Effect of W, LR, and LM Tests on the Performance of Preliminary Test Ridge Regression Estimators

by
B. M. Golam Kibria
Florida International University, Miami, USA
and
A. K. Md. E. Saleh
Carleton University, Ottawa, Canada

SUMMARY

This paper combines the idea of preliminary test and ridge regression methodology, when it is suspected that the regression coefficients may be restricted to a subspace. The preliminary test ridge regression estimators ($PTRRE$) based on the Wald (W), Likelihood Ratio (LR) and Lagrangian Multiplier (LM) tests are considered. The bias and the mean square errors (MSE) of the proposed estimators are derived under both null and alternative hypotheses. By studying the MSE criterion, the regions of optimality of the estimators are determined. Under the null hypothesis, the $PTRRE$ based on LM test has the smallest risk followed by the estimators based on LR and W tests. However, the $PTRRE$ based on W test performs the best followed by the LR and LM based estimators when the parameter moves away from the subspace of the restrictions. The conditions of superiority of the proposed estimator for both ridge parameter k and departure parameter Δ are provided. Some graphical representations have been presented which support the findings of the paper. Some tables for maximum and minimum guaranteed relative efficiency of the proposed estimators have been provided. These tables allow us to determine the optimum level of significance corresponding to the optimum estimators among proposed estimators. Finally, we concluded that the optimum choice of the level of significance becomes the traditional choice by using the W test for all non-negative ridge parameter, k.

A*MS(1991) subject classification*: 62J07, 62F03
Key Words and Phrases: Dominance; Lagrangian Multiplier; Likelihood Ratio Test; MSE; Non-central Chisquare and F; Ridge Regression; Superiority; Wald Test.

1 Introduction

Consider the following linear regression model,

$$Y \sim N(X\beta, \sigma^2 I), \tag{1.1}$$

where Y is an $n \times 1$ vector of observations on the dependent variable, which follow a normal distribution with fixed mean, $X\beta$ and unknown variance, $\sigma^2 I$, β is an $p \times 1$ vector of unknown parameters, X is an $n \times p$ known design matrix of rank p $(n \geq p)$.

Our primary interest is to estimate the regression coefficient β when it is *apriori* suspected that β may be restricted to the subspace

$$H_0 : H\beta \;\; = \;\; h, \tag{1.2}$$

where H is an $q \times p$ known matrix of full rank $q(< p)$ and h is an $q \times 1$ vector of known constants.

The unrestricted (UR) least squares estimator ($URLSE$) of β is given by

$$\hat{\beta}^{UR} = C^{-1}X'Y, \tag{1.3}$$

where $C = X'X$ matrix. The corresponding MLE of σ^2 is given by

$$\hat{\sigma}^2_{UR} = \frac{(Y - X\hat{\beta}^{UR})'(Y - X\hat{\beta}^{UR})}{n}.$$

It is observed from (1.3) that the usual least squares estimator (LSE) of β depends heavily on the characteristics of the matrix $C = X'X$. If the C matrix is ill-conditioned (near dependency among various columns of C), then the least squares estimator (LSE) produce unduly large sampling variances. Moreover, some of the regression coefficients may be statistically insignificant with wrong sign and meaningful statistical inference become impossible for the researcher. Hoerl and Kennard (1970) found that multicollinearity is a common problem in the field of engineering. To resolves this problem, they suggested to use $C(k) = X'X + kI_p, \quad (k \geq 0)$ rather than C in the estimation of β. The resulting estimator of β are known as the Ridge Regression Estimator (RRE). Hoerl and Kennard (1970) considered the following *Unrestricted Ridge Regression Estimator* ($URRE$),

$$\hat{\beta}^{UR}(k) = (X'X + kI_p)^{-1}X'y = W\hat{\beta}^{UR}, \tag{1.4}$$

where $W = [I_p + kC^{-1}]^{-1}$ and $k \geq 0$ is the ridge or biasing parameter. The bias and mean squares error of the $URRE$ of β are

$$Bias(\hat{\beta}^{UR}(k)) = B_{UR}(k) \;\; = \;\; E(\hat{\beta}^{UR}(k) - \beta) = -kC^{-1}(k)\beta \quad \text{and}$$

$$MSE_{UR}(k) = \sigma^2 tr(WC^{-1}W') + k^2\beta' C^{-2}(k)\beta, \quad (1.5)$$

respectively. Though these estimators in (1.4) result in biased, for certain value of k, they yield minimum mean square error (MSE) compared to the ordinary least squares (OLS) estimator.

In order to reduce the pain of multicollinearity, the very well known restricted least squares (RLS) method of estimation are useful in practice. The restricted least squares estimator ($RLSE$) of β and σ^2 are

$$\hat{\beta}^{RE} = \hat{\beta}^{UR} - C^{-1}H'(HC^{-1}H')^{-1}(H\hat{\beta}^{UR} - h) \quad \text{and}$$

$$\hat{\sigma}^2_{RE} = \frac{(Y - X\hat{\beta}^{RE})'(Y - X\hat{\beta}^{RE})}{n}$$

respectively. Based on the $RLSE$, Sarkar (1992) proposed the following Restricted Ridge Regression Estimator ($RRRE$),

$$\hat{\beta}^{RE}(k) = W\hat{\beta}^{RE}. \quad (1.6)$$

The bias and mean squares error of the $RRRE$ of β are, respectively,

$$B_{RE}(k) = -W\eta - kC^{-1}(k)\beta \quad \text{and}$$

$$\begin{aligned} MSE_{RE}(k) &= \sigma^2[tr(WC^{-1}W') - tr(WAW')] + \eta' W' W\eta \\ &+ 2k\eta' W' C^{-1}(k)\beta + k^2\beta' C^{-2}\beta, \end{aligned} \quad (1.7)$$

where $\eta = C^{-1}H'(HC^{-1}H')^{-1}(H\beta - h)$ and $A = C^{-1}H'(HC^{-1}H')^{-1}HC^{-1}$.

It is well known that the $RRRE$ performs better than the $URRE$, when the restrictions hold but as long as the parameters, β moves away from the subspace $H\beta = h$, the $RRRE$ becomes biased and completely inefficient while the performance of the $URRE$ remains stable. As a result, one may combine the $URRE$ and $RRRE$ to obtain a better performance of the estimators in presence of the uncertain prior information (UPI) $H\beta = h$, which leads to preliminary test ridge regression estimator ($PTRRE$). Saleh and Kibria (1993) combine the idea of $PTLSE$ and RRE and define the $PTRRE$ as,

$$\hat{\beta}^{PT}(k) = W\hat{\beta}^{PT}, \quad (1.8)$$

where $\hat{\beta}^{PT} = \hat{\beta}^{RE} I(\mathcal{L} \le \mathcal{L}_{n,\alpha}) + \hat{\beta}^{UR} I(\mathcal{L} > \mathcal{L}_{n,\alpha})$ is the usual preliminary test least squares estimator ($PTLSE$). Here, $\mathcal{L}$ is the general test-statistic for testing the null-hypothesis in

(1.2), and $\mathcal{L}_{n,\alpha}$ is the upper α-level critical value of $\mathcal{L}_n$ and $I(A)$ is the indicator function of the set A. The preliminary test approach estimation has been pioneered by Bancroft (1944), followed by Bancroft (1964), Mosteller (1948), Han and Bancroft (1968), and Giles (1991) among others. The ridge regression approach has been studied by Hoerl and Kennard (1970), McDonald and Galarneau (1975), Golub et al. (1979), Lawless (1978), Gibbons (1981), Sarkar (1992), Saleh and Kibria (1993) and Kibria (1996) to mention a few.

The main objective of this paper is to provide a finite sample theory of the $PTRRE$ based on W, LR and LM tests. We assume a Gaussian linear regression model to estimate the parameters in the model. We organize this paper as follows. In Section 2 we propose the preliminary test ridge regression estimators ($PTRRE$) based on W, LR and LM tests. Section 3 contains the bias and the MSE expressions of the estimators. In Section 4 we discuss the relative performance of the estimators. The computed risk analysis and graphs are presented in Section 5. The maximum and minimum guaranteed efficiency is discussed in Section 6. Finally, summary and concluding remarks have been added in Section 7.

2 Proposed Estimators based on W, LR and LM Tests

The usual test statistic for testing the null hypothesis in (1.2) is

$$F = \frac{(RRSS - URRE)/q}{URSS/(n-p)} = \frac{(H\hat{\beta}^{UR} - h)'(HC^{-1}H')^{-1}(H\hat{\beta}^{UR} - h)}{q\hat{\sigma}^2_{UR}},$$

where $URSS = (Y - X\hat{\beta}^{UR})'(Y - X\hat{\beta}^{UR})$ is the unrestricted residual sum of squares and $RRSS = (Y - X\hat{\beta}^{RE})'(Y - X\hat{\beta}^{RE})$ is the restricted residual sum of squares. The test-statistic F follows a central F-distribution with $(q, n-p)$ degrees of freedom (DF) under H_0. However, when H_0 does not hold the test statistic F follows a non-central F distribution with non-central parameter $\frac{1}{2}\Delta$, where

$$\Delta = \frac{(H\beta - h)'(HC^{-1}H')^{-1}(H\beta - h)}{\sigma^2} = \frac{\eta' C\eta}{\sigma^2} \tag{2.9}$$

is called the departure parameter.

The following three tests, W, LR and LM are well employed for testing the hypothesis (1.2) in Econometric Theory. Wald (1943), first introduce the W test as follows:

$$\mathcal{L}_W = \frac{(H\hat{\beta}^{UR} - h)'(HC^{-1}H')^{-1}(H\hat{\beta}^{UR})}{\hat{\sigma}^2_{UR}} = \frac{nq}{n-p}F. \tag{2.10}$$

The well known LR test is

$$\mathcal{L}_{LR} = n[\ln\hat{\sigma}^2_{RE} - \hat{\sigma}^2_{UR}] = n\ln\left(1 + \frac{\mathcal{L}_W}{n}\right). \tag{2.11}$$

Aitchison and Silvey (1958) and Silvey (1959) introduce the LM test as

$$\mathcal{L}_{LM} = \frac{(H\hat{\beta}^{UR} - h)'(HC^{-1}H')^{-1}(H\hat{\beta}^{UR} - h)}{\hat{\sigma}^2_{RE}} = \frac{\mathcal{L}_W}{1 + \mathcal{L}_W/n}. \tag{2.12}$$

It is observed that $\mathcal{L}_W$ and $\mathcal{L}_{LM}$ test statistics differ only by different estimates of σ^2. Also note that the LM test is the same as the score test of Rao (1947). Savin (1976), and Berndt and Savin (1977) have shown that the following inequality

$$\mathcal{L}_W \geq \mathcal{L}_{LR} \geq \mathcal{L}_{LM} \tag{2.13}$$

exists among these three tests. From equations (2.10) to (2.12), we also noticed that $\mathcal{L}_{LR}$ and $\mathcal{L}_{LM}$ statistics are function of $\mathcal{L}_W$ and therefore, all the test statistics are monotonic function of F statistic. Each of the test statistic has a different sampling distribution and hence the critical values. The $PTRRE$ defined in term of exact tests at a given significance level has the same bias and MSE. However, due to the inequality relation among the value of test statistics, the $PTRRE$s based on a fixed critical value may have different biases and mean squares errors.

The exact sampling distribution of the test statistics is complicated. Therefore, the critical regions of the tests are commonly based on asymptotic approximations. It can be shown that under the restrictions (1.2), all tests are asymptotically distributed as χ^2-random variable with q degrees of freedom. This asymptotic chi-square distribution has wide application in the field of Econometrics. The test based on the approximate critical values are known as large sample tests. We propose the following $PTRRE$ based on W, LR and LM tests, which are given respectively,

$$\begin{aligned}
\hat{\beta}^{PT}_{W}(k) &= \hat{\beta}^{RE}(k)I(\mathcal{L}_W \leq \chi^2_\alpha(q)) + \hat{\beta}^{UR}(k)I(\mathcal{L}_W > \chi^2_\alpha(q)), \\
\hat{\beta}^{PT}_{LR}(k) &= \hat{\beta}^{RE}(k)I(\mathcal{L}_{LR} \leq \chi^2_\alpha(q)) + \hat{\beta}^{UR}(k)I(\mathcal{L}_{LR} > \chi^2_\alpha(q)), \quad \text{and} \\
\hat{\beta}^{PT}_{LM}(k) &= \hat{\beta}^{RE}(k)I(\mathcal{L}_{LM} \leq \chi^2_\alpha(q)) + \hat{\beta}^{UR}(k)I(\mathcal{L}_{LM} > \chi^2_\alpha(q)), \qquad (2.14)
\end{aligned}$$

where $\chi^2_\alpha(q)$ is the upper percentile points of the central χ^2 distribution with q degrees of freedom. For excellent references and for various researches on W, LR and LM tests, readers are refereed to Savin (1976), Berndt and Savin (1977), Rao and Mukerjee (1977), Evans and Savin (1982), Ullah and Zinde-Walsh (1984) and most recently Billah and Saleh (2000) among others. In the following Section, we will provide the Bias and MSE expressions of the proposed estimators.

3 Biases and MSE Expressions

The biases and the MSE expressions of the proposed estimators are routinely followed from Judge and Bock (1978, Chapter 10), and Saleh and Kibria (1993).

The Biases of the proposed estimators are as follows:

$$\begin{aligned}
B_W(k,\alpha,\Delta) &= -W\eta G_{q+2,n-p}(l_1^W;\Delta) - kC^{-1}(k)\beta, \\
B_{LR}(k,\alpha,\Delta) &= -W\eta G_{q+2,n-p}(l_1^{LR};\Delta) - kC^{-1}(k)\beta, \\
B_{LM}(k,\alpha,\Delta) &= -W\eta G_{q+2,n-p}(l_1^{LM};\Delta) - kC^{-1}(k)\beta,
\end{aligned} \tag{3.15}$$

where $l_1^W = \frac{n-p}{n(q+2)}\chi^2_\alpha(q)$, $l_1^{LR} = \left(\frac{n-p}{q+2}\right)\left(e^{\frac{\chi^2_\alpha(q)}{n}} - 1\right)$, $l_1^{LM} = \frac{(n-p)\chi^2_\alpha(q)}{(q+2)(n-\chi^2_\alpha(q))}$, and $G_{q+2,n-p}(.;\Delta)$ is the cumulative distribution function (CDF) of a non-central F-distribution with $(q+2, n-p)$ degrees of freedom (DF) and non-centrality parameter $\frac{1}{2}\Delta$. Note that for $\alpha = 1$, we reject the null hypothesis, then the bias of the three estimators coincide with the bias of the $URRE$, $\hat{\beta}^{UR}(k)$, however, for $\alpha = 0$, we do not reject the null hypothesis and the bias of the proposed estimators coincide with that of the $RRRE$, $\hat{\beta}^{RE}(k)$. As $\Delta \to \infty$, $B_W(k,\alpha,\infty) = B_{LR}(k,\alpha,\infty) = B_{LM}(k,\alpha,\infty) = B_{UR}(k) = -kC^{-1}(k)\beta$, whereas, the bias of the $RRRE$ remains unbounded. Since, $l_1^{LM} \geq l_1^{LR} \geq l_1^W$, for all α, p and n, it follows that

$$G_{q+2,n-p}(l_1^{LM};\Delta) \geq G_{q+2,n-p}(l_1^{LR};\Delta) \geq G_{q+2,n-p}(l_1^W;\Delta). \tag{3.16}$$

Now, based on equations (3.15) and (3.16), we may state the following theorem.

Theorem 1: Under the null hypothesis the proposed estimators are biased and the amount of biases are same. However, under the alternative hypothesis the dominance picture of the proposed estimators is

$$\hat{\beta}^{PT}_{LM}(k) \geq \hat{\beta}^{PT}_{LR}(k) \geq \hat{\beta}^{PT}_{W}(k),$$

where $\geq$ denotes the dominance in the sense of having smaller bias. For $k = 0$, we have the dominance picture for the corresponding preliminary test least squares estimators ($PTLSE$) based on LM, LR and W tests respectively.

The MSE expressions for $\hat{\beta}^{PT}_{W}(k)$, $\hat{\beta}^{PT}_{LR}(k)$ and $\hat{\beta}^{PT}_{LM}(k)$ are respectively provided below:

$$\begin{aligned}
MSE_W(k) &= \sigma^2 tr(WC^{-1}W') - \sigma^2 tr(WAW')G_{q+2,n-p}(l_1^W;\Delta) + \eta'W'W\eta \\
&\times \left[2G_{q+2,n-p}(l_1^W;\Delta) - G_{q+4,n-p}(l_2^W;\Delta)\right] + 2kG_{q+2,n-p}(l_1^W;\Delta)\eta'W'C^{-1}(k)\beta \\
&+ k^2\beta'C^{-2}(k)\beta,
\end{aligned}$$

$$
\begin{aligned}
MSE_{LR}(k) &= \sigma^2 tr(WC^{-1}W') - \sigma^2 tr(WAW')G_{q+2,n-p}(l_1^{LR};\Delta) + \eta'W'W\eta \\
&\times \left[2G_{q+2,n-p}(l_1^{LR};\Delta) - G_{q+4,n-p}(l_2^{LR};\Delta)\right] + 2kG_{q+2,n-p}(l_1^{LR};\Delta)\eta'W'C^{-1}(k)\beta \\
&+ k^2\beta'C^{-2}(k)\beta,
\end{aligned}
$$

$$
\begin{aligned}
MSE_{LM}(k) &= \sigma^2 tr(WC^{-1}W') - \sigma^2 tr(WAW')G_{q+2,n-p}(l_1^{LM};\Delta) + \eta'W'W\eta \\
&\times \left[2G_{q+2,n-p}(l_1^{LM};\Delta) - G_{q+4,n-p}(l_2^{LM};\Delta)\right] + 2kG_{q+2,n-p}(l_1^{LM};\Delta)\eta'W'C^{-1}(k)\beta \\
&+ k^2\beta'C^{-2}(k)\beta, \qquad (3.17)
\end{aligned}
$$

where $l_2^W = \frac{n-p}{n(q+4)}\chi^2_\alpha(q)$, $l_2^{LR} = \left(\frac{n-p}{q+4}\right)\left(e^{\frac{\chi^2_\alpha(q)}{n}} - 1\right)$, $l_2^{LM} = \frac{(n-p)\chi^2_\alpha(q)}{(q+4)(n-\chi^2_\alpha(q))}$, and $G_{q+4,n-p}(.;\Delta)$ is the cumulative distribution function (CDF) of a non-central F-distribution with $(q+4, n-p)$ degrees of freedom (DF) and non-centrality parameter $\frac{1}{2}\Delta$.

4 Performance of the Estimators Under MSE Criterion

In this Section we will compare the performance of the proposed estimators by using MSE criterion. We note from (3.17) that for given α and known data, the MSEs depend on the departure parameter Δ and ridge parameter k. Therefore, we will study the relative performance of the estimators based on values of Δ and k and provided them in the following two subsections.

4.1 Performance based on Δ

We obtain from Anderson (1984, Theorem A.2.4, p.590) that

$$
\gamma_p \le \frac{\eta'W'W\eta}{\eta'C\eta} \le \gamma_1, \quad \text{or}
$$

$$
\sigma^2\Delta\gamma_p \le \eta'W'W\eta \le \sigma^2\Delta\gamma_1, \qquad (4.18)
$$

where γ_1 and γ_p are the largest and the smallest characteristic roots of the matrix $(W'WC^{-1})$ and $\Delta = \frac{\eta'C\eta}{\sigma^2}$.

Now we compare between $\hat\beta_W^{PT}(k)$ and $\hat\beta_{LR}^{PT}(k)$. The MSE difference is:

$$
\begin{aligned}
MSE_W(k) &- MSE_{LR}(k) = \\
&\sigma^2 tr(WAW')\psi - \eta'W'W\eta[2\psi - \psi^*] - 2k\sigma^{-2}\eta'W'C^{-1}(k)\beta\psi, \quad (4.19)
\end{aligned}
$$

where $\psi = G_{q+2,n-p}(l_1^{LR};\Delta) - G_{q+2,n-p}(l_1^W;\Delta)$ and $\psi^* = G_{q+4,n-p}(l_2^{LR};\Delta) - G_{q+4,n-p}(l_2^W;\Delta)$. Note from (3.16) that both ψ and ψ^* are positive for all k, Δ and α.

The difference in (4.19) is non-negative (≥ 0), whenever

$$\Delta \leq \frac{tr(WAW') - 2k\sigma^{-2}\eta' W' C^{-1}(k)\beta}{\gamma_1 \left(2 - \frac{\psi^*}{\psi}\right)} = \Delta_1(k, \alpha). \tag{4.20}$$

Thus, $\hat{\beta}_{LR}^{PT}(k)$ performs better than $\hat{\beta}_{W}^{PT}(k)$, when (4.20) holds. However, $\hat{\beta}_{W}^{PT}(k)$ performs better than $\hat{\beta}_{LR}^{PT}(k)$, whenever

$$\Delta > \frac{tr(WAW') - 2k\sigma^{-2}\eta' W' C^{-1}(k)\beta}{\gamma_p \left(2 - \frac{\psi^*}{\psi}\right)} = \Delta_2(k, \alpha). \tag{4.21}$$

Under the null hypothesis, the difference in (4.19) is always positive for all α, therefore, $\hat{\beta}_{LR}^{PT}(k)$ is superior to $\hat{\beta}_{W}^{PT}(k)$. Now we can describe the graph of $\hat{\beta}_{LR}^{PT}(k)$ as follows. At $\Delta = 0$, it assumes the value,

$$\sigma^2 tr(WC^{-1}W') - \sigma^2 tr(WAW') G_{q+2,n-p}(l_1^{LR}; 0) + k^2 \beta' C^{-2}(k)\beta,$$

then increases from 0, crossing the risk of $\hat{\beta}_{W}^{PT}(k)$ to a maximum and then drops gradually towards $MSE_{UR}(k)$ as $\Delta \to \infty$.

Now we compare the performance of $\hat{\beta}_{LR}^{PT}(k)$ with that of $\hat{\beta}_{LM}^{PT}(k)$. The MSE difference is:

$$\begin{aligned} MSE_{LR}(k) \quad - \quad & MSE_{LM}(k) = \\ & \sigma^2 tr(WAW')\psi_1 - \eta' W' W \eta [2\psi_1 - \psi_1^*] - 2k\sigma^{-2}\eta' W' C^{-1}(k)\beta\psi_1, \end{aligned} \tag{4.22}$$

where $\psi_1 = G_{q+2,n-p}(l_1^{LM}; \Delta) - G_{q+2,n-p}(l_1^{LR}; \Delta)$ and $\psi_1^* = G_{q+4,n-p}(l_2^{LM}; \Delta) - G_{q+4,n-p}(l_2^{LR}; \Delta)$. The difference in (4.22) is non-negative (≥ 0), whenever

$$\Delta \leq \frac{tr(WAW') - 2k\sigma^{-2}\eta' W' C^{-1}(k)\beta}{\gamma_1 \left(2 - \frac{\psi_1^*}{\psi_1}\right)} = \Delta_3(k, \alpha). \tag{4.23}$$

Thus, $\hat{\beta}_{LM}^{PT}(k)$ performs better than $\hat{\beta}_{LR}^{PT}(k)$ when (4.23) holds, otherwise $\hat{\beta}_{LR}^{PT}(k)$ performs better than $\hat{\beta}_{LM}^{PT}(k)$, whenever

$$\Delta > \frac{tr(WAW') - 2k\sigma^{-2}\eta' W' C^{-1}(k)\beta}{\gamma_p \left(2 - \frac{\psi_1^*}{\psi_1}\right)} = \Delta_4(k, \alpha). \tag{4.24}$$

Under the null hypothesis the difference in (4.22) is always positive for all α, therefore, $\hat{\beta}_{LM}^{PT}(k)$ is superior to $\beta_{LR}^{PT}(k)$. Now we can describe the graph of $\hat{\beta}_{LM}^{PT}(k)$ as follows. At $\Delta = 0$, it assumes a value

$$\sigma^2 tr(WC^{-1}W') - \sigma^2 tr(WAW') G_{q+2,n-p}(l_1^{LM}; 0) + k^2 \beta' C^{-2}(k)\beta,$$

then increases from 0, crossing the risk of $\hat{\beta}^{PT}_{LM}(k)$ and $\hat{\beta}^{PT}_{W}(k)$ to a maximum and then drops gradually towards $MSE_{UR}(k)$ as $\Delta \to \infty$.

Based on the above analysis we may state the following theorem:

Theorem 2: Under the null hypothesis the dominance picture of the proposed estimators is:

$$\hat{\beta}^{PT}_{LM}(k) \geq \hat{\beta}^{PT}_{LR}(k) \geq \hat{\beta}^{PT}_{W}(k),$$

where $\geq$ denotes the dominance in the sense of having smaller MSE.

Under the alternative hypothesis, the dominance picture of the proposed estimators is:

$$\hat{\beta}^{PT}_{LM}(k) \geq \hat{\beta}^{PT}_{LR}(k) \geq \hat{\beta}^{PT}_{W}(k),$$

in the interval

$$\Delta \in \left(0,\ \Delta^*_{13}(k,\alpha)\right],$$

where $\Delta^*_{13}(k,\alpha) = min\left\{\Delta_1(k,\alpha), \Delta_3(k,\alpha)\right\}$, also $\Delta_1(k,\alpha)$ and $\Delta_3(k,\alpha)$ are given in (4.20) and (4.23) respectively, while

$$\hat{\beta}^{PT}_{W}(k) \geq \hat{\beta}^{PT}_{LR}(k) \geq \hat{\beta}^{PT}_{LM}(k),$$

in the interval

$$\Delta \in \left(\Delta^*_{24}(k,\alpha), \infty\right),$$

where $\Delta^*_{24}(k,\alpha) = max\left\{\Delta_2(k,\alpha), \Delta_4(k,\alpha)\right\}$, also $\Delta_2(k,\alpha)$ and $\Delta_4(k,\alpha)$ are given in (4.21) and (4.24) respectively.

4.2 Performance based on k

In this subsection, we will compare the performance of the proposed estimators based on ridge parameter k. For this, we assume that Q be the orthogonal matrix with eigenvectors of C so that

$$Q'CQ = \Lambda = diag(\lambda_1, \lambda_2, ..., \lambda_p).$$

As C is symmetric we can write

$$WAW' = Q[\Lambda + kI_p]^{-1}\Lambda A^*\Lambda[\Lambda + kI_p]^{-1}Q', \tag{4.25}$$

where $Q'AQ = A^*$. Now without loss of generality we assume that $\lambda_1 \geq \lambda_2 \geq .....\lambda_p > 0$, and we can write,

$$tr(WC^{-1}W') = \sum_{i=1}^{p} \frac{\lambda_i}{(\lambda_i + k)^2} \quad \text{and} \quad tr(WAW') = \sum_{i=1}^{p} \frac{\lambda_i^2 a^*_{ii}}{(\lambda_i + k)^2}, \tag{4.26}$$

where $a^*_{ii} \geq 0$ is the i^{th} diagonal element of the matrix A^*. Also,

$$\beta' C^{-2}(k)\beta = \sum_{i=1}^{p} \frac{\alpha_i^2}{(\lambda_i + k)^2}, \quad \text{where} \quad \alpha = Q'\beta. \tag{4.27}$$

$$\eta' W' W \eta = \sum_{i=1}^{p} \frac{\lambda_i^2 \eta_i^{*2}}{(\lambda_i + k)^2} \quad \text{and} \quad \eta' W' C^{-1}(k)\beta = \sum_{i=1}^{p} \frac{\alpha_i \lambda_i \eta_i{}^*}{(\lambda_i + k)^2}, \tag{4.28}$$

where $\eta^* = \eta' Q$.

Thus, using equations (4.26) to (4.28), the MSE difference in equation (4.19) can be expressed in terms of the eigen values as

$$MSE_W(k) - MSE_{LR}(k) = \sum_{i=1}^{p} \frac{\lambda_i}{(\lambda_i + k)^2}\left[\sigma^2 \psi a^*_{ii}\lambda_i - (2\psi - \psi^*)\lambda_i \eta_i^{*2} - 2\psi k \eta_i^* \alpha_i\right]. \tag{4.29}$$

The difference in (4.29) will be non-negative (≥ 0) if

$$k \leq \frac{min_i\left[\sigma^2 \psi a^*_{ii}\lambda_i - (2\psi - \psi^*)\lambda_i \eta_i^{*2}\right]}{max_i[2\psi \eta_i^* \alpha_i]} = k_1(\alpha, \Delta). \tag{4.30}$$

Thus, $\hat{\beta}^{PT}_{LR}(k)$ will dominate $\hat{\beta}^{PT}_{W}(k)$ if $0 \leq k \leq k_1(\alpha, \Delta)$, while $\hat{\beta}^{PT}_{W}(k)$ will dominate $\hat{\beta}^{PT}_{LR}(k)$ whenever

$$k > \frac{max_i\left[\sigma^2 \psi a^*_{ii}\lambda_i - (2\psi - \psi^*)\lambda_i \eta_i^{*2}\right]}{min_i[2\psi \eta_i^* \alpha_i]} = k_2(\alpha, \Delta). \tag{4.31}$$

Now we compare between $\hat{\beta}^{PT}_{LR}(k)$ and $\hat{\beta}^{PT}_{LM}(k)$ estimators. Using equations (4.26) to (4.28), the MSE difference in equation (4.22) can be expressed in terms of the eigen values as

$$MSE_{LR}(k) - MSE_{LM}(k) = \sum_{i=1}^{p} \frac{\lambda_i}{(\lambda_i + k)^2}\left[\sigma^2 \psi_1 a^*_{ii}\lambda_i - (2\psi_1 - \psi_1^*)\lambda_i \eta_i^{*2} - 2\psi_1 k \eta_i^* \alpha_i\right]. \tag{4.32}$$

The difference in (4.32) will be non-negative (≥ 0) if

$$k \leq \frac{min_i\left[\sigma^2 \psi_1 a^*_{ii}\lambda_i - (2\psi_1 - \psi_1^*)\lambda_i \eta_i^{2*}\right]}{max_i[2\psi_1 \eta_i^* \alpha_i]} = k_3(\alpha, \Delta). \tag{4.33}$$

Thus, $\hat{\beta}^{PT}_{LM}(k)$ will dominate $\hat{\beta}^{PT}_{LR}(k)$ if $0 \leq k \leq k_3(\alpha, \Delta)$, while $\hat{\beta}^{PT}_{LR}(k)$ will dominate $\hat{\beta}^{PT}_{LM}(k)$ when

$$k > \frac{max_i\left[\sigma^2 \psi_1 a^*_{ii}\lambda_i - (2\psi_1 - \psi_1^*)\lambda_i \eta_i^{*2}\right]}{min_i[2\psi_1 \eta_i^* \alpha_i]} = k_4(\alpha, \Delta). \tag{4.34}$$

Based on the above results, we may state the following theorem.

Theorem 3: Under the alternative hypothesis, the dominance picture of the proposed estimators is:

$$\hat{\beta}^{PT}_{LM}(k) \geq \hat{\beta}^{PT}_{LR}(k) \geq \hat{\beta}^{PT}_{W}(k),$$

in the interval

$$k \in (0,\ k_{13}(\alpha, \Delta)]\,,$$

where $k_{13}(\alpha, \Delta) = min\,\{k_1(\alpha, \Delta), k_3(\alpha, \Delta)\}$*, also* $k_1(\alpha, \Delta)$ *and* $k_3(\alpha, \Delta)$ *are given in (4.30) and (4.33) respectively, while*

$$\hat{\beta}^{PT}_{W}(k) \geq \hat{\beta}^{PT}_{LR}(k) \geq \hat{\beta}^{PT}_{LM}(k),$$

in the interval

$$k \in (k_{24}(\alpha, \Delta), \infty)\,,$$

where $k_{24}(\alpha, \Delta) = max\,\{k_2(\alpha, \Delta), k_4(\alpha, \Delta)\}$*, also* $k_2(\alpha, \Delta)$ *and* $k_4(\alpha, \Delta)$ *are given in (4.31) and (4.34) respectively.*

Now, considering the conditions on Δ and k simultaneously, we may state the following theorem:

Theorem 4: Under the alternative hypothesis, the dominance picture of the proposed estimators is:

$$\hat{\beta}^{PT}_{LM}(k) \geq \hat{\beta}^{PT}_{LR}(k) \geq \hat{\beta}^{PT}_{W}(k),$$

in the interval,

$$(\Delta, k) \in (0, \Delta_{13}(k, \alpha)] \times (0, k_{13}(\alpha, \Delta)]\,,$$

while

$$\beta^{PT}_{W}(k) \geq \beta^{PT}_{LR}(k) \geq \beta^{PT}_{LM}(k),$$

in the interval,

$$(\Delta, k) \in (\Delta_{13}(k, \alpha), \infty) \times (k_{13}(\alpha, \Delta), \infty).$$

5 Computed Risk Analysis

In this section we will provide some graphical representations of the proposed estimators via calculating the MSE. Note that, for given α, the MSE of the estimators depend on observed data and unknown parameters k and Δ. Thus the dominance pictures of the estimators are data dependent. In order to avoid data dependent condition, we will consider the orthonormal regression, $X'X = I$. Furthermore, to facilitate numerical computation of MSE functions of the proposed estimators, we consider $H'H = I$, $\beta'\beta$=1, and $h = 0$. Using these restrictions in (3.17), the MSE of the proposed estimators become:

$$
\begin{aligned}
MSE_W(k) &= \frac{\sigma^2}{(1+k)^2}\Big[p - qG_{q+2,n-p}(l_1^W;\Delta) + \Delta[2G_{q+2,n-p}(l_1^W;\Delta) - G_{q+4,n-p}(l_2^W;\Delta)] \\
&+ 2k\Delta G_{q+2,n-p}(l_1^W;\Delta) + k^2\Big]. \\
MSE_{LR}(k) &= \frac{\sigma^2}{(1+k)^2}\Big[p - qG_{q+2,n-p}(l_1^{LR};\Delta) + \Delta[2G_{q+2,n-p}(l_1^{LR};\Delta) - G_{q+4,n-p}(l_2^{LR};\Delta)] \\
&+ 2k\Delta G_{q+2,n-p}(l_1^{LR};\Delta) + k^2\Big]. \\
MSE_{LM}(k) &= \frac{\sigma^2}{(1+k)^2}\Big[p - qG_{q+2,n-p}(l_1^{LM};\Delta) + \Delta[2G_{q+2,n-p}(l_1^{LM};\Delta) - G_{q+4,n-p}(l_2^{LM};\Delta)] \\
&+ 2k\Delta G_{q+2,n-p}(l_1^{LM};\Delta) + k^2\Big]. \qquad (5.35)
\end{aligned}
$$

When we compare the performance of the proposed estimators based on departure parameter Δ, we see that the $URRE$ has constant risk as it does not depend on the restriction. Thus, for given k, $\hat{\beta}_{LR}^{PT}(k)$ is superior to $\hat{\beta}_W^{PT}(k)$ if $\Delta \in \left(0, \frac{q}{2-\frac{\psi^*}{\psi}+2k}\right]$, otherwise $\hat{\beta}_W^{PT}(k)$ is superior to $\hat{\beta}_{LR}^{PT}(k)$ if $\Delta \in \left(\frac{q}{2-\frac{\psi^*}{\psi}+2k},\ \infty\right)$. Now, $\hat{\beta}_{LM}^{PT}(k)$ is superior to $\hat{\beta}_{LR}^{PT}(k)$ if $\Delta \in \left(0,\ \frac{p}{2-\frac{\psi_1^*}{\psi_1}+2k}\right)$, otherwise $\hat{\beta}_{LR}^{PT}(k)$ is superior to $\hat{\beta}_{LM}^{PT}(k)$ if $\Delta \in \left[\frac{q}{2-\frac{\psi_1^*}{\psi_1}+2k},\ \infty\right)$. It is clear that when Δ moves away from H_0 beyond the value $\frac{q}{2-\frac{\psi^*}{\psi}+2k}$, the risk of $\hat{\beta}_{LR}^{PT}(k)$ becomes unbounded for any $k \geq 0$. However, when Δ moves away from H_0 beyond the value $\frac{q}{2-\frac{\psi_1^*}{\psi_1}+2k}$, the risk of $\hat{\beta}_{LM}^{PT}(k)$ becomes unbounded for any $k \geq 0$. Similarly, for given Δ, $\hat{\beta}_{LR}^{PT}(k)$ is superior to $\hat{\beta}_W^{PT}(k)$ if $k \in \left(0, \frac{1}{2}\left(\frac{q}{\Delta} - (2 - \frac{\Psi^*}{\Psi}\right)\right]$ otherwise $\hat{\beta}_W^{PT}(k)$ is superior to $\hat{\beta}_{LR}^{PT}(k)$. Similarly, $\hat{\beta}_{LM}^{PT}(k)$ is superior to $\hat{\beta}_{LR}^{PT}(k)$ if $k \in \left(0, \frac{1}{2}\left(\frac{q}{\Delta} - (2 - \frac{\Psi_1^*}{\Psi_1}\right)\right]$ otherwise $\hat{\beta}_{LR}^{PT}(k)$ is superior to $\hat{\beta}_{LM}^{PT}(k)$.

Thus it is evident that the performance of the $PTRRE$ strongly depends on the restriction of the parameters in the model and ridge parameter k. We have plotted the MSE functions versus Δ for fixed $p = 4$ and $q = 3$ and for different values of n, α and k and presented them in figures 1-4. We have also plotted the MSE functions versus k for fixed $p = 4$ and $q = 3$ and for different values of n, α and Δ and have presented them in figures 5-8. The computation of the figures have been done by Splus software. From these figures we observed that the graphical analysis support the findings of the paper.

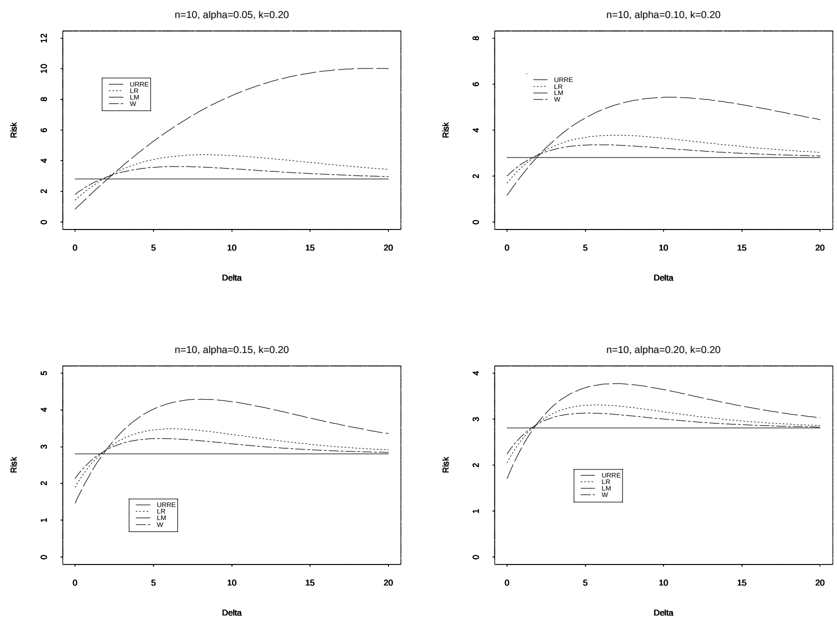

Figure 1. Risk function of the PTRRE based on the W, LR and LM tests for different significance levels and fixed k.

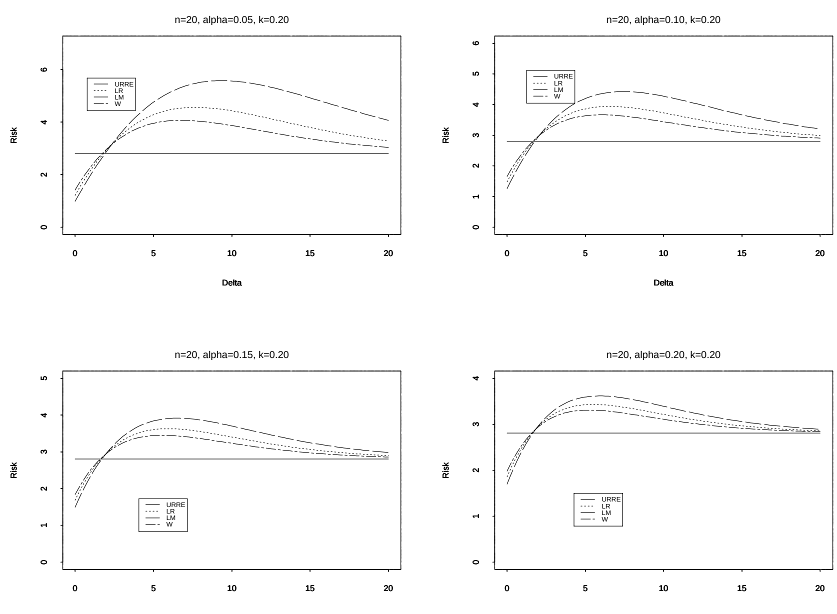

Figure 2. Risk function of the PTRRE based on the W, LR and LM tests for different significance levels and fixed k.

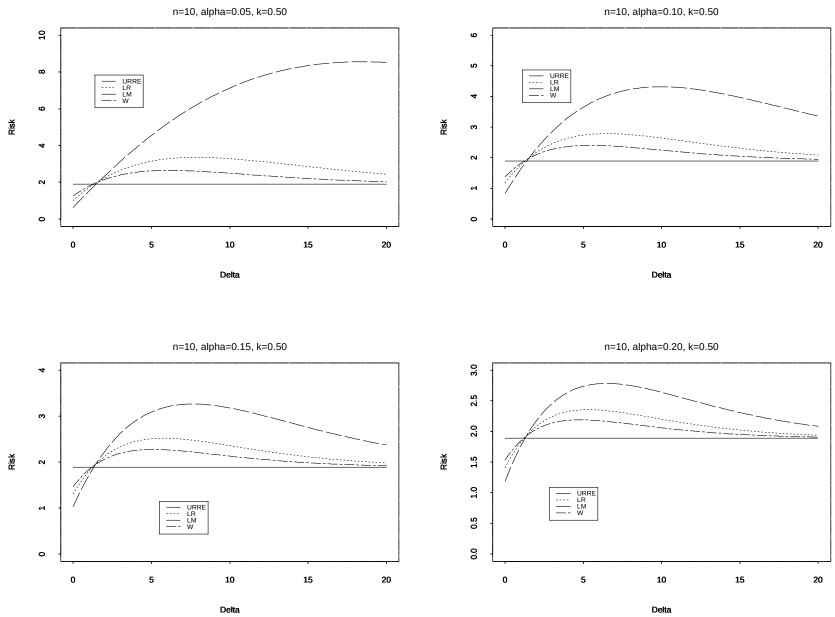

Figure 3. Risk function of the $PTRRE$ based on the W, LR and LM tests for different significance levels and fixed k.

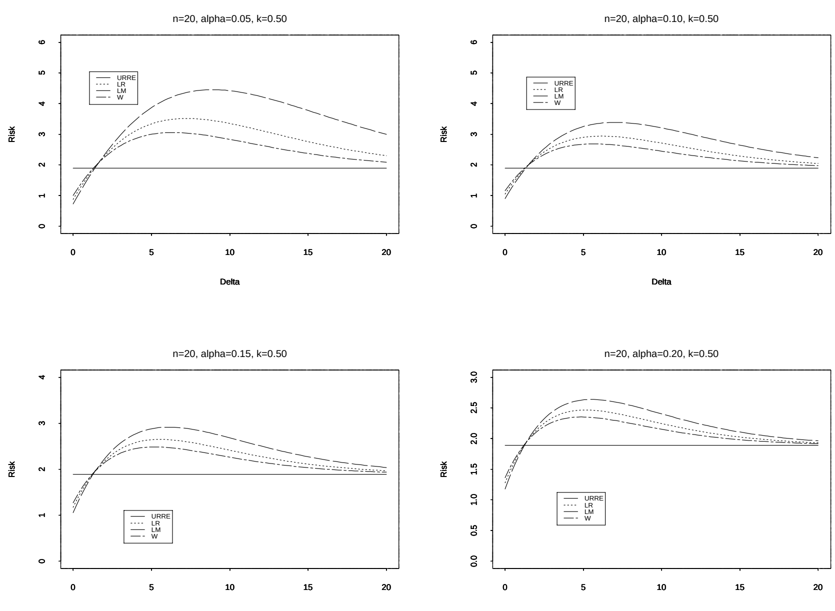

Figure 4. Risk function of the $PTRRE$ based on the W, LR and LM tests for different significance levels and fixed k.

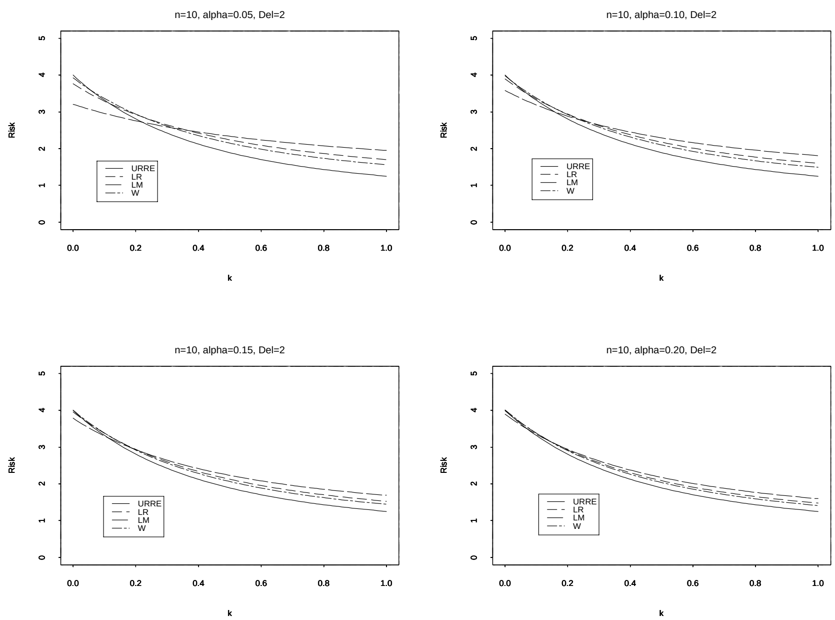

Figure 5. Risk function of the PTRRE based on the W, LR and LM tests for different significance levels and fixed delta.

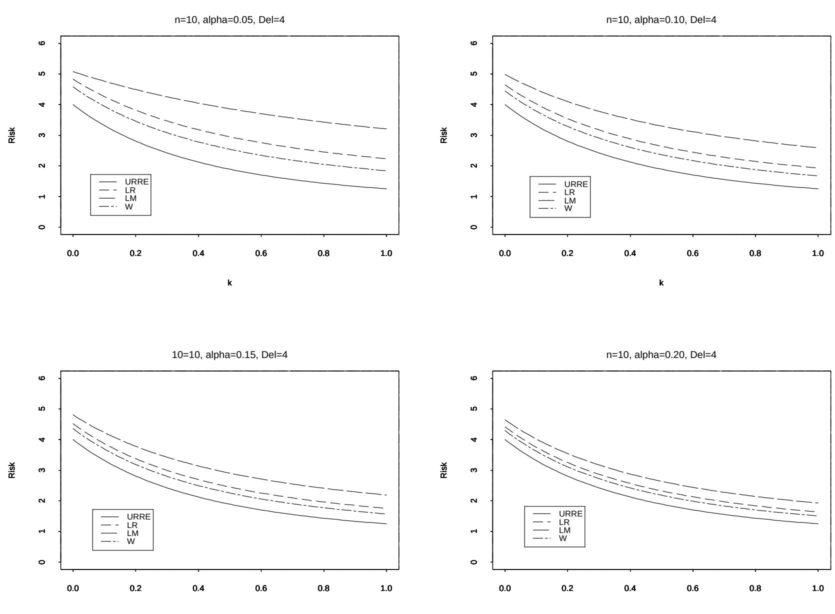

Figure 6. Risk function of the PTRRE based on the W, LR and LM tests for different significance levels and fixed delta.

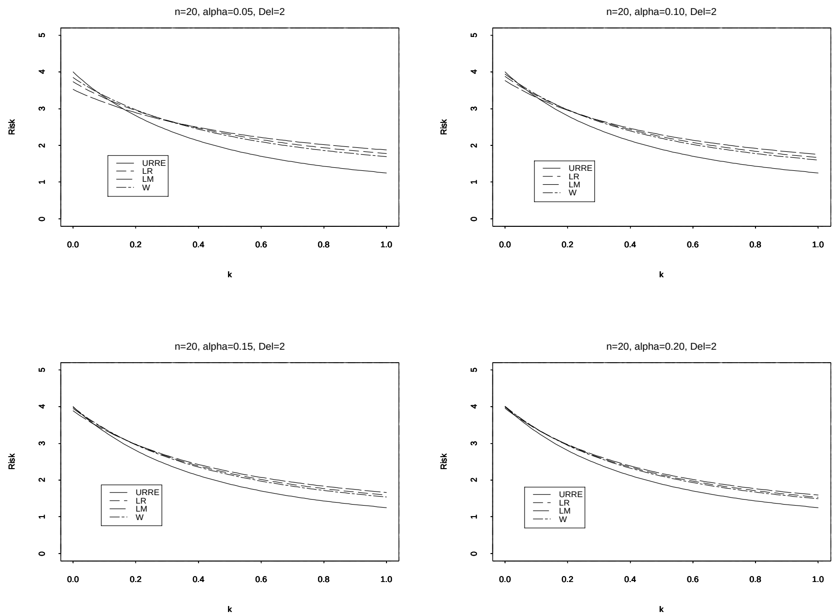

Figure 7. Risk function of the $PTRRE$ based on the W, LR and LM tests for different significance levels and fixed delta.

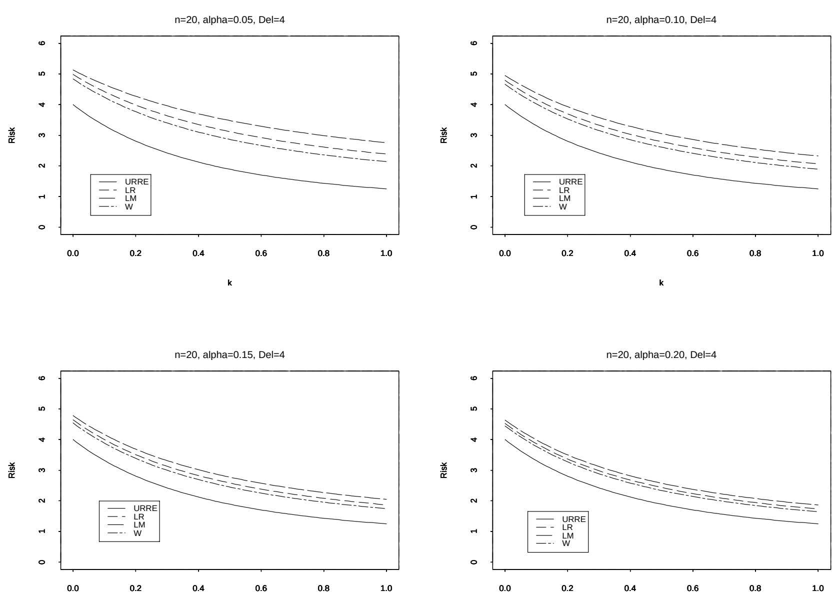

Figure 8. Risk function of the $PTRRE$ based on the W, LR and LM tests for different significance levels and fixed delta.

6 Relative Efficiency and Optimum Significance Level

In this section, we describe the relative efficiency of the proposed estimators for β. Accordingly, we provide maximum and minimum (Max & Min) rule for the optimum choice of the level of significance of the $PTRRE$ for testing the null hypothesis (1.2). For a fixed value of $k(>0)$, the relative efficiency of the $PTRRE$ ($\hat{\beta}_*^{PT}(k)$) compared to the $URRE$ ($\hat{\beta}^{UR}(k)$) is a function of α, and Δ. Let us denote this by

$$E(k,\alpha,\Delta) = \frac{MSE_{UR}(k)}{MSE_*(k)} = \Big[1 - h(k,\alpha,\Delta)\Big]^{-1}, \tag{6.36}$$

where

$$h(k,\alpha,\Delta) = \frac{g(k,\alpha,\Delta)}{\sigma^2 tr(WC^{-1}W') + k^2\beta' C^{-2}(k)\beta}, \tag{6.37}$$

and

$$\begin{aligned} g(k,\alpha,\Delta) &= \sigma^2 tr(WAW')G_{q+2,n-p}(l_1^*;\Delta) \\ &- \eta' W' W\eta\{2G_{q+2,n-p}(l_1^*;\Delta) - G_{q+4,n-p}(l_2^*;\Delta)\} \\ &- 2kG_{q+2,n-p}(l_1^*;\Delta)\eta' W' C^{-1}(k)\beta. \end{aligned} \tag{6.38}$$

For a given n, p, q and k, $E(k,\alpha,\Delta)$, is a function of α and Δ. For $\alpha \neq 0$, it has maximum at $\Delta = 0$ with value

$$E_{max}(k,\alpha,0) = \left[1 - \frac{\sigma^2 tr(WAW')G_{q+2,n-p}(l_1^*;0)}{\sigma^2 tr(WC^{-1}W') + k^2\beta' C^{-2}(k)\beta}\right]^{-1}.$$

As Δ increases from 0, $E(k,\alpha,\Delta)$ decreases and crossing the line $E(k,\alpha,\Delta) = 1$ to a minimum $E(k,\alpha,\Delta^0)$ at $\Delta = \Delta^0$, then increases towards 1 as $\Delta \to \infty$. For $\Delta = 0$ and varying α, we obtain,

$$max_{0\leq\alpha\leq 1} E(k,\alpha,0) = E(k,0,0) = \Big[1 - \frac{\sigma^2 tr(WAW')}{\sigma^2 tr(WC^{-1}W') + k^2\beta' C^{-2}(k)\beta}\Big]^{-1}. \tag{6.39}$$

The value $E(k,\alpha,0)$ decreases as α increases. On the other hand, for $\alpha \neq 0$, as Δ varies the graphs of $E(k,0,\Delta)$ and $E(k,1,\Delta)$ intersect in the range $0 \leq \Delta \leq \Delta_1(k,0)$, where

$$\Delta_1(k,0) = \frac{tr(WAW') - 2k\sigma^{-2}\eta' W' C^{-1}(k)\beta}{Ch_{max}(W'WC^{-1})},$$

and $Ch_{max}(W'WC^{-1})$ is the latgest characteritic root of the matrix $(W'WC^{-1})$. Also for $k = 0$, $E(0,0,\Delta)$ intersect at $\Delta = q$. For a general α, $E(0,0,\Delta)$ and $E(0,1,\Delta)$ will intersect in the interval $0 \leq \Delta \leq q$; the value of Δ decrease at the intersection decreases as α increase.

Thus in order to choose an estimator with optimum relative efficiency, we adopt the following rule for given k values. If $0 < \Delta < \Delta_1(k,0)$, $\hat{\beta}_*^{PT}(k)$ is chosen since $E(k,0,\Delta)$ is largest in this interval. However, in general Δ is unknown and may not lie in the interval and there is no way of choosing a uniformly best estimator. In such case we pre-assign a value of the efficiency $E^0(k)$ (minimum guaranteed efficiency) and consider the set $\mathcal{A} = \left\{\alpha | E(k,\alpha,\Delta) \geq E^0(k)\right\}$ and choose an estimator which maximizes $E(k,\alpha,\Delta)$ for all $\alpha \in \mathcal{A}$ and $\Delta \in [0,\infty)$. Thus we solve the following equation

$$max_{\alpha\in\mathcal{A}} min_{\Delta} E(k,\alpha,\Delta) = E^0(k). \tag{6.40}$$

The solution α^* for (6.40) gives the optimum choice of α and the value of $\Delta = \Delta_{min}(k)$ for which (6.40) is satisfied. At the same time these values $(\alpha^*, \Delta_{min}(k))$ yield the corresponding value of optimum k, which can be estimated from the following equation.

$$\hat{k}(\alpha,\Delta) = \frac{min_i \left[\sigma^2 a_{ii}^* \lambda_i G_{q+2,n-p}(l_1^*;\Delta) - \lambda_i \eta_i^{2*}\left\{2G_{q+2,n-p}(l_1^*;\Delta) - G_{q+4,n-p}(l_1^{**};\Delta)\right\}\right]}{max_i[2\eta_i^* \alpha_i G_{q+2,n-p}(l_1^*;\Delta)]}.$$

The above equation is obtained from the difference of $MSE_{UR}(k)$(URRE) and $MSE_*(k)$ $(PTRRE)$ and based on the smaller MSE criterion.

Thus for each estimator we can find the optimum significance level say $\alpha_*^W, \alpha_*^{LR}, \alpha_*^{LM}$ respectively, with minimum guaranteed efficiency $E^0(k)$. Then, we choose $\alpha_*^W = min(\alpha_*^W, \alpha_*^{LR}, \alpha_*^{LM})$ as optimum level of significance since $\alpha_*^W \leq \alpha_*^{LR} \leq \alpha_*^{LM}$. Note that our main goal is to choose the smallest level of significance (α) which gives the best estimator in the sense of having highest efficiency. Imposing the restrictions as, $X'X = I$, $H'H = I$, and $\beta'\beta = 1$, in equation (6.36), we obtain the maximum and minimum (Max & Min) guranteed efficency of the proposed estimators compared to $URRE$ (for $k = 0$, we obtain $URLSE$). Tables 1 through 3 provide the value of the maximum and minimum guaranteed relative efficiency and recommended corresponding size of α of the proposed estimators for $p = 4$, $q = 3$ and $n = 10, 15, 20, 30$, and $k = 0.10, 0.50, 0.75$. Also Tables 4 and 5 provide the value of the maximum and minimum guaranteed relative efficiency of $PTRRE$ compared to $URLSE$ for $p = 4$, $q = 3$ and $n = 10, 15, 20, 30$, and $k = 0.10, 0.75$. We found that for all Δ and k, the $PTRRE$s are more efficient compared to $URLSE$ than $URRE$. How can one use the table? For example, if $n = 20$, $p = 4$, $k = 0.10$, and the experimenter wishes to have an estimator with a minimum guaranteed efficiency of 0.75. Now using Table 1, we recommend him/her to select $\alpha = 0.05$, corresponding to $\hat{\beta}_W^{PT}(k)$, because such a choice of α would yield an estimator with a minimum efficiency of 0.75077 and maximum efficiency 1.99859. The corresponding minimum and maximum relative efficient of the $PTRRE$ compared to $URLSE$ are 0.90616 and 2.41226 respectively. It is interesting to note that $\alpha = 0.05$ is the traditional level of significance used by Sir R. A. Fisher, the founder of classical statistics.

7 Summary and Concluding Remarks

In this paper we studied the effect of three tests, W, LR and LM on the performance of the $PTRRE$ for the regression parameters when there exist a uncertain prior information in the parameter space. In literature, it is known that $W \leq LR \leq LM$. Thus there may exists conflict in the resulting test conclusions when certain critical value is chosen. We have effectively determined some conditions on Δ, the departure parameter and k, the ridge parameter for the superiority of the proposed estimators. Note that the superiority of the proposed estimators depend on data and the information about the hypothesis. We have also discussed the method of choosing optimum level of significance to obtain mimimum guaranteed efficient estimators. The $PTRRE$ based on Wald test is found to be the most efficient in the choice of the smallest level of significance. The most interesting result of the paper is the optimum choice of the level of significance becomes the traditional choice by using W test.

References

Aitchison, J. and D. Silvey (1958). Maximum likelihood estimation of parameters subject to restraints. *Annals of Mathematical Statistics*, 29, 813-828.

Anderson, T. W. (1984). An introduction to multivariate statistical analysis. Second Edition. John Wiley, NY.

Bancroft, T. A. (1944). On biases in estimation due to use of preliminary tests of significance. *Annals of Mathematics and Statistics*, 15, 190-204.

Bancroft, T. A. (1964). Analysis and inference for incompletely specified models involving the use of preliminary test(s) of significance. *Biometrics*, 20, 427-442.

Berndt, E. and Savin, N. E. (1977). Conflict among criteria for testing Hypothesis in the multivariate linear regression model, *Econometrica*, 45, 1263-1278.

Billah, B. and Saleh, A. K. Md. E. Saleh (2000). Performance of the PTE based on the conflicting W, LR and LM Tests in Regression model, *Advances on Methodologicaal and Applied Aspects of Probability and Statistics* (N.Balakrishnan Editor.Gordon and Breach Science Publishers.

Evans, G. B. A. and Savin, N. E. (1982). Conflict among the criteria revisited; The W, LR and LM Tests, *Econometrica*, 50, 737-748.

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

Giles, A. J. (1991). Pretesting for Linear Restrictions in a Regression Model with Spherically Symmetric Distributions. *J. Econometrics*. 50, 377-398.

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.

Han, C-P, and Bancroft, T. A. (1968). On Poolingmeans when variance is unknown. *Journal of the American Statistical Association*, 63, 1333-1342.

Hoerl, A.E. and Kennard, R.W. (1970). Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics*, 12, 55-67.

Judge, G.G. and Bock, M.E. (1978). The Statistical Implications of Pre-test and Stein-rule Estimators in Econometrics, North-Holland Publishing Company, Amsterdam.

Kibria, B. M. G. (1996). On Preliminary test ridge regression estimators for linear restrictions in a regression model with non-normal disturbances. *Communications in Statistics*, A 25, 2349-2369.

Lawless, J. F. (1978). Ridge and related estimation procedure. *Communications in Statistics*, A 7, 139-164.

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

Mosteller, F. (1948). On pooling data. *Journal of the American Statistical Association*, 43, 231-242.

Rao, C. R. (1947). Large sample tests of statistical hypothesis concerning several parameters. Proceedings of the Cambridge Philosophical Society, 44, 50-57.

Rao, C. R. and Mukerjee, R. (1977). Comparison of LR, Score and Wald Tests in a Non-ID setting, *Journal of Multivariate Analysis*, 60, 99-110.

Saleh, A. K. Md. and Kibria, B. M. G. (1993). Performances of some new preliminary test ridge regression estimators and their properties. *Communications in Statistics*, **22**, 2747-2764.

Sarkar, N. (1992). A new Estimator Combining the Ridge Regression and the Restricted Least Squares Method of Estimation. *Communications in Statistics*, A. 21(7), 1987-2000.

Savin, N. E. (1976). Conflict among testing procedures in a linear regression model with autoregressive disturbances. *Econometrica*, 44, 1303-1315.

Silvey, S. D. (1959). The Lagrange Multiplier test. *Annals of Mathematical Statistics*, 30, 389-407.

Ullah, A. and Zinde-Walsh, V. (1984). On the robustness of LM, LR and W Tests in Rgression Model, *Econometoca*, 52, 1055-1066.

Wald, A. (1943). Tests of hypothesis concerning several parameters when the number of observations is large. *Transation of the American Mathematical Society*, 54, 426-482.

Dr. B. M. Golam Kibria
Department of Statistics
Florida International University
Miami, FL 33199, USA
Email: kibriag@fiu.edu
and
Professor A. K. Md. E. Saleh
Department of Mathematics and Statistics
Carleton University
Ottawa Ontario K1S 5B6, Canada.

Table 1: Max & Min Guaranteed Efficiency of $PTRRE$ compared to $URRE$ ($k = 0.10$)

		n=10						
Test	α:	5%	10%	15%	20%	25%	30%	50%
W	Max	1.56396	1.40452	1.31381	1.25158	1.20502	1.16843	1.07581
	Min	0.82649	0.86469	0.88679	0.90944	0.93083	0.95489	0.97830
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LR	Max	1.98363	1.65488	1.48185	1.37054	1.29167	1.23250	1.09514
	Min	0.67904	0.77332	0.82801	0.86937	0.90494	0.94102	0.97308
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LM	Max	3.43551	2.43850	1.93330	1.65251	1.47724	1.35867	1.12509
	Min	0.38030	0.57241	0.70822	0.79240	0.85876	0.91848	0.96534
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
		n=15						
W	Max	1.83172	1.58811	1.45108	1.35828	1.28972	1.23641	1.10432
	Min	0.77656	0.82432	0.85274	0.88165	0.90924	0.94060	0.97102
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LR	Max	2.21955	1.82231	1.60781	1.46859	1.36954	1.29508	1.12168
	Min	0.66690	0.75466	0.80712	0.85000	0.88849	0.92932	0.96659
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LM	Max	3.02054	2.29117	1.89905	1.65995	1.50006	1.38623	1.14484
	Min	0.48927	0.64030	0.73644	0.80272	0.85881	0.91407	0.96089
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
		n=20						
W	Max	1.99859	1.7017	1.53497	1.42272	1.34030	1.27661	1.12065
	Min	0.75077	0.8032	0.83485	0.86694	0.89775	0.93294	0.96704
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LR	Max	2.33411	1.90680	1.67227	1.51912	1.40984	1.32755	1.13557
	Min	0.66479	0.74801	0.79844	0.84148	0.88093	0.92374	0.96335
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LM	Max	2.89281	2.24537	1.88835	1.66382	1.50993	1.39823	1.15406
	Min	0.54084	0.66865	0.74842	0.80742	0.85913	0.91229	0.95891
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
		n=30						
W	Max	2.19088	1.83293	1.63124	1.49606	1.39742	1.32167	1.13854
	Min	0.72458	0.78156	0.81647	0.85175	0.88583	0.92497	0.96283
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LR	Max	2.44533	1.99128	1.73747	1.57050	1.45095	1.36076	1.14985
	Min	0.66497	0.74294	0.79081	0.83368	0.87381	0.91835	0.96012
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LM	Max	2.79492	2.20990	1.88024	1.66770	1.51903	1.40929	1.16290
	Min	0.58870	0.69398	0.75931	0.81187	0.85960	0.91073	0.95706
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000

Table 2: Max & Min Guaranteed Efficiency of $PTRRE$ compared to $URRE$ ($k = 0.50$)

		n=10						
Test	α:	5%	10%	15%	20%	25%	30%	50%
W	Max	1.51569	1.37315	1.29093	1.23405	1.19123	1.15742	1.07123
	Min	0.75944	0.80801	0.83390	0.86286	0.88943	0.91867	0.96227
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LR	Max	1.87925	1.59587	1.44259	1.34245	1.27074	1.21653	1.08929
	Min	0.57544	0.68602	0.75029	0.80263	0.84724	0.89175	0.95275
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LM	Max	3.02014	2.25530	1.83651	1.59378	1.43847	1.33169	1.11720
	Min	0.26502	0.44629	0.58958	0.69012	0.77198	0.84620	0.93841
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
		n=15						
W	Max	1.74955	1.53706	1.41503	1.33135	1.26895	1.22012	1.09785
	Min	0.69733	0.75549	0.78755	0.82333	0.85667	0.89384	0.94988
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LR	Max	2.07654	1.74144	1.55445	1.43073	1.34155	1.27385	1.11402
	Min	0.56453	0.66521	0.72454	0.77702	0.82362	0.87234	0.94190
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LM	Max	2.71119	2.13546	1.80729	1.60031	1.45887	1.35664	1.13556
	Min	0.37245	0.52682	0.63051	0.70918	0.77649	0.84260	0.93153
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
		n=20						
W	Max	1.89190	1.63684	1.48996	1.38955	1.31503	1.25697	1.11307
	Min	0.66637	0.72880	0.76383	0.80287	0.83954	0.88072	0.94316
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LR	Max	2.17058	1.81391	1.61111	1.47585	1.37795	1.30344	1.12694
	Min	0.56361	0.65827	0.71423	0.76609	0.81306	0.86334	0.93654
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LM	Max	2.61346	2.09784	1.79814	1.60370	1.46767	1.36747	1.14411
	Min	0.42706	0.56173	0.64797	0.71763	0.77885	0.84136	0.92853
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
		n=30						
W	Max	2.05283	1.75059	1.57509	1.45530	1.36674	1.29809	1.12970
	Min	0.63567	0.70201	0.73989	0.78207	0.82201	0.86718	0.93609
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LR	Max	2.26080	1.88572	1.66800	1.52148	1.41492	1.33359	1.14020
	Min	0.56530	0.65330	0.70541	0.75629	0.80330	0.85480	0.93126
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LM	Max	2.53770	2.06856	1.79120	1.60710	1.47577	1.37746	1.15230
	Min	0.47982	0.59349	0.66385	0.72550	0.78123	0.84038	0.92579
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000

Table 3: Max & Min Guaranteed Efficiency of $PTRRE$ compared to $URRE$ ($k = 0.75$)

		n=10						
Test	α:	5%	10%	15%	20%	25%	30%	50%
W	Max	1.46398	1.33893	1.26571	1.21458	1.17583	1.14507	1.06603
	Min	0.73305	0.78511	0.81209	0.84325	0.87157	0.90247	0.95496
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LR	Max	1.77250	1.53329	1.40015	1.31168	1.2476	1.19875	1.08266
	Min	0.53845	0.65284	0.71959	0.77539	0.8229	0.87008	0.94352
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LM	Max	2.65305	2.07674	1.73699	1.53149	1.39653	1.30211	1.10831
	Min	0.23240	0.40524	0.54722	0.65114	0.73706	0.81542	0.92625
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
		n=15						
W	Max	1.66412	1.48253	1.37592	1.30180	1.24600	1.20200	1.09054
	Min	0.66713	0.72842	0.76134	0.79932	0.83442	0.87329	0.94031
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LR	Max	1.93394	1.65728	1.49758	1.38973	1.31088	1.25040	1.10539
	Min	0.52818	0.63162	0.69242	0.74770	0.79667	0.84764	0.93076
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LM	Max	2.42676	1.98137	1.71260	1.53711	1.41441	1.32429	1.12511
	Min	0.33591	0.48757	0.59156	0.67307	0.74327	0.81226	0.91835
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
		n=20						
W	Max	1.78298	1.56842	1.44158	1.35344	1.28725	1.23523	1.10452
	Min	0.63471	0.69995	0.73564	0.77679	0.81517	0.85798	0.93239
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LR	Max	2.00947	1.71813	1.54638	1.42927	1.34318	1.27690	1.11723
	Min	0.52770	0.62470	0.68168	0.73600	0.78507	0.83734	0.92450
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LM	Max	2.35339	1.95113	1.70494	1.54002	1.42212	1.33390	1.13293
	Min	0.38995	0.52389	0.61060	0.68279	0.74642	0.81129	0.91495
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
		n=30						
W	Max	1.91475	1.66499	1.51540	1.41129	1.33325	1.27212	1.11976
	Min	0.60285	0.67159	0.70989	0.75404	0.79557	0.84226	0.92410
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LR	Max	2.08109	1.77787	1.59502	1.46901	1.37582	1.30380	1.12936
	Min	0.52988	0.61988	0.67257	0.72559	0.77440	0.82761	0.91835
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LM	Max	2.29589	1.92749	1.69912	1.54294	1.42920	1.34274	1.14040
	Min	0.44302	0.55723	0.62800	0.69183	0.74954	0.81060	0.91184
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000

Table 4: Max & Min Guaranteed Efficiency of $PTRRE$ compared to $URLSE$ ($k = 0.10$)

		n=10						
Test	α:	5%	10%	15%	20%	25%	30%	50%
W	Max	1.88767	1.69523	1.58575	1.51063	1.45444	1.41027	1.29849
	Min	0.99756	1.04366	1.07034	1.09768	1.12349	1.15254	1.18079
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LR	Max	2.39421	1.99742	1.78856	1.65422	1.55902	1.48761	1.32181
	Min	0.81958	0.93339	0.99940	1.04931	1.09225	1.13580	1.17449
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LM	Max	4.14660	2.94323	2.33346	1.99455	1.78300	1.63989	1.35797
	Min	0.45902	0.69089	0.85481	0.95641	1.03651	1.10859	1.16514
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
		n=15						
W	Max	2.21086	1.91682	1.75143	1.63942	1.55666	1.49232	1.33289
	Min	0.93729	0.99494	1.02924	1.06414	1.09744	1.13529	1.17200
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LR	Max	2.67895	2.19950	1.94060	1.77256	1.65301	1.56313	1.35385
	Min	0.80494	0.91086	0.97419	1.02594	1.07239	1.12167	1.16666
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LM	Max	3.64574	2.76540	2.29212	2.00353	1.81055	1.67315	1.38181
	Min	0.59054	0.77284	0.88887	0.96886	1.03656	1.10327	1.15978
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
		n=20						
W	Max	2.41226	2.05392	1.85268	1.71719	1.61771	1.54084	1.35261
	Min	0.90616	0.96945	1.00765	1.04639	1.08357	1.12605	1.16720
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LR	Max	2.81723	2.30147	2.01841	1.83355	1.70165	1.60233	1.37061
	Min	0.80239	0.90284	0.96370	1.01565	1.06327	1.11494	1.16274
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LM	Max	3.49157	2.71013	2.27920	2.00820	1.82246	1.68763	1.39294
	Min	0.65278	0.80705	0.90332	0.97454	1.03696	1.10112	1.15738
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
		n=30						
W	Max	2.64435	2.21232	1.96888	1.80572	1.68666	1.59523	1.37420
	Min	0.87455	0.94333	0.98547	1.02805	1.06918	1.11642	1.16211
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LR	Max	2.95147	2.40344	2.09709	1.89557	1.75127	1.64241	1.38785
	Min	0.80260	0.89672	0.95450	1.00624	1.05467	1.10843	1.15885
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LM	Max	3.37342	2.66731	2.26941	2.01289	1.83344	1.70099	1.40360
	Min	0.71055	0.83763	0.91647	0.97991	1.03752	1.09923	1.15515
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000

Table 5: Max & Min Guaranteed Efficiency of $PTRRE$ compared to $URLSE$ ($k = 0.75$)

		n=10						
Test	α:	5%	10%	15%	20%	25%	30%	50%
W	Max	3.93068	3.59493	3.39835	3.26105	3.15703	3.07444	2.86221
	Min	1.96818	2.10797	2.18041	2.26406	2.34011	2.42308	2.56401
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LR	Max	4.75905	4.11678	3.75930	3.52178	3.34973	3.21856	2.90688
	Min	1.44571	1.75283	1.93204	2.08187	2.20943	2.33612	2.53328
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LM	Max	7.12325	5.57590	4.66369	4.11196	3.74959	3.49608	2.97572
	Min	0.62398	1.08804	1.46925	1.74827	1.97895	2.18934	2.48692
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
		n=15						
W	Max	4.46804	3.98049	3.69425	3.49525	3.34543	3.22728	2.92802
	Min	1.79120	1.95576	2.04415	2.14611	2.24037	2.34472	2.52467
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LR	Max	5.19249	4.44968	4.02090	3.73132	3.51962	3.35723	2.96790
	Min	1.41814	1.69586	1.85911	2.00751	2.13900	2.27587	2.49902
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LM	Max	6.51569	5.31984	4.5982	4.12703	3.79760	3.55563	3.02084
	Min	0.90189	1.30909	1.5883	1.80714	1.99563	2.18087	2.46571
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
		n=20						
W	Max	4.78718	4.21111	3.87054	3.63388	3.45618	3.31651	2.96555
	Min	1.70416	1.87932	1.97514	2.08564	2.18866	2.30361	2.50342
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LR	Max	5.39528	4.61306	4.15193	3.83749	3.60634	3.42840	2.99968
	Min	1.41684	1.67728	1.83026	1.97611	2.10785	2.24819	2.48223
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LM	Max	6.31868	5.23864	4.57765	4.13486	3.81829	3.58143	3.04183
	Min	1.04699	1.40661	1.63943	1.83324	2.00408	2.17825	2.45657
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
		n=30						
W	Max	5.14098	4.47040	4.06875	3.78920	3.57969	3.41555	3.00647
	Min	1.61862	1.80316	1.90601	2.02454	2.13604	2.26142	2.48113
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LR	Max	5.58758	4.77345	4.28253	3.94420	3.69398	3.50061	3.03225
	Min	1.42270	1.66432	1.80581	1.94815	2.07921	2.22209	2.46571
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000
LM	Max	6.16431	5.17518	4.56203	4.14270	3.83730	3.60517	3.06191
	Min	1.18948	1.49613	1.68613	1.85753	2.01246	2.17640	2.44823
	Δ	10.00000	8.00000	6.00000	5.00000	4.00000	3.00000	3.00000