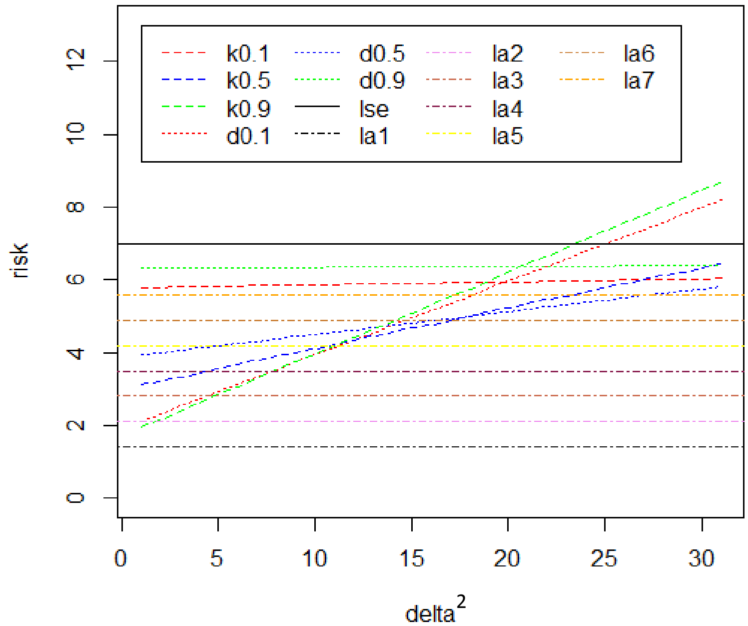

Figure 3.5: Risks of all estimators as a function of Δ^2 for p=7

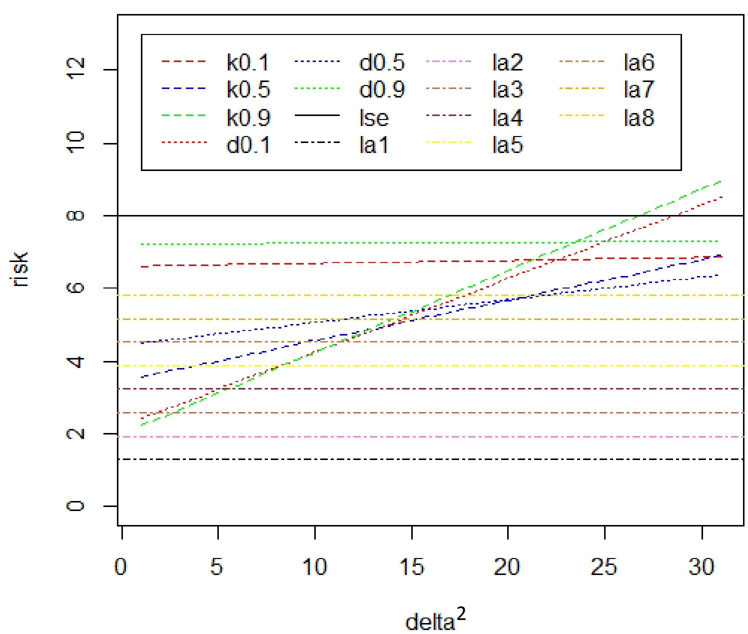

Figure 3.6: Risks of all estimators as a function of Δ^2 for p=8

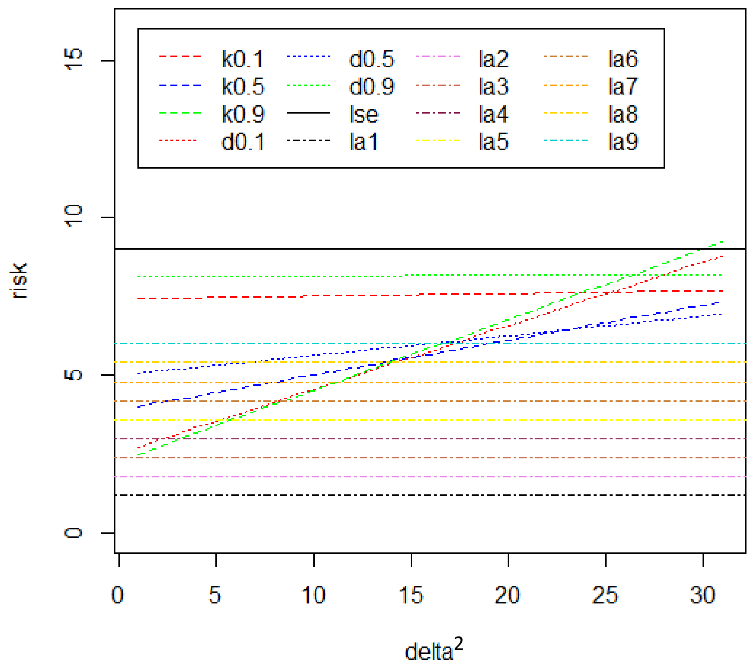

Figure 3.7: Risks of all estimators as a function of Δ^2 for p=9

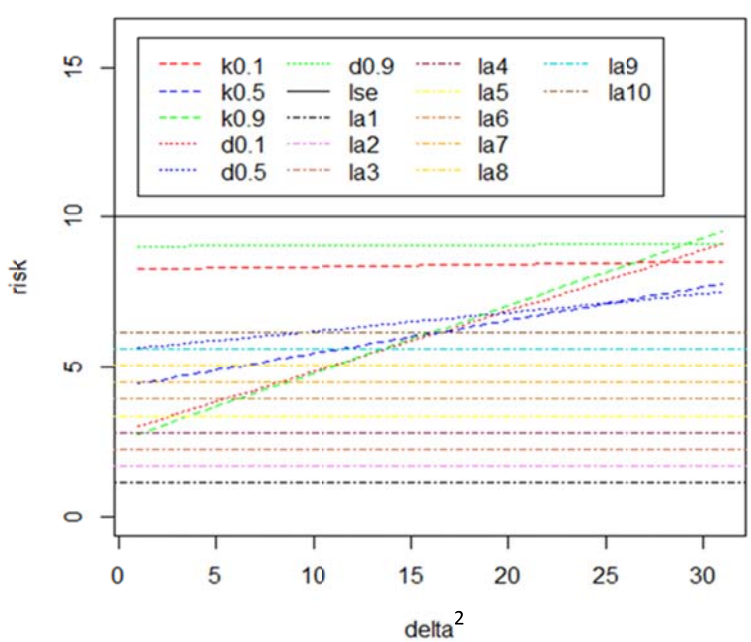

Figure 3.8: Risks of all estimators as a function of Δ^2 for p=10

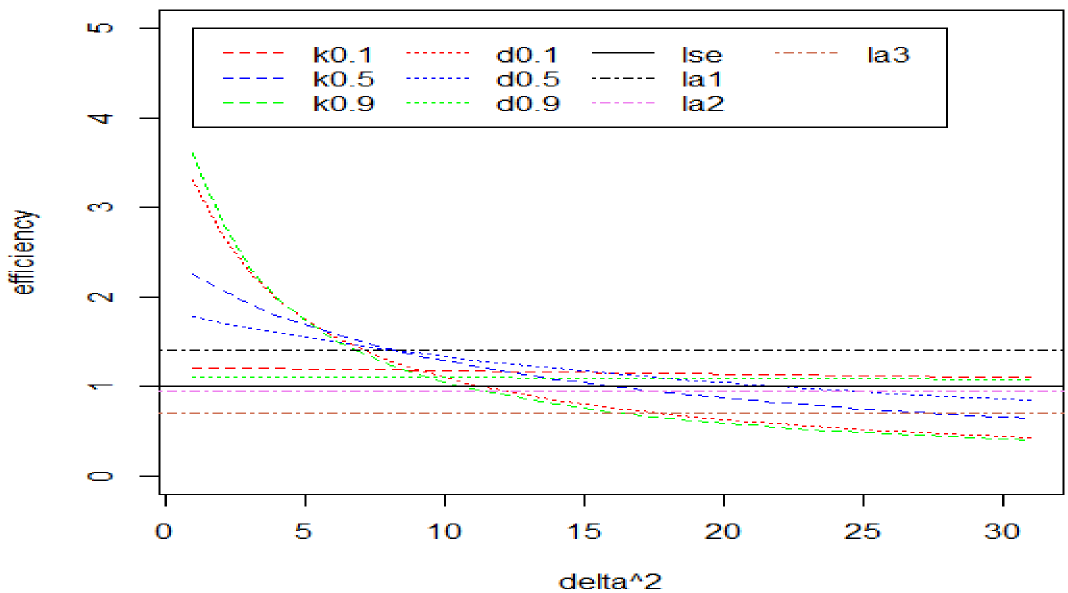

Figure 3.9: Efficiency of all estimators as a function of Δ^2 for p=3

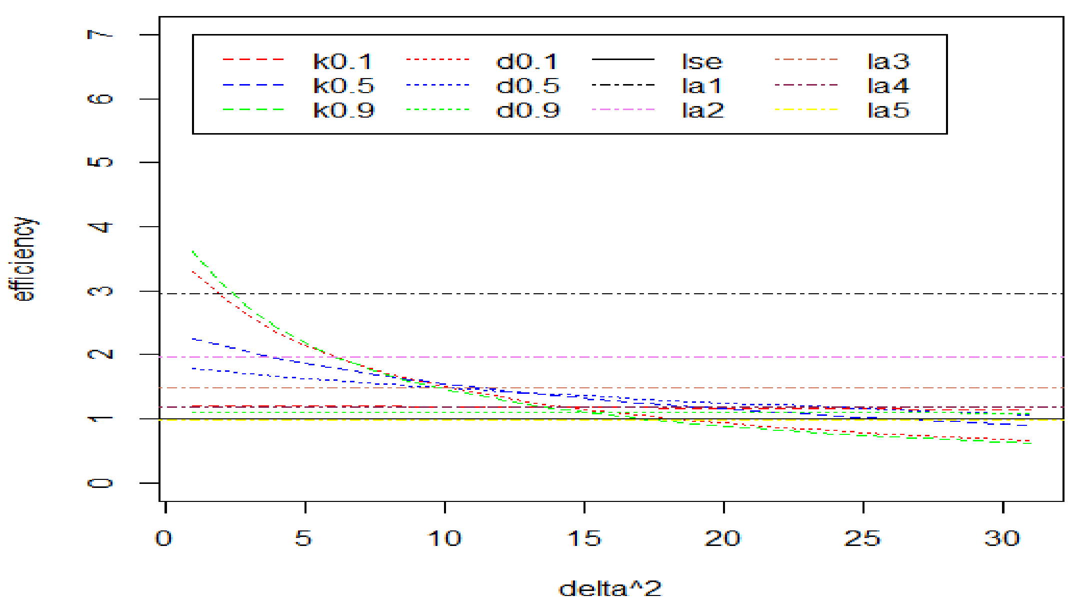

Figure 3.10: Efficiency of all estimators as a function of Δ^2 for p=5

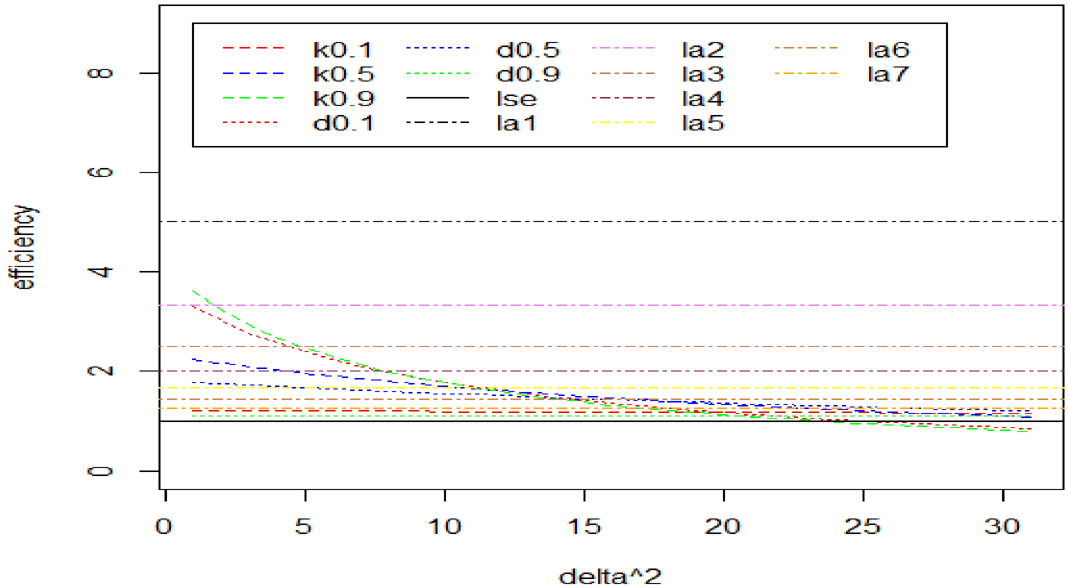

Figure 3.11: Efficiency of all estimators as a function of Δ^2 for p=7

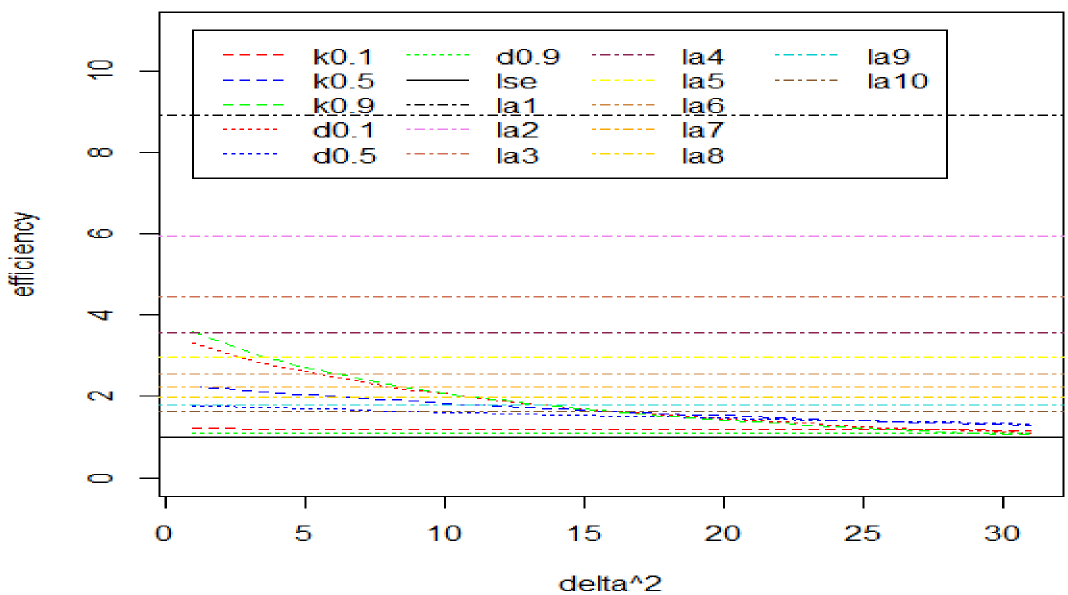

Figure 3.12: Efficiency of all estimators as a function of Δ^2 for p=10

Table 3.1: Risk for different values of Δ^2 at p=3 and k=d=0.1, 0.5 and 0.9

Δ^2	LSE	LASSO			Ridge			Liu		
		q=1	*q=2*	*q=3*	*k=0.1*	*k=0.5*	*k=0.9*	*d=0.1*	*d=0.5*	*d=0.9*
0.00	3.00	2.13	3.20	4.26	2.48	1.33	0.83	0.91	1.69	2.71
1.00	3.00	2.13	3.20	4.26	2.49	1.44	1.06	1.11	1.75	2.71
2.00	3.00	2.13	3.20	4.26	2.50	1.56	1.28	1.31	1.81	2.71
3.00	3.00	2.13	3.20	4.26	2.50	1.67	1.50	1.52	1.88	2.72
4.00	3.00	2.13	3.20	4.26	2.51	1.78	1.73	1.72	1.94	2.72
5.00	3.00	2.13	3.20	4.26	2.52	1.89	1.95	1.92	2.00	2.72
6.00	3.00	2.13	3.20	4.26	2.53	2.00	2.18	2.12	2.06	2.72
7.00	3.00	2.13	3.20	4.26	2.54	2.11	2.40	2.33	2.13	2.73
8.00	3.00	2.13	3.20	4.26	2.55	2.22	2.63	2.53	2.19	2.73
9.00	3.00	2.13	3.20	4.26	2.55	2.33	2.85	2.73	2.25	2.73
10.00	3.00	2.13	3.20	4.26	2.56	2.44	3.07	2.93	2.31	2.73
11.00	3.00	2.13	3.20	4.26	2.57	2.56	3.30	3.14	2.38	2.74
12.00	3.00	2.13	3.20	4.26	2.58	2.67	3.52	3.34	2.44	2.74
13.00	3.00	2.13	3.20	4.26	2.59	2.78	3.75	3.54	2.50	2.74
14.00	3.00	2.13	3.20	4.26	2.60	2.89	3.97	3.74	2.56	2.74
15.00	3.00	2.13	3.20	4.26	2.60	3.00	4.20	3.95	2.63	2.75
16.00	3.00	2.13	3.20	4.26	2.61	3.11	4.42	4.15	2.69	2.75
17.00	3.00	2.13	3.20	4.26	2.62	3.22	4.65	4.35	2.75	2.75
18.00	3.00	2.13	3.20	4.26	2.63	3.33	4.87	4.55	2.81	2.75
19.00	3.00	2.13	3.20	4.26	2.64	3.44	5.09	4.76	2.88	2.76
20.00	3.00	2.13	3.20	4.26	2.64	3.56	5.32	4.96	2.94	2.76
21.00	3.00	2.13	3.20	4.26	2.65	3.67	5.54	5.16	3.00	2.76
22.00	3.00	2.13	3.20	4.26	2.66	3.78	5.77	5.36	3.06	2.76
23.00	3.00	2.13	3.20	4.26	2.67	3.89	5.99	5.57	3.13	2.77
24.00	3.00	2.13	3.20	4.26	2.68	4.00	6.22	5.77	3.19	2.77
25.00	3.00	2.13	3.20	4.26	2.69	4.11	6.44	5.97	3.25	2.77
26.00	3.00	2.13	3.20	4.26	2.69	4.22	6.66	6.17	3.31	2.77
27.00	3.00	2.13	3.20	4.26	2.70	4.33	6.89	6.38	3.38	2.78
28.00	3.00	2.13	3.20	4.26	2.71	4.44	7.11	6.58	3.44	2.78
29.00	3.00	2.13	3.20	4.26	2.72	4.56	7.34	6.78	3.50	2.78
30.00	3.00	2.13	3.20	4.26	2.73	4.67	7.56	6.98	3.56	2.78

Table 3.2: Risk for different values of Δ^2 at p=4 and k=d=0.1, 0.5 and 0.9

Δ^2	LSE	LASSO			Ridge			Liu		
		q=1	*q=2*	*q=4*	*k=0.1*	*k=0.5*	*k=0.9*	*d=0.1*	*d=0.5*	*d=0.9*
0.00	4.00	1.89	3.77	4.72	3.31	1.78	1.11	1.21	2.25	3.61
1.00	4.00	1.89	3.77	4.72	3.31	1.89	1.33	1.41	2.31	3.61
2.00	4.00	1.89	3.77	4.72	3.32	2.00	1.56	1.62	2.38	3.62
3.00	4.00	1.89	3.77	4.72	3.33	2.11	1.78	1.82	2.44	3.62
4.00	4.00	1.89	3.77	4.72	3.34	2.22	2.01	2.02	2.50	3.62
5.00	4.00	1.89	3.77	4.72	3.35	2.33	2.23	2.22	2.56	3.62
6.00	4.00	1.89	3.77	4.72	3.36	2.44	2.45	2.43	2.63	3.63
7.00	4.00	1.89	3.77	4.72	3.36	2.56	2.68	2.63	2.69	3.63
8.00	4.00	1.89	3.77	4.72	3.37	2.67	2.90	2.83	2.75	3.63
9.00	4.00	1.89	3.77	4.72	3.38	2.78	3.13	3.03	2.81	3.63
10.00	4.00	1.89	3.77	4.72	3.39	2.89	3.35	3.24	2.88	3.64
11.00	4.00	1.89	3.77	4.72	3.40	3.00	3.58	3.44	2.94	3.64
12.00	4.00	1.89	3.77	4.72	3.40	3.11	3.80	3.64	3.00	3.64
13.00	4.00	1.89	3.77	4.72	3.41	3.22	4.02	3.84	3.06	3.64
14.00	4.00	1.89	3.77	4.72	3.42	3.33	4.25	4.05	3.13	3.65
15.00	4.00	1.89	3.77	4.72	3.43	3.44	4.47	4.25	3.19	3.65
16.00	4.00	1.89	3.77	4.72	3.44	3.56	4.70	4.45	3.25	3.65
17.00	4.00	1.89	3.77	4.72	3.45	3.67	4.92	4.65	3.31	3.65
18.00	4.00	1.89	3.77	4.72	3.45	3.78	5.15	4.86	3.38	3.66
19.00	4.00	1.89	3.77	4.72	3.46	3.89	5.37	5.06	3.44	3.66
20.00	4.00	1.89	3.77	4.72	3.47	4.00	5.60	5.26	3.50	3.66
21.00	4.00	1.89	3.77	4.72	3.48	4.11	5.82	5.46	3.56	3.66
22.00	4.00	1.89	3.77	4.72	3.49	4.22	6.04	5.67	3.63	3.67
23.00	4.00	1.89	3.77	4.72	3.50	4.33	6.27	5.87	3.69	3.67
24.00	4.00	1.89	3.77	4.72	3.50	4.44	6.49	6.07	3.75	3.67
25.00	4.00	1.89	3.77	4.72	3.51	4.56	6.72	6.27	3.81	3.67
26.00	4.00	1.89	3.77	4.72	3.52	4.67	6.94	6.48	3.88	3.68
27.00	4.00	1.89	3.77	4.72	3.53	4.78	7.17	6.68	3.94	3.68
28.00	4.00	1.89	3.77	4.72	3.54	4.89	7.39	6.88	4.00	3.68
29.00	4.00	1.89	3.77	4.72	3.55	5.00	7.61	7.08	4.06	3.68
30.00	4.00	1.89	3.77	4.72	3.55	5.11	7.84	7.29	4.13	3.69

Table 3.3: Risk for different values of Δ^2 at p=5 and k=d=0.1, 0.5 and 0.9

Δ^2	LSE	LASSO			Ridge			Liu		
		q=1	*q=3*	*q=5*	*k=0.1*	*k=0.5*	*k=0.9*	*d=0.1*	*d=0.5*	*d=0.9*
0.00	5.00	1.69	3.38	5.06	4.13	2.22	1.39	1.51	2.81	4.51
1.00	5.00	1.69	3.38	5.06	4.14	2.33	1.61	1.72	2.88	4.52
2.00	5.00	1.69	3.38	5.06	4.15	2.44	1.83	1.92	2.94	4.52
3.00	5.00	1.69	3.38	5.06	4.16	2.56	2.06	2.12	3.00	4.52
4.00	5.00	1.69	3.38	5.06	4.17	2.67	2.28	2.32	3.06	4.52
5.00	5.00	1.69	3.38	5.06	4.17	2.78	2.51	2.53	3.13	4.53
6.00	5.00	1.69	3.38	5.06	4.18	2.89	2.73	2.73	3.19	4.53
7.00	5.00	1.69	3.38	5.06	4.19	3.00	2.96	2.93	3.25	4.53
8.00	5.00	1.69	3.38	5.06	4.20	3.11	3.18	3.13	3.31	4.53
9.00	5.00	1.69	3.38	5.06	4.21	3.22	3.40	3.34	3.38	4.54
10.00	5.00	1.69	3.38	5.06	4.21	3.33	3.63	3.54	3.44	4.54
11.00	5.00	1.69	3.38	5.06	4.22	3.44	3.85	3.74	3.50	4.54
12.00	5.00	1.69	3.38	5.06	4.23	3.56	4.08	3.94	3.56	4.54
13.00	5.00	1.69	3.38	5.06	4.24	3.67	4.30	4.15	3.63	4.55
14.00	5.00	1.69	3.38	5.06	4.25	3.78	4.53	4.35	3.69	4.55
15.00	5.00	1.69	3.38	5.06	4.26	3.89	4.75	4.55	3.75	4.55
16.00	5.00	1.69	3.38	5.06	4.26	4.00	4.98	4.75	3.81	4.55
17.00	5.00	1.69	3.38	5.06	4.27	4.11	5.20	4.96	3.88	4.56
18.00	5.00	1.69	3.38	5.06	4.28	4.22	5.42	5.16	3.94	4.56
19.00	5.00	1.69	3.38	5.06	4.29	4.33	5.65	5.36	4.00	4.56
20.00	5.00	1.69	3.38	5.06	4.30	4.44	5.87	5.56	4.06	4.56
21.00	5.00	1.69	3.38	5.06	4.31	4.56	6.10	5.77	4.13	4.57
22.00	5.00	1.69	3.38	5.06	4.31	4.67	6.32	5.97	4.19	4.57
23.00	5.00	1.69	3.38	5.06	4.32	4.78	6.55	6.17	4.25	4.57
24.00	5.00	1.69	3.38	5.06	4.33	4.89	6.77	6.37	4.31	4.57
25.00	5.00	1.69	3.38	5.06	4.34	5.00	6.99	6.58	4.38	4.58
26.00	5.00	1.69	3.38	5.06	4.35	5.11	7.22	6.78	4.44	4.58
27.00	5.00	1.69	3.38	5.06	4.36	5.22	7.44	6.98	4.50	4.58
28.00	5.00	1.69	3.38	5.06	4.36	5.33	7.67	7.18	4.56	4.58
29.00	5.00	1.69	3.38	5.06	4.37	5.44	7.89	7.39	4.63	4.59
30.00	5.00	1.69	3.38	5.06	4.38	5.56	8.12	7.59	4.69	4.59

Table 3.4: Risk for different values of Δ^2 at p=7 and k=d=0.1, 0.5 and 0.9

Δ^2	LSE	LASSO				Ridge			Liu		
		q=1	*q=3*	*q=5*	*q=7*	*k=0.1*	*k=0.5*	*k=0.9*	*d=0.1*	*d=0.5*	*d=0.9*
0.00	7.00	1.40	2.80	4.19	5.59	5.79	3.11	1.94	2.12	3.94	6.32
1.00	7.00	1.40	2.80	4.19	5.59	5.79	3.22	2.16	2.32	4.00	6.32
2.00	7.00	1.40	2.80	4.19	5.59	5.80	3.33	2.39	2.52	4.06	6.32
3.00	7.00	1.40	2.80	4.19	5.59	5.81	3.44	2.61	2.73	4.13	6.33
4.00	7.00	1.40	2.80	4.19	5.59	5.82	3.56	2.84	2.93	4.19	6.33
5.00	7.00	1.40	2.80	4.19	5.59	5.83	3.67	3.06	3.13	4.25	6.33
6.00	7.00	1.40	2.80	4.19	5.59	5.83	3.78	3.29	3.33	4.31	6.33
7.00	7.00	1.40	2.80	4.19	5.59	5.84	3.89	3.51	3.54	4.38	6.34
8.00	7.00	1.40	2.80	4.19	5.59	5.85	4.00	3.73	3.74	4.44	6.34
9.00	7.00	1.40	2.80	4.19	5.59	5.86	4.11	3.96	3.94	4.50	6.34
10.00	7.00	1.40	2.80	4.19	5.59	5.87	4.22	4.18	4.14	4.56	6.34
11.00	7.00	1.40	2.80	4.19	5.59	5.88	4.33	4.41	4.35	4.63	6.35
12.00	7.00	1.40	2.80	4.19	5.59	5.88	4.44	4.63	4.55	4.69	6.35
13.00	7.00	1.40	2.80	4.19	5.59	5.89	4.56	4.86	4.75	4.75	6.35
14.00	7.00	1.40	2.80	4.19	5.59	5.90	4.67	5.08	4.95	4.81	6.35
15.00	7.00	1.40	2.80	4.19	5.59	5.91	4.78	5.30	5.16	4.88	6.36
16.00	7.00	1.40	2.80	4.19	5.59	5.92	4.89	5.53	5.36	4.94	6.36
17.00	7.00	1.40	2.80	4.19	5.59	5.93	5.00	5.75	5.56	5.00	6.36
18.00	7.00	1.40	2.80	4.19	5.59	5.93	5.11	5.98	5.76	5.06	6.36
19.00	7.00	1.40	2.80	4.19	5.59	5.94	5.22	6.20	5.97	5.13	6.37
20.00	7.00	1.40	2.80	4.19	5.59	5.95	5.33	6.43	6.17	5.19	6.37
21.00	7.00	1.40	2.80	4.19	5.59	5.96	5.44	6.65	6.37	5.25	6.37
22.00	7.00	1.40	2.80	4.19	5.59	5.97	5.56	6.88	6.57	5.31	6.37
23.00	7.00	1.40	2.80	4.19	5.59	5.98	5.67	7.10	6.78	5.38	6.38
24.00	7.00	1.40	2.80	4.19	5.59	5.98	5.78	7.32	6.98	5.44	6.38
25.00	7.00	1.40	2.80	4.19	5.59	5.99	5.89	7.55	7.18	5.50	6.38
26.00	7.00	1.40	2.80	4.19	5.59	6.00	6.00	7.77	7.38	5.56	6.38
27.00	7.00	1.40	2.80	4.19	5.59	6.01	6.11	8.00	7.59	5.63	6.39
28.00	7.00	1.40	2.80	4.19	5.59	6.02	6.22	8.22	7.79	5.69	6.39
29.00	7.00	1.40	2.80	4.19	5.59	6.02	6.33	8.45	7.99	5.75	6.39
30.00	7.00	1.40	2.80	4.19	5.59	6.03	6.44	8.67	8.19	5.81	6.39

Table 3.5: Risk for different values of Δ^2 at *p*=6 and *k*=*d*=0.1, 0.5 and 0.9

Δ^2	LSE	LASSO				Ridge			Liu		
		q=1	*q=3*	*q=5*	*q=6*	*k=0.1*	*k=0.5*	*k=0.9*	*d=0.1*	*d=0.5*	*d=0.9*
0.00	6.00	1.53	3.06	4.58	5.35	4.96	2.67	1.66	1.82	3.38	5.42
1.00	6.00	1.53	3.06	4.58	5.35	4.97	2.78	1.89	2.02	3.44	5.42
2.00	6.00	1.53	3.06	4.58	5.35	4.98	2.89	2.11	2.22	3.50	5.42
3.00	6.00	1.53	3.06	4.58	5.35	4.98	3.00	2.34	2.42	3.56	5.42
4.00	6.00	1.53	3.06	4.58	5.35	4.99	3.11	2.56	2.63	3.63	5.43
5.00	6.00	1.53	3.06	4.58	5.35	5.00	3.22	2.78	2.83	3.69	5.43
6.00	6.00	1.53	3.06	4.58	5.35	5.01	3.33	3.01	3.03	3.75	5.43
7.00	6.00	1.53	3.06	4.58	5.35	5.02	3.44	3.23	3.23	3.81	5.43
8.00	6.00	1.53	3.06	4.58	5.35	5.02	3.56	3.46	3.44	3.88	5.44
9.00	6.00	1.53	3.06	4.58	5.35	5.03	3.67	3.68	3.64	3.94	5.44
10.00	6.00	1.53	3.06	4.58	5.35	5.04	3.78	3.91	3.84	4.00	5.44
11.00	6.00	1.53	3.06	4.58	5.35	5.05	3.89	4.13	4.04	4.06	5.44
12.00	6.00	1.53	3.06	4.58	5.35	5.06	4.00	4.35	4.25	4.13	5.45
13.00	6.00	1.53	3.06	4.58	5.35	5.07	4.11	4.58	4.45	4.19	5.45
14.00	6.00	1.53	3.06	4.58	5.35	5.07	4.22	4.80	4.65	4.25	5.45
15.00	6.00	1.53	3.06	4.58	5.35	5.08	4.33	5.03	4.85	4.31	5.45
16.00	6.00	1.53	3.06	4.58	5.35	5.09	4.44	5.25	5.06	4.38	5.46
17.00	6.00	1.53	3.06	4.58	5.35	5.10	4.56	5.48	5.26	4.44	5.46
18.00	6.00	1.53	3.06	4.58	5.35	5.11	4.67	5.70	5.46	4.50	5.46
19.00	6.00	1.53	3.06	4.58	5.35	5.12	4.78	5.93	5.66	4.56	5.46
20.00	6.00	1.53	3.06	4.58	5.35	5.12	4.89	6.15	5.87	4.63	5.47
21.00	6.00	1.53	3.06	4.58	5.35	5.13	5.00	6.37	6.07	4.69	5.47
22.00	6.00	1.53	3.06	4.58	5.35	5.14	5.11	6.60	6.27	4.75	5.47
23.00	6.00	1.53	3.06	4.58	5.35	5.15	5.22	6.82	6.47	4.81	5.47
24.00	6.00	1.53	3.06	4.58	5.35	5.16	5.33	7.05	6.68	4.88	5.48
25.00	6.00	1.53	3.06	4.58	5.35	5.17	5.44	7.27	6.88	4.94	5.48
26.00	6.00	1.53	3.06	4.58	5.35	5.17	5.56	7.50	7.08	5.00	5.48
27.00	6.00	1.53	3.06	4.58	5.35	5.18	5.67	7.72	7.28	5.06	5.48
28.00	6.00	1.53	3.06	4.58	5.35	5.19	5.78	7.94	7.49	5.13	5.49
29.00	6.00	1.53	3.06	4.58	5.35	5.20	5.89	8.17	7.69	5.19	5.49
30.00	6.00	1.53	3.06	4.58	5.35	5.21	6.00	8.39	7.89	5.25	5.49

Table 3.6: Risk for different values of Δ^2 at p=8 and k=d=0.1, 0.5 and 0.9

Δ^2	LSE	LASSO					Ridge			Liu		
		q=1	*q=3*	*q=5*	*q=7*	*q=8*	*k=0.1*	*k=0.5*	*k=0.9*	*d=0.1*	*d=0.5*	*d=0.9*
0.00	8.00	1.29	2.58	3.87	5.16	5.80	6.61	3.56	2.22	2.42	4.50	7.22
1.00	8.00	1.29	2.58	3.87	5.16	5.80	6.62	3.67	2.44	2.62	4.56	7.22
2.00	8.00	1.29	2.58	3.87	5.16	5.80	6.63	3.78	2.66	2.83	4.63	7.23
3.00	8.00	1.29	2.58	3.87	5.16	5.80	6.64	3.89	2.89	3.03	4.69	7.23
4.00	8.00	1.29	2.58	3.87	5.16	5.80	6.64	4.00	3.11	3.23	4.75	7.23
5.00	8.00	1.29	2.58	3.87	5.16	5.80	6.65	4.11	3.34	3.43	4.81	7.23
6.00	8.00	1.29	2.58	3.87	5.16	5.80	6.66	4.22	3.56	3.64	4.88	7.24
7.00	8.00	1.29	2.58	3.87	5.16	5.80	6.67	4.33	3.79	3.84	4.94	7.24
8.00	8.00	1.29	2.58	3.87	5.16	5.80	6.68	4.44	4.01	4.04	5.00	7.24
9.00	8.00	1.29	2.58	3.87	5.16	5.80	6.69	4.56	4.24	4.24	5.06	7.24
10.00	8.00	1.29	2.58	3.87	5.16	5.80	6.69	4.67	4.46	4.45	5.13	7.25
11.00	8.00	1.29	2.58	3.87	5.16	5.80	6.70	4.78	4.68	4.65	5.19	7.25
12.00	8.00	1.29	2.58	3.87	5.16	5.80	6.71	4.89	4.91	4.85	5.25	7.25
13.00	8.00	1.29	2.58	3.87	5.16	5.80	6.72	5.00	5.13	5.05	5.31	7.25
14.00	8.00	1.29	2.58	3.87	5.16	5.80	6.73	5.11	5.36	5.26	5.38	7.26
15.00	8.00	1.29	2.58	3.87	5.16	5.80	6.74	5.22	5.58	5.46	5.44	7.26
16.00	8.00	1.29	2.58	3.87	5.16	5.80	6.74	5.33	5.81	5.66	5.50	7.26
17.00	8.00	1.29	2.58	3.87	5.16	5.80	6.75	5.44	6.03	5.86	5.56	7.26
18.00	8.00	1.29	2.58	3.87	5.16	5.80	6.76	5.56	6.25	6.07	5.63	7.27
19.00	8.00	1.29	2.58	3.87	5.16	5.80	6.77	5.67	6.48	6.27	5.69	7.27
20.00	8.00	1.29	2.58	3.87	5.16	5.80	6.78	5.78	6.70	6.47	5.75	7.27
21.00	8.00	1.29	2.58	3.87	5.16	5.80	6.79	5.89	6.93	6.67	5.81	7.27
22.00	8.00	1.29	2.58	3.87	5.16	5.80	6.79	6.00	7.15	6.88	5.88	7.28
23.00	8.00	1.29	2.58	3.87	5.16	5.80	6.80	6.11	7.38	7.08	5.94	7.28
24.00	8.00	1.29	2.58	3.87	5.16	5.80	6.81	6.22	7.60	7.28	6.00	7.28
25.00	8.00	1.29	2.58	3.87	5.16	5.80	6.82	6.33	7.83	7.48	6.06	7.28
26.00	8.00	1.29	2.58	3.87	5.16	5.80	6.83	6.44	8.05	7.69	6.13	7.29
27.00	8.00	1.29	2.58	3.87	5.16	5.80	6.83	6.56	8.27	7.89	6.19	7.29
28.00	8.00	1.29	2.58	3.87	5.16	5.80	6.84	6.67	8.50	8.09	6.25	7.29
29.00	8.00	1.29	2.58	3.87	5.16	5.80	6.85	6.78	8.72	8.29	6.31	7.29
30.00	8.00	1.29	2.58	3.87	5.16	5.80	6.86	6.89	8.95	8.50	6.38	7.30

Table 3.7: Risk for different values of Δ^2 at p=9 and k=d=0.1, 0.5 and 0.9

Δ^2	LSE	LASSO					Ridge			Liu		
		q=1	*q=3*	*q=5*	*q=7*	*q=9*	*k=0.1*	*k=0.5*	*k=0.9*	*d=0.1*	*d=0.5*	*d=0.9*
0.00	9.00	1.20	2.40	3.60	4.80	5.99	7.44	4.00	2.49	2.72	5.06	8.12
1.00	9.00	1.20	2.40	3.60	4.80	5.99	7.45	4.11	2.72	2.93	5.13	8.13
2.00	9.00	1.20	2.40	3.60	4.80	5.99	7.45	4.22	2.94	3.13	5.19	8.13
3.00	9.00	1.20	2.40	3.60	4.80	5.99	7.46	4.33	3.17	3.33	5.25	8.13
4.00	9.00	1.20	2.40	3.60	4.80	5.99	7.47	4.44	3.39	3.53	5.31	8.13
5.00	9.00	1.20	2.40	3.60	4.80	5.99	7.48	4.56	3.61	3.74	5.38	8.14
6.00	9.00	1.20	2.40	3.60	4.80	5.99	7.49	4.67	3.84	3.94	5.44	8.14
7.00	9.00	1.20	2.40	3.60	4.80	5.99	7.50	4.78	4.06	4.14	5.50	8.14
8.00	9.00	1.20	2.40	3.60	4.80	5.99	7.50	4.89	4.29	4.34	5.56	8.14
9.00	9.00	1.20	2.40	3.60	4.80	5.99	7.51	5.00	4.51	4.55	5.63	8.15
10.00	9.00	1.20	2.40	3.60	4.80	5.99	7.52	5.11	4.74	4.75	5.69	8.15
11.00	9.00	1.20	2.40	3.60	4.80	5.99	7.53	5.22	4.96	4.95	5.75	8.15
12.00	9.00	1.20	2.40	3.60	4.80	5.99	7.54	5.33	5.19	5.15	5.81	8.15
13.00	9.00	1.20	2.40	3.60	4.80	5.99	7.55	5.44	5.41	5.36	5.88	8.16
14.00	9.00	1.20	2.40	3.60	4.80	5.99	7.55	5.56	5.63	5.56	5.94	8.16
15.00	9.00	1.20	2.40	3.60	4.80	5.99	7.56	5.67	5.86	5.76	6.00	8.16
16.00	9.00	1.20	2.40	3.60	4.80	5.99	7.57	5.78	6.08	5.96	6.06	8.16
17.00	9.00	1.20	2.40	3.60	4.80	5.99	7.58	5.89	6.31	6.17	6.13	8.17
18.00	9.00	1.20	2.40	3.60	4.80	5.99	7.59	6.00	6.53	6.37	6.19	8.17
19.00	9.00	1.20	2.40	3.60	4.80	5.99	7.60	6.11	6.76	6.57	6.25	8.17
20.00	9.00	1.20	2.40	3.60	4.80	5.99	7.60	6.22	6.98	6.77	6.31	8.17
21.00	9.00	1.20	2.40	3.60	4.80	5.99	7.61	6.33	7.20	6.98	6.38	8.18
22.00	9.00	1.20	2.40	3.60	4.80	5.99	7.62	6.44	7.43	7.18	6.44	8.18
23.00	9.00	1.20	2.40	3.60	4.80	5.99	7.63	6.56	7.65	7.38	6.50	8.18
24.00	9.00	1.20	2.40	3.60	4.80	5.99	7.64	6.67	7.88	7.58	6.56	8.18
25.00	9.00	1.20	2.40	3.60	4.80	5.99	7.64	6.78	8.10	7.79	6.63	8.19
26.00	9.00	1.20	2.40	3.60	4.80	5.99	7.65	6.89	8.33	7.99	6.69	8.19
27.00	9.00	1.20	2.40	3.60	4.80	5.99	7.66	7.00	8.55	8.19	6.75	8.19
28.00	9.00	1.20	2.40	3.60	4.80	5.99	7.67	7.11	8.78	8.39	6.81	8.19
29.00	9.00	1.20	2.40	3.60	4.80	5.99	7.68	7.22	9.00	8.60	6.88	8.20
30.00	9.00	1.20	2.40	3.60	4.80	5.99	7.69	7.33	9.22	8.80	6.94	8.20

Table 3.8: Risk for different values of Δ^2 at p=10 and k=d=0.1, 0.5 and 0.9

Δ^2	LSE	LASSO						Ridge			Liu		
		q=1	*q=3*	*q=5*	*q=7*	*q=9*	*q=10*	*k=0.1*	*k=0.5*	*k=0.9*	*d=0.1*	*d=0.5*	d=0.9
0.00	10.00	1.12	2.24	3.36	4.48	5.61	6.17	8.26	4.44	2.77	3.03	5.63	9.03
1.00	10.00	1.12	2.24	3.36	4.48	5.61	6.17	8.27	4.56	2.99	3.23	5.69	9.03
2.00	10.00	1.12	2.24	3.36	4.48	5.61	6.17	8.28	4.67	3.22	3.43	5.75	9.03
3.00	10.00	1.12	2.24	3.36	4.48	5.61	6.17	8.29	4.78	3.44	3.63	5.81	9.03
4.00	10.00	1.12	2.24	3.36	4.48	5.61	6.17	8.30	4.89	3.67	3.84	5.88	9.04
5.00	10.00	1.12	2.24	3.36	4.48	5.61	6.17	8.31	5.00	3.89	4.04	5.94	9.04
6.00	10.00	1.12	2.24	3.36	4.48	5.61	6.17	8.31	5.11	4.12	4.24	6.00	9.04
7.00	10.00	1.12	2.24	3.36	4.48	5.61	6.17	8.32	5.22	4.34	4.44	6.06	9.04
8.00	10.00	1.12	2.24	3.36	4.48	5.61	6.17	8.33	5.33	4.57	4.65	6.13	9.05
9.00	10.00	1.12	2.24	3.36	4.48	5.61	6.17	8.34	5.44	4.79	4.85	6.19	9.05
10.00	10.00	1.12	2.24	3.36	4.48	5.61	6.17	8.35	5.56	5.01	5.05	6.25	9.05
11.00	10.00	1.12	2.24	3.36	4.48	5.61	6.17	8.36	5.67	5.24	5.25	6.31	9.05
12.00	10.00	1.12	2.24	3.36	4.48	5.61	6.17	8.36	5.78	5.46	5.46	6.38	9.06
13.00	10.00	1.12	2.24	3.36	4.48	5.61	6.17	8.37	5.89	5.69	5.66	6.44	9.06
14.00	10.00	1.12	2.24	3.36	4.48	5.61	6.17	8.38	6.00	5.91	5.86	6.50	9.06
15.00	10.00	1.12	2.24	3.36	4.48	5.61	6.17	8.39	6.11	6.14	6.06	6.56	9.06
16.00	10.00	1.12	2.24	3.36	4.48	5.61	6.17	8.40	6.22	6.36	6.27	6.63	9.07
17.00	10.00	1.12	2.24	3.36	4.48	5.61	6.17	8.40	6.33	6.58	6.47	6.69	9.07
18.00	10.00	1.12	2.24	3.36	4.48	5.61	6.17	8.41	6.44	6.81	6.67	6.75	9.07
19.00	10.00	1.12	2.24	3.36	4.48	5.61	6.17	8.42	6.56	7.03	6.87	6.81	9.07
20.00	10.00	1.12	2.24	3.36	4.48	5.61	6.17	8.43	6.67	7.26	7.08	6.88	9.08
21.00	10.00	1.12	2.24	3.36	4.48	5.61	6.17	8.44	6.78	7.48	7.28	6.94	9.08
22.00	10.00	1.12	2.24	3.36	4.48	5.61	6.17	8.45	6.89	7.71	7.48	7.00	9.08
23.00	10.00	1.12	2.24	3.36	4.48	5.61	6.17	8.45	7.00	7.93	7.68	7.06	9.08
24.00	10.00	1.12	2.24	3.36	4.48	5.61	6.17	8.46	7.11	8.16	7.89	7.13	9.09
25.00	10.00	1.12	2.24	3.36	4.48	5.61	6.17	8.47	7.22	8.38	8.09	7.19	9.09
26.00	10.00	1.12	2.24	3.36	4.48	5.61	6.17	8.48	7.33	8.60	8.29	7.25	9.09
27.00	10.00	1.12	2.24	3.36	4.48	5.61	6.17	8.49	7.44	8.83	8.49	7.31	9.09
28.00	10.00	1.12	2.24	3.36	4.48	5.61	6.17	8.50	7.56	9.05	8.70	7.38	9.10
29.00	10.00	1.12	2.24	3.36	4.48	5.61	6.17	8.50	7.67	9.28	8.90	7.44	9.10
30.00	10.00	1.12	2.24	3.36	4.48	5.61	6.17	8.51	7.78	9.50	9.10	7.50	9.10

Table 3.9: Efficiency for different values of Δ^2 at *p*=3 and *k*=*d*=0.1, 0.5 and 0.9

Δ^2	LSE	LASSO			Ridge			Liu		
		q=1	*q=2*	*q=3*	*k=0.1*	*k=0.5*	*k=0.9*	*d=0.1*	*d=0.5*	*d=0.9*
0.00	1.00	1.41	0.94	0.70	1.21	2.25	3.61	3.31	1.78	1.11
1.00	1.00	1.41	0.94	0.70	1.21	2.08	2.84	2.70	1.71	1.11
2.00	1.00	1.41	0.94	0.70	1.20	1.93	2.34	2.29	1.66	1.11
3.00	1.00	1.41	0.94	0.70	1.20	1.80	1.99	1.98	1.60	1.10
4.00	1.00	1.41	0.94	0.70	1.19	1.69	1.74	1.75	1.55	1.10
5.00	1.00	1.41	0.94	0.70	1.19	1.59	1.54	1.56	1.50	1.10
6.00	1.00	1.41	0.94	0.70	1.19	1.50	1.38	1.41	1.45	1.10
7.00	1.00	1.41	0.94	0.70	1.18	1.42	1.25	1.29	1.41	1.10
8.00	1.00	1.41	0.94	0.70	1.18	1.35	1.14	1.19	1.37	1.10
9.00	1.00	1.41	0.94	0.70	1.17	1.29	1.05	1.10	1.33	1.10
10.00	1.00	1.41	0.94	0.70	1.17	1.23	0.98	1.02	1.30	1.10
11.00	1.00	1.41	0.94	0.70	1.17	1.17	0.91	0.96	1.26	1.10
12.00	1.00	1.41	0.94	0.70	1.16	1.13	0.85	0.90	1.23	1.10
13.00	1.00	1.41	0.94	0.70	1.16	1.08	0.80	0.85	1.20	1.09
14.00	1.00	1.41	0.94	0.70	1.16	1.04	0.76	0.80	1.17	1.09
15.00	1.00	1.41	0.94	0.70	1.15	1.00	0.71	0.76	1.14	1.09
16.00	1.00	1.41	0.94	0.70	1.15	0.96	0.68	0.72	1.12	1.09
17.00	1.00	1.41	0.94	0.70	1.15	0.93	0.65	0.69	1.09	1.09
18.00	1.00	1.41	0.94	0.70	1.14	0.90	0.62	0.66	1.07	1.09
19.00	1.00	1.41	0.94	0.70	1.14	0.87	0.59	0.63	1.04	1.09
20.00	1.00	1.41	0.94	0.70	1.13	0.84	0.56	0.61	1.02	1.09
21.00	1.00	1.41	0.94	0.70	1.13	0.82	0.54	0.58	1.00	1.09
22.00	1.00	1.41	0.94	0.70	1.13	0.79	0.52	0.56	0.98	1.09
23.00	1.00	1.41	0.94	0.70	1.12	0.77	0.50	0.54	0.96	1.08
24.00	1.00	1.41	0.94	0.70	1.12	0.75	0.48	0.52	0.94	1.08
25.00	1.00	1.41	0.94	0.70	1.12	0.73	0.47	0.50	0.92	1.08
26.00	1.00	1.41	0.94	0.70	1.11	0.71	0.45	0.49	0.91	1.08
27.00	1.00	1.41	0.94	0.70	1.11	0.69	0.44	0.47	0.89	1.08
28.00	1.00	1.41	0.94	0.70	1.11	0.68	0.42	0.46	0.87	1.08
29.00	1.00	1.41	0.94	0.70	1.10	0.66	0.41	0.44	0.86	1.08
30.00	1.00	1.41	0.94	0.70	1.10	0.64	0.40	0.43	0.84	1.08

Table 3.10: Efficiency for different values of Δ^2 at *p*=5 and *k*=*d*=0.1, 0.5 and 0.9

Δ^2	LSE	LASSO			Ridge			Liu		
		q=1	*q=3*	*q=5*	*k=0.1*	*k=0.5*	*k=0.9*	*d=0.1*	*d=0.5*	*d=0.9*
0.00	1.00	2.96	1.48	0.99	1.21	2.25	3.61	3.31	1.78	1.11
1.00	1.00	2.96	1.48	0.99	1.21	2.14	3.11	2.92	1.74	1.11
2.00	1.00	2.96	1.48	0.99	1.21	2.05	2.73	2.61	1.70	1.11
3.00	1.00	2.96	1.48	0.99	1.20	1.96	2.43	2.36	1.67	1.11
4.00	1.00	2.96	1.48	0.99	1.20	1.88	2.19	2.15	1.63	1.11
5.00	1.00	2.96	1.48	0.99	1.20	1.80	1.99	1.98	1.60	1.10
6.00	1.00	2.96	1.48	0.99	1.20	1.73	1.83	1.83	1.57	1.10
7.00	1.00	2.96	1.48	0.99	1.19	1.67	1.69	1.71	1.54	1.10
8.00	1.00	2.96	1.48	0.99	1.19	1.61	1.57	1.60	1.51	1.10
9.00	1.00	2.96	1.48	0.99	1.19	1.55	1.47	1.50	1.48	1.10
10.00	1.00	2.96	1.48	0.99	1.19	1.50	1.38	1.41	1.45	1.10
11.00	1.00	2.96	1.48	0.99	1.18	1.45	1.30	1.34	1.43	1.10
12.00	1.00	2.96	1.48	0.99	1.18	1.41	1.23	1.27	1.40	1.10
13.00	1.00	2.96	1.48	0.99	1.18	1.36	1.16	1.21	1.38	1.10
14.00	1.00	2.96	1.48	0.99	1.18	1.32	1.10	1.15	1.36	1.10
15.00	1.00	2.96	1.48	0.99	1.17	1.29	1.05	1.10	1.33	1.10
16.00	1.00	2.96	1.48	0.99	1.17	1.25	1.01	1.05	1.31	1.10
17.00	1.00	2.96	1.48	0.99	1.17	1.22	0.96	1.01	1.29	1.10
18.00	1.00	2.96	1.48	0.99	1.17	1.18	0.92	0.97	1.27	1.10
19.00	1.00	2.96	1.48	0.99	1.17	1.15	0.89	0.93	1.25	1.10
20.00	1.00	2.96	1.48	0.99	1.16	1.13	0.85	0.90	1.23	1.10
21.00	1.00	2.96	1.48	0.99	1.16	1.10	0.82	0.87	1.21	1.10
22.00	1.00	2.96	1.48	0.99	1.16	1.07	0.79	0.84	1.19	1.09
23.00	1.00	2.96	1.48	0.99	1.16	1.05	0.76	0.81	1.18	1.09
24.00	1.00	2.96	1.48	0.99	1.15	1.02	0.74	0.78	1.16	1.09
25.00	1.00	2.96	1.48	0.99	1.15	1.00	0.71	0.76	1.14	1.09
26.00	1.00	2.96	1.48	0.99	1.15	0.98	0.69	0.74	1.13	1.09
27.00	1.00	2.96	1.48	0.99	1.15	0.96	0.67	0.72	1.11	1.09
28.00	1.00	2.96	1.48	0.99	1.15	0.94	0.65	0.70	1.10	1.09
29.00	1.00	2.96	1.48	0.99	1.14	0.92	0.63	0.68	1.08	1.09
30.00	1.00	2.96	1.48	0.99	1.14	0.90	0.62	0.66	1.07	1.09

Table 3.11: Efficiency for different values of Δ^2 at p=7 and k=d=0.1, 0.5 and 0.9

Δ^2	LSE	LASSO				Ridge			Liu		
		q=1	*q=3*	*q=5*	*q=7*	*k=0.1*	*k=0.5*	*k=0.9*	*d=0.1*	*d=0.5*	*d=0.9*
0.00	1.00	5.01	2.50	1.67	1.25	1.21	2.25	3.61	3.31	1.78	1.11
1.00	1.00	5.01	2.50	1.67	1.25	1.21	2.17	3.24	3.02	1.75	1.11
2.00	1.00	5.01	2.50	1.67	1.25	1.21	2.10	2.93	2.78	1.72	1.11
3.00	1.00	5.01	2.50	1.67	1.25	1.20	2.03	2.68	2.57	1.70	1.11
4.00	1.00	5.01	2.50	1.67	1.25	1.20	1.97	2.47	2.39	1.67	1.11
5.00	1.00	5.01	2.50	1.67	1.25	1.20	1.91	2.29	2.24	1.65	1.11
6.00	1.00	5.01	2.50	1.67	1.25	1.20	1.85	2.13	2.10	1.62	1.11
7.00	1.00	5.01	2.50	1.67	1.25	1.20	1.80	1.99	1.98	1.60	1.10
8.00	1.00	5.01	2.50	1.67	1.25	1.20	1.75	1.87	1.87	1.58	1.10
9.00	1.00	5.01	2.50	1.67	1.25	1.19	1.70	1.77	1.78	1.56	1.10
10.00	1.00	5.01	2.50	1.67	1.25	1.19	1.66	1.67	1.69	1.53	1.10
11.00	1.00	5.01	2.50	1.67	1.25	1.19	1.62	1.59	1.61	1.51	1.10
12.00	1.00	5.01	2.50	1.67	1.25	1.19	1.58	1.51	1.54	1.49	1.10
13.00	1.00	5.01	2.50	1.67	1.25	1.19	1.54	1.44	1.47	1.47	1.10
14.00	1.00	5.01	2.50	1.67	1.25	1.19	1.50	1.38	1.41	1.45	1.10
15.00	1.00	5.01	2.50	1.67	1.25	1.18	1.47	1.32	1.36	1.44	1.10
16.00	1.00	5.01	2.50	1.67	1.25	1.18	1.43	1.27	1.31	1.42	1.10
17.00	1.00	5.01	2.50	1.67	1.25	1.18	1.40	1.22	1.26	1.40	1.10
18.00	1.00	5.01	2.50	1.67	1.25	1.18	1.37	1.17	1.21	1.38	1.10
19.00	1.00	5.01	2.50	1.67	1.25	1.18	1.34	1.13	1.17	1.37	1.10
20.00	1.00	5.01	2.50	1.67	1.25	1.18	1.31	1.09	1.13	1.35	1.10
21.00	1.00	5.01	2.50	1.67	1.25	1.17	1.29	1.05	1.10	1.33	1.10
22.00	1.00	5.01	2.50	1.67	1.25	1.17	1.26	1.02	1.07	1.32	1.10
23.00	1.00	5.01	2.50	1.67	1.25	1.17	1.24	0.99	1.03	1.30	1.10
24.00	1.00	5.01	2.50	1.67	1.25	1.17	1.21	0.96	1.00	1.29	1.10
25.00	1.00	5.01	2.50	1.67	1.25	1.17	1.19	0.93	0.97	1.27	1.10
26.00	1.00	5.01	2.50	1.67	1.25	1.17	1.17	0.90	0.95	1.26	1.10
27.00	1.00	5.01	2.50	1.67	1.25	1.17	1.15	0.88	0.92	1.24	1.10
28.00	1.00	5.01	2.50	1.67	1.25	1.16	1.13	0.85	0.90	1.23	1.10
29.00	1.00	5.01	2.50	1.67	1.25	1.16	1.11	0.83	0.88	1.22	1.10
30.00	1.00	5.01	2.50	1.67	1.25	1.16	1.09	0.81	0.85	1.20	1.10

Table 3.12: Efficiency for different values of Δ^2 at *p*=10 and *k*=*d*=0.1, 0.5 and 0.9

Δ^2	LSE	LASSO						Ridge			Liu		
		q=1	*q=3*	*q=5*	*q=7*	*q=9*	*q=10*	*k=0.1*	*k=0.5*	*k=0.9*	*d=0.1*	*d=0.5*	d=0.9
0.00	1.00	8.92	4.46	2.97	2.23	1.78	1.62	1.21	2.25	3.61	3.31	1.78	1.11
1.00	1.00	8.92	4.46	2.97	2.23	1.78	1.62	1.21	2.20	3.34	3.10	1.76	1.11
2.00	1.00	8.92	4.46	2.97	2.23	1.78	1.62	1.21	2.14	3.11	2.92	1.74	1.11
3.00	1.00	8.92	4.46	2.97	2.23	1.78	1.62	1.21	2.09	2.90	2.75	1.72	1.11
4.00	1.00	8.92	4.46	2.97	2.23	1.78	1.62	1.21	2.05	2.73	2.61	1.70	1.11
5.00	1.00	8.92	4.46	2.97	2.23	1.78	1.62	1.20	2.00	2.57	2.48	1.68	1.11
6.00	1.00	8.92	4.46	2.97	2.23	1.78	1.62	1.20	1.96	2.43	2.36	1.67	1.11
7.00	1.00	8.92	4.46	2.97	2.23	1.78	1.62	1.20	1.91	2.30	2.25	1.65	1.11
8.00	1.00	8.92	4.46	2.97	2.23	1.78	1.62	1.20	1.88	2.19	2.15	1.63	1.11
9.00	1.00	8.92	4.46	2.97	2.23	1.78	1.62	1.20	1.84	2.09	2.06	1.62	1.11
10.00	1.00	8.92	4.46	2.97	2.23	1.78	1.62	1.20	1.80	1.99	1.98	1.60	1.10
11.00	1.00	8.92	4.46	2.97	2.23	1.78	1.62	1.20	1.76	1.91	1.90	1.58	1.10
12.00	1.00	8.92	4.46	2.97	2.23	1.78	1.62	1.20	1.73	1.83	1.83	1.57	1.10
13.00	1.00	8.92	4.46	2.97	2.23	1.78	1.62	1.19	1.70	1.76	1.77	1.55	1.10
14.00	1.00	8.92	4.46	2.97	2.23	1.78	1.62	1.19	1.67	1.69	1.71	1.54	1.10
15.00	1.00	8.92	4.46	2.97	2.23	1.78	1.62	1.19	1.64	1.63	1.65	1.52	1.10
16.00	1.00	8.92	4.46	2.97	2.23	1.78	1.62	1.19	1.61	1.57	1.60	1.51	1.10
17.00	1.00	8.92	4.46	2.97	2.23	1.78	1.62	1.19	1.58	1.52	1.55	1.50	1.10
18.00	1.00	8.92	4.46	2.97	2.23	1.78	1.62	1.19	1.55	1.47	1.50	1.48	1.10
19.00	1.00	8.92	4.46	2.97	2.23	1.78	1.62	1.19	1.53	1.42	1.46	1.47	1.10
20.00	1.00	8.92	4.46	2.97	2.23	1.78	1.62	1.19	1.50	1.38	1.41	1.45	1.10
21.00	1.00	8.92	4.46	2.97	2.23	1.78	1.62	1.19	1.48	1.34	1.37	1.44	1.10
22.00	1.00	8.92	4.46	2.97	2.23	1.78	1.62	1.18	1.45	1.30	1.34	1.43	1.10
23.00	1.00	8.92	4.46	2.97	2.23	1.78	1.62	1.18	1.43	1.26	1.30	1.42	1.10
24.00	1.00	8.92	4.46	2.97	2.23	1.78	1.62	1.18	1.41	1.23	1.27	1.40	1.10
25.00	1.00	8.92	4.46	2.97	2.23	1.78	1.62	1.18	1.38	1.19	1.24	1.39	1.10
26.00	1.00	8.92	4.46	2.97	2.23	1.78	1.62	1.18	1.36	1.16	1.21	1.38	1.10
27.00	1.00	8.92	4.46	2.97	2.23	1.78	1.62	1.18	1.34	1.13	1.18	1.37	1.10
28.00	1.00	8.92	4.46	2.97	2.23	1.78	1.62	1.18	1.32	1.10	1.15	1.36	1.10
29.00	1.00	8.92	4.46	2.97	2.23	1.78	1.62	1.18	1.30	1.08	1.12	1.34	1.10
30.00	1.00	8.92	4.46	2.97	2.23	1.78	1.62	1.17	1.29	1.05	1.10	1.33	1.10

IV. COMPARISON OF RIDGE, LIU AND TWO PARAMETER BIASED ESTIMATORS

Since the comparison of Ridge, Liu and Two parameter biased estimator is limited in literature, in this chapter, I review some estimators for estimating ridge parameter *k* and optimum value of shrinkage parameter *d*. Since a theoretical comparison is not possible, I will do a simulation study to compare the performance of the estimators in the sense of smaller MSE.

4.1. Ridge, Liu and Two parameter biased Estimators.

Using the canonical form of linear model, we know from (2.4) that the MSE of generalized ridge regression estimator is,

$$\text{MSE}\ (\hat{\alpha}_k) = \sigma^2 \sum_{i=1}^{p} \frac{\lambda_i}{(\lambda_i+k_i)^2} + \sum_{i=1}^{p} \frac{k_i^2 \lambda_i^2}{(\lambda_i+k_i)^2}\ .$$

Note that in the previous Chapter 3, I compare the estimators based on the orthonormal regression model because of LASSO estimator, as the risk function is only available in orthonormal form.

It follows from Hoerl and Kennard (1970) that the value of k_i which minimizes the MSE($\hat{\alpha}_k$) is

$$k_i = \frac{\sigma^2}{\alpha_i^2},$$

where σ^2 represents the error variance of the model and α_i is the *i* th element of α. Hoerl and Kennard (1970), suggested to replace $\sigma^2$and α_i^2 by their corresponding unbiased estimators. That is,

$$\hat{k}_i = \frac{\hat{\sigma}^2}{\hat{\alpha}_i^2}\ . \tag{4.1}$$

Now I will review some estimators, they are as follows:

1. Estimator based on Hoerl, Kennard and Baldwin (1975) (thereafter $\hat{k}_{HKB}$ or HKB), proposed an estimator of *k* by taking harmonic mean of $\hat{k}_i$ in (4.1).

$$\hat{k}_{HKB} = \frac{p\hat{\sigma}^2}{\hat{\alpha}'\hat{\alpha}} . \qquad (4.2)$$

2. Estimator based on Lawless and Wang (1976) (thereafter $\hat{k}_{LW}$ or LW), proposed the following estimator:

$$\hat{k}_{LW} = \frac{p\hat{\sigma}^2}{\hat{\alpha}'X'X\hat{\alpha}} . \qquad (4.3)$$

3. Kibria (2003) proposed an estimator by taking the geometric mean of $\hat{k}_i$, which produced the following estimator:

$$\hat{k}_{GM} = \frac{\hat{\sigma}^2}{(\prod_{i=1}^{p} \hat{\alpha}_i^2)^{\frac{1}{p}}} . \qquad (4.4)$$

4. Muniz and Kibria (2009) proposed estimators by taking geometric mean and square root of estimator proposed by Alkhamisi and Shukur (2006). Suppose $m_i = \sqrt{\frac{\hat{\sigma}^2}{\hat{\alpha}_i^2}}$, then following estimators were proposed:

$$\hat{k}_{KM4} = \left(\prod_{i=1}^{p} \frac{1}{m_i}\right)^{\frac{1}{p}} . \qquad (4.5)$$

and

$$\hat{k}_{KM5} = \left(\prod_{i=1}^{p} m_i\right)^{\frac{1}{p}} . \qquad (4.6)$$

5. Estimator based on Alkhamisi and Shukur (2006) (thereafter $\hat{k}_{AS}$ or AS), proposed the following estimator:

$$\hat{k}_{AS} = \left(\frac{\hat{\sigma}^2}{max(\hat{\alpha}_i^2)} + \frac{1}{\lambda_i}\right), \ i = 1,2,\ldots,p. \qquad (4.7)$$

These are the few among many estimators suggested by researchers that will be compared in the study. For more in the estimation of *k*, I refer our readers to Kibria (2003), Khalaf and Shukur (2005), Muniz and Kibria (2009), Alkhamisi and Shukur

(2006), Khalaf (2012), Aslam (2014), Dorugade (2013) and very recently Kibria and Banik (2015) among others.

From (2.6), we know the MSE of Liu estimator in canonical form is given by,

$$\text{MSE}\ (\hat{\alpha}_d) = \sigma^2 \sum_{i=1}^{p} \frac{(\lambda_i + d)^2}{\lambda_i(\lambda_i+1)^2} + (d-1)^2 \sum_{i=1}^{p} \frac{\alpha_i^2}{(\lambda_i+1)^2}\ .$$

Liu (1993) suggested that the MSE($\hat{\alpha}_d$) is minimized at

$$d_i = \frac{\alpha_i^2 - \sigma^2}{\alpha_i^2 + \left(\sigma^2/\lambda_i\right)}, \qquad i = 1,2,\ldots,p. \tag{4.8}$$

Now, considering the canonical form in (2.1), the estimate for two parameter biased estimator is obtained as,

$$\hat{\alpha}_{k,d} = (\mathbf{\Lambda} + k\mathbf{I})^{-1}(\mathbf{Q}'\mathrm{y} - d\ \hat{\alpha})\ . \tag{4.9}$$

The relationship between linear regression model and orthogonal model is as,

$\hat{\beta}_{k,d} = \mathbf{Q}\ \hat{\alpha}_{k,d}$. MSE($\hat{\alpha}_{k,d}$) is obtained as,

$$\text{MSE}\ (\hat{\alpha}_{k,d}) = \sigma^2 \sum_{i=1}^{p} \frac{(d-\lambda_i)^2}{\lambda_i(\lambda_i+k)^2} + \sum_{i=1}^{p} \frac{(d+k)^2\alpha_i^2}{(\lambda_i+k)^2}\ . \tag{4.10}$$

It can be shown that (4.10) is minimized at

$$d_{opt} = \frac{\sum_{i=1}^{p} \frac{\sigma^2 - k\alpha_i^2}{(\lambda_i+k)^2}}{\sum_{i=1}^{p} \frac{(\lambda_i\alpha_i^2 + \sigma^2)}{\lambda_i(\lambda_i+k)^2}}\ . \tag{4.11}$$

As mentioned earlier that the two parameter biased estimator has less MSE than ridge regression estimator, also it allows larger values of *k* and thus can fully address the problem of ill conditioning**.** The understanding of the superior performance of two parameter biased estimator over ridge regression estimator can be theoretically explained as follows, we know that adding a value of k deals with ill conditioning of $\mathbf{X'X}$ in the model but practically ridge regression does not allow a very large value of k as it creates a bias. Because of this bias the problem of ill conditioning is not fully addressed. In the two parameter biased estimator of $\hat{\beta}_{k,d}$, k can be used exclusively to control the ill-conditioning of $\mathbf{X'X} + k\mathbf{I}$, inevitably some bias is generated and hence the second parameter d is used to improve the fit . I choose the ridge regression estimators discussed earlier from equation (4.2) – (4.7). After the k is selected, we can use d_{opt} to choose d from (4.11). Thus the two parameters in $\hat{\beta}_{k,d}$ are selected.

4.2. Monte Carlo Simulation.

In this section, I want to use a simulation study to illustrate the behavior of all the estimators discussed in section 4.1. The simulation is carried out under different degrees of multicollinearity, following McDonald and Galarneau (1975) which was also adopted by Gibbons (1981) and Kibria (2003).The explanatory variables were generated using the following equation:

$$x_{ij} = (1-\gamma^2)^{\frac{1}{2}} z_{ij} + \gamma z_{ip}, \; i = 1,2,\ldots,n \; ; \; j = 1,2,\ldots,p, \tag{4.12}$$

where z_{ij} are independent standard normal pseudo-random numbers, and γ^2 is the theoretical correlation between any two explanatory variables. These variables are standardized so that $\mathbf{X'X}$ and $\mathbf{X}'y$ are in correlation forms. The n observations for the dependent variable are determined by,

$$y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} + \varepsilon_i, \; i = 1,2,\ldots,n, \tag{4.13}$$

where ε_i are independent normal pseudo-random numbers with mean 0 and variance σ^2.

Since my primary interest lies in the performance of our proposed estimators according to strength of multicollinearity, I considered three sets of correlation corresponding to γ = 0.7, 0.8, 0.9. I also want to see the effect of the sample size on the number of regressors so I vary sample size between 15 and 50, and explanatory variables between 4 and 10. I investigate five values of sigma σ : 0.1, 0.5, 1, 4, 10; or equivalently, five signal-to-noise ratios: 100, 4, 1, 0.0625, 0.01. For each set of explanatory variables, I follow Newhouse and Oman (1971) conclusion for choosing the coefficient vector to minimize the MSE. When the MSE is function of $\beta, \sigma^2\ and\ k$ and explanatory variables are fixed, they suggested to choose the coefficient vector corresponding to the largest eigen value of **X′X** matrix subject to constraint β′β=1. One can also use the coefficient vector corresponding to the smallest eigen value but the results about performance of estimators do not differ significantly. The eigen values and the regression coefficients of **X′X** for different set on *n, p, γ* and ρ^2 are given in Table 4.1.

For the given values of *n, p, β, λ, γ* and ρ^2, the set of explanatory variables are generated. Then the experiment was repeated 2000 times by generating new error terms in (4.13). Then the values of ridge parameters *k* of the different estimators, *d* for Liu estimator and optimum *ds* for two parameter estimators and their corresponding estimators as well as average MSEs were estimated. The MSEs for the estimators are calculated as follows

$$\text{MSE}(\hat{\alpha}) = \frac{1}{2000}\sum_{r=1}^{2000}(\hat{\alpha}_{(r)} - \alpha)'(\hat{\alpha}_{(r)} - \alpha). \tag{4.14}$$

In this simulation study, twelve estimators are compared and their simulated MSE are presented in Tables 4.2-4.13 respectively. For a more in depth idea about which estimator performs uniformly better than LSE can be obtained from Tables 4.14 – 4.16. Along with MSEs, average values of *k,* standard deviation of *k* and the percentage for which the given estimator out performs LSE are provided. The twelve estimators compared are:

1. LSE: Least square estimator.
2. HKB: Ridge regression with $\hat{k}_{HKB}$
3. LW: Ridge regression with $\hat{k}_{LW}$
4. GM: Ridge regression with $\hat{k}_{GM}$
5. KM4: Ridge regression with $\hat{k}_{KM4}$
6. KM5: Ridge regression with $\hat{k}_{KM5}$
7. AS: Ridge regression with $\hat{k}_{AS}$
8. TPHKB: Two parameter biased estimator with $\hat{k}_{HKB}$, $\hat{\beta} = \hat{\beta}_{LS}$ and d_{opt}
9. TPGM: Two parameter biased estimator with $\hat{k}_{GM}$ and $\hat{\beta} = \hat{\beta}_{LS}$ and d_{opt}
10. TPKM5: Two parameter biased estimator with $\hat{k}_{KM5}$ and $\hat{\beta} = \hat{\beta}_{LS}$ and d_{opt}
11. TPAS: Two parameter biased estimator with $\hat{k}_{AS}$ and $\hat{\beta} = \hat{\beta}_{LS}$ and d_{opt}
12. Liu: Liu estimator with optimum d

Table 4.1: Values of λ and β used in simulation for n = 50 and different p

γ	0.7	0.8	0.9	0.7	0.8	0.9
n		*n=50, p=4*			*n=50, p=10*	
λ_1	408.24	378.011	338.597	295.141	344.529	395.777
λ_2	27.3165	19.2807	10.1746	48.5102	34.2557	18.0848
λ_3	21.0514	14.8643	7.84855	39.0282	27.749	14.7464
λ_4	10.1791	6.69967	3.13001	31.5056	22.1267	11.6163
λ_5				25.2457	17.8418	9.42814
λ_6				21.1363	14.8799	7.83493
λ_7				19.5329	13.8011	7.29219
λ_8				17.3463	12.154	6.37162
λ_9				10.5365	7.23757	3.76573
λ_{10}				8.04895	4.9286	2.10146
β_1	-0.3897	-0.405	-0.4286	-0.271	-0.2849	-0.2977
β_2	-0.4015	-0.4152	-0.4365	-0.2063	-0.2359	-0.2655
β_3	-0.4353	-0.4442	-0.4582	-0.2904	-0.3002	-0.3085
β_4	-0.7053	-0.6828	-0.6448	-0.3923	-0.3693	-0.3485
β_5				-0.2645	-0.2765	-0.2896
β_6				-0.271	-0.282	-0.2938
β_7				-0.3012	-0.3051	-0.3093
β_8				-0.2845	-0.2893	-0.2966
β_9				-0.2545	-0.2674	-0.2828
β_{10}				-0.5157	-0.4822	-0.4365

Table 4.2: Estimated MSE with n = 15, p =4, γ = 0.7

sigma	LS	HKB	LW	GM	KM4	KM5	AS	TPHKB	TPGM	TPKM5	TPAS	LIU
0.1	0.0099914	0.0096636	0.0099753	0.0131883	0.0071161	0.006417	0.0073855	0.0051448	0.0062297	0.0056696	0.0051374	0.0036321
0.5	0.2497838	0.1579487	0.2376406	0.1505154	0.1872188	0.1316426	0.1687266	0.1168612	0.1206345	0.1145756	0.1205216	0.0913687
1	0.9991353	0.4852051	0.8653898	0.5663049	0.7727402	0.50718	0.5704635	0.4424265	0.4480063	0.4406052	0.4593035	0.3576757
4	15.986164	7.429741	11.907569	10.309655	13.109458	8.333695	9.040674	6.969378	6.895514	7.000231	7.054533	5.745691
10	99.91353	45.80064	72.06577	69.54592	84.51883	52.99509	58.57629	42.65201	42.06875	42.7154	43.05988	34.62554

Table 4.3: Estimated MSE with n = 15, p =4, γ = 0.8

sigma	LS	HKB	LW	GM	KM4	KM5	AS	TPHKB	TPGM	TPKM5	TPAS	LIU
0.1	0.0149578	0.0142023	0.0149196	0.0117327	0.0098214	0.0084132	0.0095694	0.0072895	0.0079226	0.0073661	0.0074577	0.0053886
0.5	0.3739443	0.2128902	0.3441016	0.2152643	0.2550967	0.1972852	0.2285522	0.1705009	0.168675	0.1652844	0.1803241	0.1346386
1	1.4957771	0.7116152	1.2111367	0.8746838	1.0431655	0.7869669	0.8420659	0.6617411	0.6591206	0.6576636	0.696571	0.5408818
4	23.932433	11.426643	17.242251	16.984002	17.895037	13.749235	15.085919	10.529856	10.393935	10.511884	10.797726	8.642682
10	149.57771	69.78893	106.47177	112.01861	115.43902	87.36506	95.34936	64.26612	63.45536	63.9611	65.93505	52.79462

Table 4.4: Estimated MSE with n = 15, p =4, γ = 0.9

sigma	LS	HKB	LW	GM	KM4	KM5	AS	TPHKB	TPGM	TPKM5	TPAS	LIU
0.1	0.0308692	0.0276922	0.030685	0.0174065	0.0201844	0.0161961	0.0209403	0.0146239	0.0140715	0.0135773	0.0161779	0.0109332
0.5	0.7717296	0.3795884	0.6431317	0.4331345	0.4877646	0.4228886	0.5235205	0.3367361	0.3270864	0.3262764	0.387198	0.2712389
1	3.086918	1.393182	2.255883	1.83192	1.924166	1.727677	2.066281	1.287673	1.270758	1.274197	1.449438	1.067342
4	49.39069	22.48665	33.66357	34.98506	31.08769	30.24301	35.08837	20.41577	20.17477	20.22026	22.26608	16.94436
10	308.6918	145.0339	208.3476	247.3803	204.9827	208.3298	233.1699	130.9143	129.0636	129.3599	143.0726	108.503

Table 4.5: Estimated MSE with n = 15, p =10, γ = 0.7

sigma	LS	HKB	LW	GM	KM4	KM5	AS	TPHKB	TPGM	TPKM5	TPAS	LIU
0.1	0.046798	0.0367564	0.0449406	0.0504945	0.0301879	0.0341312	0.0357979	0.0250479	0.0254408	0.0225386	0.029043	0.0169642
0.5	1.169951	0.6419494	0.8767881	0.9961206	0.7611975	0.8354619	0.8682475	0.5528522	0.5351144	0.5354022	0.6850901	0.4183946
1	4.679804	2.634919	3.274944	3.893886	3.04824	3.292386	3.352814	2.158172	2.092785	2.120595	2.606198	1.659148
4	74.87686	41.9769	50.30539	62.59726	49.30746	52.40536	52.35224	33.63644	32.75617	33.1739	38.8979	25.91715
10	467.9804	275.8618	317.3916	421.0058	317.7537	353.4829	352.3179	216.6397	209.405	212.5535	252.7841	166.5883

Table 4.6: Estimated MSE with n = 15, p =10, γ = 0.8

Sigma	LS	HKB	LW	GM	KM4	KM5	AS	TPHKB	TPGM	TPKM5	TPAS	LIU
0.1	0.0719235	0.0511077	0.0671433	0.0651366	0.0456673	0.0538282	0.0586689	0.0372895	0.034643	0.0329088	0.0474027	0.0255868
0.5	1.7980877	1.0060598	1.2464344	1.51001	1.1214231	1.3375363	1.4361408	0.8303941	0.7946134	0.8015447	1.1409943	0.6363006
1	7.192351	4.222495	4.709176	6.283716	4.585819	5.556869	5.886123	3.305163	3.189974	3.225929	4.460466	2.565731
4	115.07761	70.12109	74.91071	106.35673	74.02306	93.24769	97.48011	53.47523	51.6721	52.20775	69.87783	41.88652
10	719.2351	436.202	465.3851	679.7789	467.2957	595.0823	615.6397	331.7632	320.5374	323.4525	435.8849	258.3578

Table 4.7: Estimated MSE with n = 15, p =10, γ = 0.9

Sigma	LS	HKB	LW	GM	KM4	KM5	AS	TPHKB	TPGM	TPKM5	TPAS	LIU
0.1	0.1529433	0.0918929	0.1307873	0.1276389	0.110417	0.1207058	0.1364077	0.0757656	0.0670437	0.0665439	0.1138187	0.0538967
0.5	3.823582	2.249317	2.447658	3.359379	2.770679	3.166831	3.505374	1.73034	1.667656	1.673758	2.851258	1.364865
1	15.294327	9.135178	9.477046	13.425396	10.804042	12.582917	13.777292	6.884303	6.649564	6.679168	10.921922	5.484152
4	244.70923	150.31716	151.87863	229.57841	174.72612	212.79638	229.73494	111.2674	107.79617	108.12833	175.16839	88.38628
10	1529.4327	948.2393	947.8756	1476.0187	1086.8153	1363.6703	1456.1289	699.2951	677.4846	678.6354	1108.8051	554.5174

Table 4.8: Estimated MSE with n = 50, p =4, γ = 0.7

Sigma	LS	HKB	LW	GM	KM4	KM5	AS	TPHKB	TPGM	TPKM5	TPAS	LIU
0.1	0.0020854	0.0020716	0.002085	0.0111715	0.0019603	0.0019372	0.0020471	0.0016525	0.0018626	0.0017617	0.0016391	0.0007601
0.5	0.0521353	0.0451748	0.0518841	0.0529514	0.0498268	0.0344652	0.0492616	0.0267332	0.0324605	0.0273893	0.0267306	0.0190356
1	0.2085411	0.1402165	0.204994	0.1448037	0.200697	0.1289923	0.1770852	0.1001898	0.1113137	0.0999975	0.1018641	0.0763115
4	3.336658	1.528862	3.041341	1.927381	3.247668	1.854959	1.756108	1.45235	1.470178	1.491496	1.470347	1.192816
10	20.85411	9.373097	18.168696	13.010668	20.426873	10.952904	11.361576	8.850362	8.754806	9.111642	8.877008	7.168491

Table 4.9: Estimated MSE with n = 50, p =4, γ = 0.8

sigma	LS	HKB	LW	GM	KM4	KM5	AS	TPHKB	TPGM	TPKM5	TPAS	LIU
0.1	0.0030391	0.0030086	0.0030381	0.0075621	0.0027307	0.0023205	0.0029178	0.0019287	0.0023604	0.0021346	0.0019158	0.0011038
0.5	0.0759784	0.0620279	0.0753433	0.0579413	0.0703267	0.0461828	0.0690433	0.0371801	0.0415205	0.0367882	0.0375378	0.0277636
1	0.3039136	0.1844956	0.2951541	0.1796017	0.2846096	0.1744956	0.2390452	0.1401776	0.1456594	0.138994	0.1443675	0.1092562
4	4.862617	2.232179	4.281019	2.861068	4.639933	2.57811	2.578353	2.106984	2.105738	2.161598	2.131829	1.740549
10	30.39136	13.85273	25.79729	19.99025	29.2936	15.71448	17.0922	12.997	12.87357	13.27169	13.0375	10.6188

Table 4.10: Estimated MSE with n = 50, p =4, γ = 0.9

sigma	LS	HKB	LW	GM	KM4	KM5	AS	TPHKB	TPGM	TPKM5	TPAS	LIU
0.1	0.0060537	0.0059273	0.0060492	0.0061443	0.0047564	0.0037309	0.0051914	0.0030589	0.003498	0.0031532	0.0030656	0.0021582
0.5	0.1513422	0.1066449	0.1481297	0.087148	0.1261314	0.0818556	0.118864	0.0701819	0.069795	0.0675112	0.0720618	0.0537323
1	0.6053686	0.3128591	0.5648157	0.3342905	0.5182376	0.3166532	0.3927557	0.2672132	0.2641652	0.2647211	0.2797057	0.2121683
4	9.685898	4.401896	7.982438	5.807901	8.623953	4.845213	5.203612	4.108073	4.07844	4.166959	4.139735	3.400016
10	60.53686	27.35599	48.81411	41.26191	55.1935	30.91536	35.1995	25.35054	25.10372	25.55998	25.6013	20.87807

Table 4.11: Estimated MSE with n = 50, p =10, γ = 0.7

sigma	LS	HKB	LW	GM	KM4	KM5	AS	TPHKB	TPGM	TPKM5	TPAS	LIU
0.1	0.0057045	0.0055999	0.0057021	0.0504328	0.0055077	0.005087	0.0055631	0.0039923	0.0052433	0.0044939	0.0039707	0.0020024
0.5	0.1426117	0.1051151	0.1412005	0.1551477	0.1383003	0.0845454	0.1333832	0.0780738	0.0896033	0.0766101	0.0801394	0.0501204
1	0.5704467	0.327528	0.5517138	0.4475692	0.5543025	0.3317808	0.4764835	0.2974342	0.3065563	0.2969289	0.3157866	0.2021973
4	9.127148	4.638433	8.1892	6.174592	8.903293	5.092452	5.185746	4.416422	4.305396	4.578231	4.600128	3.186709
10	57.04467	29.22233	50.32267	39.98599	55.77135	31.54804	32.59912	27.56371	26.66685	28.53667	28.55483	19.76678

Table 4.12: Estimated MSE with n = 50, p =10, γ = 0.8

sigma	LS	HKB	LW	GM	KM4	KM5	AS	TPHKB	TPGM	TPKM5	TPAS	LIU
0.1	0.0084228	0.0081829	0.0084165	0.0274848	0.0078995	0.0057491	0.0079422	0.0050875	0.0068132	0.0054223	0.005091	0.0029807
0.5	0.2105689	0.142061	0.206779	0.1723412	0.199381	0.1195448	0.1870799	0.1128457	0.1145735	0.1081953	0.1182638	0.0733781
1	0.8422755	0.4552908	0.7944422	0.589147	0.8000254	0.4683981	0.6451712	0.4240638	0.4147371	0.4231201	0.4589172	0.2921224
4	13.476407	6.99257	11.664827	9.308667	12.874578	7.446916	7.690046	6.539806	6.317763	6.739574	6.823498	4.717
10	84.22755	44.43721	72.14139	62.51251	80.83463	46.97573	49.31967	41.27454	39.78072	42.27507	42.78561	29.50965

Table 4.13: Estimated MSE with n = 50, p =10, γ = 0.9

sigma	LS	HKB	LW	GM	KM4	KM5	AS	TPHKB	TPGM	TPKM5	TPAS	LIU
0.1	0.0171385	0.0160878	0.0171032	0.0171974	0.0144543	0.0098896	0.0137651	0.009742	0.0101154	0.0089067	0.0098869	0.0059323
0.5	0.4284616	0.2488431	0.4064802	0.2882771	0.3689336	0.2365009	0.3186997	0.2197035	0.2028688	0.2086741	0.2403167	0.1475976
1	1.7138465	0.8863237	1.4985533	1.1480594	1.4888807	0.9410616	1.094845	0.8323048	0.7914599	0.8283718	0.920826	0.5886096
4	27.421545	14.219395	21.991581	19.189275	24.204195	15.180668	15.447081	12.998545	12.501353	13.132833	13.694628	9.434745
10	171.3846	90.22121	137.76715	126.50171	152.86653	96.26811	98.91006	82.2925	79.04653	82.74551	86.72064	59.44376

Table 4.14: Estimated MSE, average *k,* s.d. of *k* with *n* = 30, *p* =6, *γ* = 0.7

sigma	LS	HKB	LW	GM	KM4	KM5	AS	LIU	TPHKB	TPGM	TPKM5	TPAS
0.1	0.005169	0.00509	0.005165	0.014476	0.004765	0.003741	0.004956	0.001869	0.003189	0.004116	0.003537	0.003174
		(0.059663, 0.001006)	(0.002583, 0.000325)	(14.70597, 19.41113)	(0.343651, 0.12918)	(3.441961, 1.691243)	(0.096147, 0.062497)					
		100	100	41.1	100	97.2	100	100	100	100	100	100
0.5	0.129224	0.09813	0.127123	0.097298	0.121113	0.075808	0.117125	0.047013	0.066782	0.071607	0.065659	0.067821
		(1.3326, 0.142297)	(0.063327, 0.027755)	(24.80631, 37.68962)	(0.264596, 0.098772)	(4.461042, 2.215371)	(0.337185, 0.06573)					
		100	100	82.8	100	99	100	100	100	100	100	100
1	0.516895	0.295662	0.492707	0.332272	0.487246	0.293787	0.404687	0.188729	0.255735	0.259074	0.254466	0.264526
		(4.117401, 1.002566)	(0.191533, 0.114045)	(30.11345, 37.27834)	(0.239132, 0.090408)	(4.946125, 2.377419)	(1.099273, 0.184969)					
		99.95	100	88.05	100	98.8	99.95	100	100	100	100	100
4	8.27032	4.045102	7.274132	5.384369	7.898347	4.480938	4.5821	2.9468	3.880841	3.821265	3.959094	3.939506
		(14.84304, 10.75586)	(0.601523, 0.67383)	(52.90355, 79.00465)	(0.183814, 0.071468)	(6.487171, 3.290223)	(11.13381, 16.35435)					
		99.55	100	84.45	100	97.9	96	100	100	100	100	100
10	51.6895	25.58743	44.88862	36.46683	49.68654	27.78412	29.80301	18.48397	24.38181	23.86778	24.79131	24.7117
		(18.33704, 18.98134)	(0.714358, 1.055777)	(79.27496, 125.5558)	(0.156943, 0.064798)	(7.803691, 4.287955)	(370.7379, 10728.73)					
		99.25	100	80.45	100	97.25	93.85	100	100	100	100	100

Table 4.15: Estimated MSE, average *k*, s.d. of *k* with n = 30, p =6, γ = 0.8

sigma	LS	HKB	LW	GM	KM4	KM5	AS	LIU	TPHKB	TPGM	TPKM5	TPAS
0.1	0.00772	0.007536	0.007712	0.009061	0.006696	0.00475	0.007009	0.002767	0.004173	0.005031	0.004325	0.004178
		(0.05951, 0.000983)	(0.00245, 0.000458)	(10.0616, 11.87183)	(0.40939, 0.153969)	(2.87419, 1.342212)	(0.13867, 0.100266)					
		100	100	72.75	100	98.55	100	100	100	100	100	100
0.5	0.193005	0.133358	0.187882	0.124091	0.17234	0.107451	0.162221	0.069253	0.097041	0.095998	0.093982	0.099737
		(1.26493, 0.16284)	(0.06727, 0.041156)	(18.0407, 25.77398)	(0.31117, 0.116853)	(3.80373, 1.890526)	(0.37912, 0.102015)					
		99.95	100	87.4	99.95	97.4	99.95	100	100	100	100	100
1	0.772019	0.407981	0.714584	0.469636	0.697078	0.414901	0.54007	0.27583	0.371452	0.364961	0.369513	0.388307
		(3.64658, 1.006989)	(0.20708, 0.157627)	(23.5827, 40.22962)	(0.27566, 0.103902)	(4.31292, 2.232464)	(1.13507, 0.198965)					
		99.95	100	89.2	100	97.85	99.95	100	100	100	100	100
4	12.3523	6.110453	10.53628	8.276557	11.376	6.602244	6.945109	4.428472	5.807665	5.700715	5.889751	5.906828
		(10.8656, 9.563891)	(0.53368, 0.722205)	(38.2894, 58.54542)	(0.21817, 0.085112)	(5.49570, 2.844425)	(8.55986, 11.48202)					
		98.65	100	83.7	100	96.2	94.8	100	100	100	100	100
10	77.20187	38.19354	65.3593	55.4644	71.99711	41.06761	44.07596	27.38699	36.07819	35.35801	36.43697	36.63483
		(12.5978, 13.1616)	(0.59567, 1.017023)	(63.3459, 144.2615)	(0.18281, 0.077798)	(6.80775, 4.124177)	(373.032, 11347.91)					
		98.75	100	79.75	100	94.95	94.65	100	100	100	100	100

Table 4.16: Estimated MSE, average *k,* s.d. of *k* with *n* = 30, *p* =6, *γ* = 0.9

sigma	LS	HKB	LW	GM	KM4	KM5	AS	LIU	TPHKB	TPGM	TPKM5	TPAS
0.1	0.015896	0.015083	0.015862	0.010644	0.01159	0.008824	0.011565	0.005737	0.008379	0.008221	0.007861	0.008528
		(0.05903, 0.001088)	(0.002302, 0.000718)	(5.94673, 8.652064)	(0.54461, 0.203563)	(2.17560, 1.101852)	(0.27493, 0.229402)					
		100	100	88.9	99.35	95.7	98.95	100	100	100	100	99.9
0.5	0.397398	0.232756	0.371915	0.245112	0.30433	0.216285	0.266974	0.142213	0.196652	0.186199	0.189202	0.208166
		(1.10031, 0.209141)	(0.08144, 0.075145)	(10.2829, 15.64252)	(0.41195, 0.153058)	(2.87138, 1.427968)	(0.51559, 0.230153)					
		99.9	100	87.05	99.8	93.75	97.95	100	100	100	100	99.95
1	1.589593	0.807853	1.378215	1.022296	1.24303	0.87692	0.940344	0.573312	0.761482	0.737442	0.75578	0.80943
		(2.66120, 1.075268)	(0.20777, 0.228632)	(12.8825, 19.27105)	(0.370784, 0.139741)	(3.20172, 1.622612)	(1.19422 0.332469)					
		98.55	100	844.5	99.85	93.05	96.55	100	100	100	100	100
4	25.43348	12.71991	20.56304	17.82145	20.68535	14.16653	14.61344	9.070651	11.82114	11.54778	11.83807	12.14534
		(5.99370, 6.254837)	(0.42112, 0.738302)	(21.6235, 44.11132)	(0.29274, 0.110116)	(4.07913, 2.233101)	(6.03878, 8.509219)					
		96.95	100	79.55	99.95	90.2	91.45	100	100	100	100	100
10	158.9593	78.41805	126.2277	117.4674	133.3351	89.05708	91.10128	55.44889	72.71819	70.7814	72.40383	74.72872
		(6.79058, 8.346396)	(0.50858, 1.167226)	(34.0470, 47.7129)	(0.23686, 0.096607)	(5.15030, 2.743206)	(4318.49, 181606.3)					
		97.2	100	77.75	99.95	89.55	91.5	100	100	100	100	99.95

4.3. Simulated Results.

I will discuss the simulation results in this section. A comparison will be made among the estimators based on the smaller MSE criterion for different values of *p, k, d and ρ.*

4.3.1. Performance as a function of σ.

From Tables 4.2-4.13, we can compare MSEs of the estimators as a function of the variance of the errors (σ^2). When the value of σ increases, the MSE of the estimators also increases. For all values of σ, the ridge regression estimators and the two parameter biased estimators have smaller MSE compared with the LSE. However, the performance of the two parameter biased estimators is better than the performance of the corresponding ridge regression estimators. I also observe that the MSE of Liu estimator is the smallest which is a special case of two parameter estimator for *d* = 0. This behavior was almost constant for any sample size and number of variables considered.

Amongst ridge estimators, KM5 performs better than others estimators for $\sigma < 1$, however for $\sigma > 1$, HKB performs better than the rest of ridge estimators closely followed by KM5. But to note here, from Tables 4.14-4.16, we see from the percentages that the MSE of KM5 is not always less than that of LSE. There is decrease in the percentage times the MSE of KM5 outperform the MSE of LSE with increase in sigma. For $\sigma < 1$, amongst HKB and KM5, better choice is that to choose HKB as the average *k* and s.d of *k* corresponding to HKB are smaller and thus it is more reliable. Mostly performance of all the estimator are between HKB and LW except that of KM4 for larger values of σ. All the two parameter biased estimators perform better than ridge estimators, TPGM performs better than rest of the two parameter biased estimators for $\sigma > 1$ and for $\sigma < 1$

TPKM5 performs better than the rest the two parameter biased estimators. Liu estimator out performs all the estimators in all cases.

For given γ = 0.70 and n= 15, the performance of estimators as a function of the standard deviation of the errors for p= 4 and p= 10 are provided in Figures 4.1 and 4.2 respectively. From these figures I observe that as the standard deviation increases, the MSE also increases.

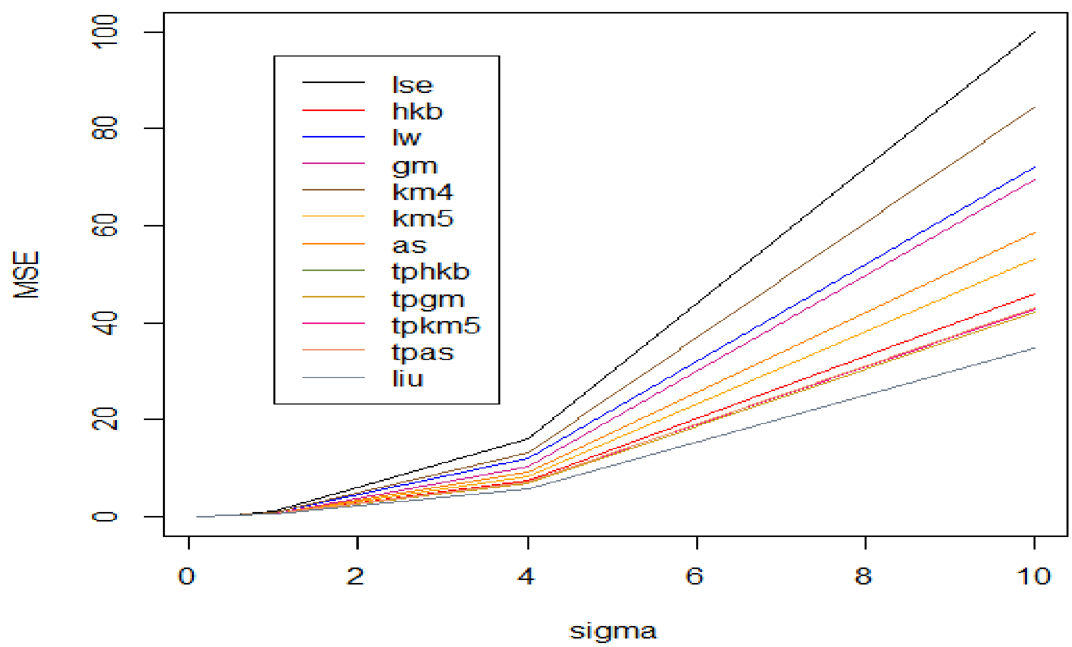

Figure 4.1: Performance of estimators as a function of σ, for p = 4, γ = 0.70 and n= 15

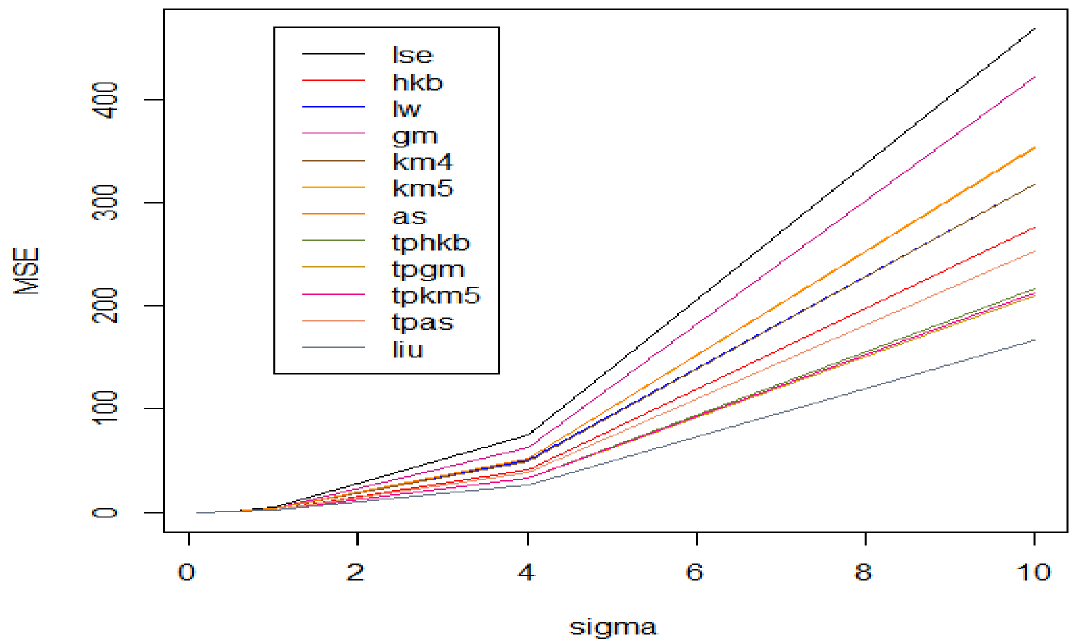

Figure 4.2: Performance of estimators as a function of σ, for p = 10 ,γ = 0.70 and n= 15

4.3.2. Performance as a function of γ.

From, Tables 4.2-4.13, I observe that for smaller sigma (σ = 0.01) the change in the correlation between the explanatory variables had almost no effect on the MSEs. In all situations they remained almost the same for any sample size or number of parameters, and their MSEs are very small. When σ increases, the higher correlation between the independent variables, results in an increase of the MSE of the all estimators.

In general, all the two parameter biased estimators except TPAS perform better than rest of estimators other than KM4. For given σ =1 and p = 10, the performance of estimators as a function of the correlation between the explanatory variables for n = 15 and n = 50 are provided in Figures 4.3 and 4.4 respectively. From these figures I observed that as correlation increases, the MSE also increases. All of the estimators have smaller MSE compared with LSE. Liu estimator again outperforms all other estimators, as special case of two parameter biased estimator. The MSEs of all the two parameter biased estimators are very close.

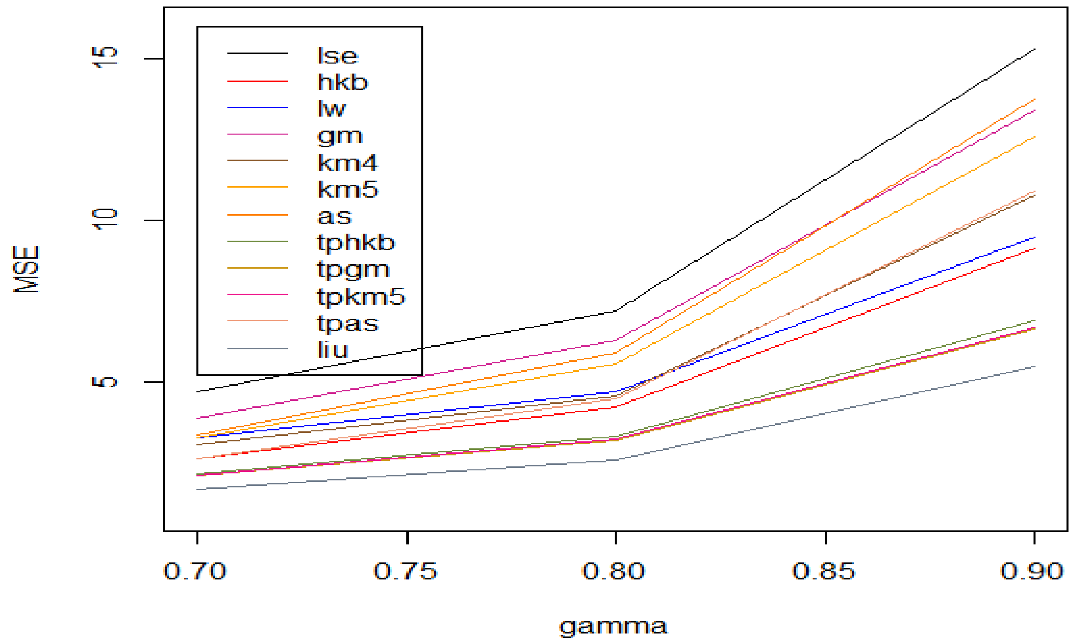

Figure 4.3: Performance of estimators as a function of γ, for n = 15

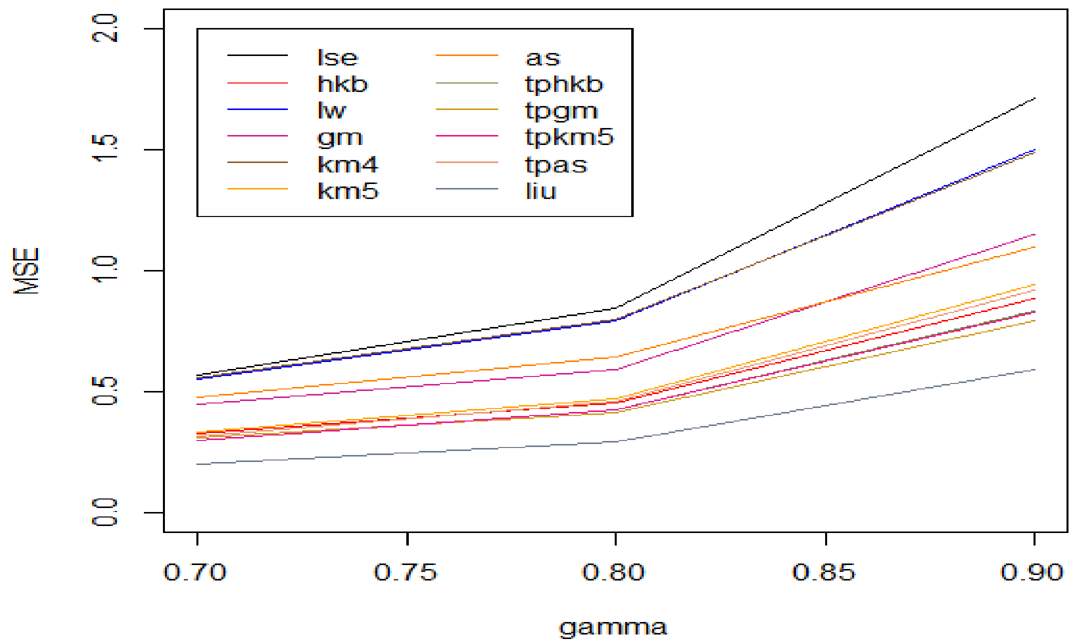

Figure 4.4: Performance of estimators as a function of γ, for n = 50

4.3.3. Performance as a function of *n* and *p*.

From Tables 4.2 to 4.13, I observed that, in general, when the sample size increases, the MSE decreases, or remained the same. Even for the large values of γ and σ, if I increase the sample size, the MSE of estimators decrease. Again in this situation, as n increases the performance of TPHKB, TPGM, TPKM5, TPAS and Liu better than the rest of the ridge estimators, also TPGM performs better amongst all the two parameter biased estimator and Liu performs better overall. From Tables 4.14 to 4.16, we also observe that amongst all the Two parameter biased estimators TPHKB and TPKM5 perform equally good but the s.d of k for TPKM5 is smaller than that of TPHKB for $\sigma > 1$, thus TPKM5 is recommended to be used than TPHKB.

For given σ and γ, as the number of explanatory variables increase, the MSE of proposed estimator increases and the performance of estimators is similar to that when compared for different σs.

.

V. SUMMARY AND CONCLUDING REMARKS

The purpose of this research is two fold. Firstly I made an analytical comparison of LASSO, ridge regression estimator, Liu estimator and least square estimator. For selected values of *k, p* and Δ^2, I compared the risks (MSE) and relative efficiencies and presented them in tabular form in Tables 3.1 to 3.12. I, also provided risk graphs for a visual comparison.

Based on the analyses of risks and relative efficiencies, I found that none of estimator uniformly dominate each other. I compared all the estimators for their dominance criteria in terms of Δ^2, each criteria is found to be an increasing function of *p*. I found LSE mostly being dominated uniformly by rest of the estimators over a wider sub-space except for smaller values of *p* at 3 and 4. Also as value of *k* increases the sub-space where ridge estimator is dominated by LASSO and Liu estimator, where *d* equals *k,* increases. The results are similar for Liu estimator for decreasing values of *d,* at small value of *d,* LASSO and ridge estimator with *k* equal to *d* dominate it over a wider sub-space. This phenomenon increases for LASSO with smaller values of *q,* as *p* increases. Neither estimator dominates one another uniformly except for LASSO at larger values of *p* and small *q*. Finally, neither LASSO, ridge regression nor Liu estimator perform uniformly better than one other.

Secondly, I compared the Ridge regression, Liu, two parameter estimator and least square estimators under multicollinear model with error distribution being normal. The performance of the estimators depends on the variance of the random error, the correlations among the explanatory variables, the sample size and the unknown coefficients vectors. Based on the simulation study, some conclusions might be drawn. However, these conclusions might be restricted to the set of experimental conditions which are investigated. I used the MSE criteria to measure the goodness of the estimators. Increase in the value of σ and the increase of the correlation between the independent variables have a negative effect on the MSE, in the sense that it also increases. When the sample size increases the MSE decreases, even when the correlation between the independent variables and σ are large. The two parameter biased estimator gave better performance than the corresponding ridge regression estimator. Comparing the choice of k, k_{GM} performed better than rest of estimators for σ < 1 and k_{KM5} performed better for σ >1. In conclusion, two parameter biased estimator with appropriate k might be considered over ridge regression, as observed from simulated results. Finally, Liu estimator is a special case of two parameter biased estimator at d =0, outperforms every estimator. However, more study is required before making any definite statement.

For future researcher one may consider comparing the estimators under different error distribution, also the experiment was restricted to 2000 replications in the study which can be increased to attain more precise result in future. One variation that can be applied to the two parameter biased estimator is to replace $\tilde{\beta}$ by $\hat{\beta}_{ridge}$ instead of $\hat{\beta}_{LS}$. Also, lot of ridge estimators are available in the literature which can be considered and a comparison can be made amongst them. Variations such as the mean, geometric mean can be considered in estimating d, the shrinkage parameter for Liu estimator.

LIST OF REFERENCES

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

Aslam, M. (2014). “Performance of Kibria’s Method for the Heteroscedastic Ridge Regression Model: Some Monte Carlo Evidence”. *Communications in Statistics—Simulation and Computation, 43: 673-686.*

Arashi, M. and Valizadeh, T. (2015).“Performance of Kibria’s methods in partial linear ridge regression model”. *To appear in Statistical Papers*.

Donoho, D. and Johnstone, I. (1994). “Ideal spatial adaption by wavelet shrinkage”. *Biometrika*, 81, 425-455..

Dorugade, A. V. (2013). “New ridge parameters for ridge regression”. *Journal of the Association of Arab Universities for Basic and Applied Sciences.* Vol.15. 94-99.

Galton, Sir Francis. (1885). “Regression towards mediocrity in heredity stature”. *Journal of Anthropological Institute 15*: 246-263.

Gibbons, D.G. (1981). “A simulation study of some ridge estimators”. *Journal of the American Statistical Association* 76:131-139.

Golub, G. H., Health, M. and Wahaba, G. (1979). “Generalized cross-validation as a method for choosing a good ridge parameter”. *Technometrics*, 21: 215–223.

Hefnawy, A. E and Farag, A. (2014). “A Combined Nonlinear Programming Model and Kibria Method for Choosing Ridge Parameter Regression”. *Communications in Statistics—Simulation and Computation,* 43: 1442–1470, 2014

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

Hoerl, A.E., Kennard, R.W., Baldwin, K.F. (1975). “Ridge regression: some simulation”. *Communications in Statistics* 4: 105-123

Khalaf, G. and Shukur, G. (2005). “Choosing ridge parameters for regression problems”. *Communication in Statistics- Theory and Methods*, vol.34, pp.1177-1182.

Khalaf, G. (2012). “A proposed ridge parameter to improve the least squares estimator”. *Journal of Modern Applied Statistical Methods*, vol.11, pp.443-449.

Kibria B. M. G. (2003). “Performance of some new ridge regression estimators”. *Communications in Statistics- Theory and Methods*, 32:419-435.

Kibria, B. M. G. and Banik, S. (2015). “A Comprehensive Simulation Study on Some Ridge Regression Estimators and Their Performances”. To appear in *Journal of Modern Applied Statistical Methods*.

Lawless, J.F., Wang, P.(1976).” A simulation study of ridge and other regression estimators”. *Communications in Statistics* A5: 307-323.

Liu, K. (1993). “A new class of biased estimate in linear regression”. *Communications in Statistics- Theory and Methods*, Vol.22, No.2. 393-402

Liu, K. (2003). “Using Liu-Type Estimator to Combat Collinearity”. *Communications in Statistics - Theory and Methods,* Vol.32, No.5. 1009-1020

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

Muniz, Gisela and Kibria, B.M. G. (2009). “On Some Ridge Regression Estimators : An Empirical Comparisons”. *Communications in Statistics- Theory and Methods*, Vol.38, No.3. 621-630.

Mansson, K., Kibria, B.M. G., Shukur, G. (2012). “On Liu estimators for the logit regression model.” *Economic Modelling*. Vol.29 No.4., 1483–1488.

Newhouse, J. P. and Oman, S. D. 1971. “An evaluation of ridge estimators”. *Rand Corporation*. P-716-PR

Saleh, A. K. Md. E. (2006). “Theory of Preliminary Test and Stein-Type Estimation with Applications”. *Wiley, New York,* 2006.

Tibshirani, R. (1996). “Regression Shrinkage and Slection via the LASSO”. *Journal of the Royal Statistical Society. Series B(Methodological,* Vol.58, No.1*).* 267-288.