J. Japan Statist. Soc.
Vol. 33 No. 1 2003 119–136

EFFECT OF W, LR, AND LM TESTS ON THE PERFORMANCE OF PRELIMINARY TEST RIDGE REGRESSION ESTIMATORS

B. M. Golam Kibria* and A. K. Md. E. Saleh**

The preliminary test ridge regression estimators ($PTRRE$) based on the Wald (W), Likelihood Ratio (LR) and Lagrangian Multiplier (LM) tests are considered in this paper. Using risks, 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 shrinkage parameter and departure parameter are given. 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 conclude that the optimum choice of the level of significance becomes the traditional choice by using the W test for all non-negative shrinkage parameter.

Key words and phrases: Lagrangian multiplier, likelihood ratio test, preliminary test, ridge regression, risk, 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 parameters β when it is *apriori* suspected but not certain 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.

Received May 27, 2002. Revised November 29, 2002. Accepted February 3, 2003.
*Department of Statistics, Florida International University, Miami, FL 33199, U.S.A.
Email: kibriag@fiu.edu
**Department of Mathematics and Statistics, Carleton University, Ottawa, Ontario K1S 5B6, Canada.
Email: esaleh@math.carleton.ca

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

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

where $C = X'X$ is the information matrix. The corresponding maximum likelihood estimator of σ^2 is given by

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

It is observed from (1.3) that the $URLSE$ 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 resolve this problem, they suggested to use $C(k) = X'X + kI_p, (k \geq 0)$ 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* ($URRE$) of β,

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

where $R = [I_p + kC^{-1}]^{-1}$ is the ridge or biasing parameter and $k \geq 0$ is the shrinkage parameter.

Suppose β^* be any estimator of β, D be the positive semi definite matrix and consider the following quadratic loss function

$$L(\beta^*, \beta) = (\beta^* - \beta)'D(\beta^* - \beta).$$

Then the risk function of β^* is defined by

$$Risk = E[L(\beta^*, \beta)] = E[(\beta^* - \beta)'D(\beta^* - \beta)] = tr(DM) = Tr(M),$$

for $D = I_p$, where M is the mean-squared error matrix of β^*.

The bias and the risk of the $URRE$ of β are

$$B(\tilde{\beta}(k)) = E(\tilde{\beta}(k) - \beta) = -kC^{-1}(k)\beta \quad \text{and}$$
$$R(\tilde{\beta}(k)) = \sigma^2 tr(WC^{-1}W') + k^2\beta'C^{-2}(k)\beta, \tag{1.5}$$

respectively. Though these estimators in (1.4) result in biased, for certain value of k, they yield minimum mean square error compared to unrestricted least squares 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 the maximum likelihood estimator of σ^2 are

$$\hat{\beta} = \tilde{\beta} - C^{-1}H'(HC^{-1}H')^{-1}(H\tilde{\beta} - h) \quad \text{and}$$
$$\hat{\sigma}^2 = \frac{(Y - X\hat{\beta})'(Y - X\hat{\beta})}{n}$$

respectively. Based on the $RLSE$, Sarkar (1992) proposed the following *restricted ridge regression estimator* ($RRRE$) of β,

$$\hat{\beta}(k) = R\hat{\beta}. \tag{1.6}$$

The bias and risk of the $RRRE$ of β are, respectively,

$$\begin{aligned} B(\hat{\beta}(k)) =& - R\eta - kC^{-1}(k)\beta \quad \text{and} \\ R(\hat{\beta}(k)) =& \sigma^2\left[tr(RC^{-1}R') - tr(RAR')\right] + \eta' R' R\eta \\ &+ 2k\eta' R' C^{-1}(k)\beta + k^2\beta' C^{-2}\beta, \end{aligned} \tag{1.7}$$

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 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) define the $PTRRE$ of β as,

$$\hat{\beta}^{PT}(k) = R\hat{\beta}^{PT}, \tag{1.8}$$

where $\hat{\beta}^{PT} = \hat{\beta} + \{1 - I(\mathcal{L}_n \leq \mathcal{L}_{n,\alpha})\}(\tilde{\beta} - \hat{\beta})$ is the usual preliminary test least squares estimator ($PTLSE$). Here, $\mathcal{L}_n$ 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), Han and Bancroft (1968), Giles (1991) and Kibria and Saleh (1993) among others. The ridge regression approach has been studied by Hoerl and Kennard (1970), McDonald and Galarneau (1975), Lawless (1978), Gibbons (1981), Sarkar (1992), Saleh and Kibria (1993) and Kibria (1996) to mention a few.

The $PTRRE$ proposed by Saleh and Kibria (1993) have used the test statistic $\mathcal{L}_n$ for testing the null hypothesis (1.2), which follows a central F-distribution with appropriate degrees of freedoms. 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 based on W, LR and LM tests. Section 3 contains the biases and risks 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

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

where $URSS = (Y - X\tilde{\beta})'(Y - X\tilde{\beta})$ is the unrestricted residual sum of squares and $RRSS = (Y - X\hat{\beta})'(Y - X\hat{\beta})$ 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

$$\Delta = \frac{(H\beta - h)'(HC^{-1}H')^{-1}(H\beta - h)}{\sigma^2} = \frac{\eta' C \eta}{\sigma^2} \tag{2.1}$$

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:

$$\mathcal{L}_W = \frac{(H\tilde{\beta} - h)'(HC^{-1}H')^{-1}(H\tilde{\beta} - h)}{\tilde{\sigma}^2} = \frac{nq}{n-p}F. \tag{2.2}$$

The well known LR test is

$$\mathcal{L}_{LR} = n\Big[\ln\hat{\sigma}^2 - \ln\tilde{\sigma}^2\Big] = n\ln\left(1 + \frac{\mathcal{L}_W}{n}\right). \tag{2.3}$$

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

$$\mathcal{L}_{LM} = \frac{(H\tilde{\beta} - h)'(HC^{-1}H')^{-1}(H\tilde{\beta} - h)}{\hat{\sigma}^2} = \frac{\mathcal{L}_W}{1 + \mathcal{L}_W/n}. \tag{2.4}$$

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

$$\mathcal{L}_W \geq \mathcal{L}_{LR} \geq \mathcal{L}_{LM} \tag{2.5}$$

exists among these three tests. From equations (2.2) to (2.4), we observed 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 risk. However, due to the inequality relation among the value of test statistics, the *PTRREs* based on a fixed critical value may have different biases and risks.

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 restriction (1.2), all tests are asymptotically distributed as χ^2-random variable with q degrees of freedom. We propose the following *PTRRE* based on W, LR and LM tests, which are given below,

$$\hat{\beta}_*^{PT}(k) = \hat{\beta}(k) I(\mathcal{L}_* \leq \chi_q^2(\alpha)) + \tilde{\beta}(k) I(\mathcal{L}_* > \chi_q^2(\alpha)), \tag{2.6}$$

where $(*)$ stands for either of W, LR and LM tests and $\chi_q^2(\alpha)$ is the upper percentile of the central χ^2 distribution with q degrees of freedom. For $k = 0$, we obtain the *PTLSE* based on W, LR and LM tests, which have been considered by Billah and Saleh (2000).

For excellent reference on W, LR and LM tests, readers are refereed to Judge *et al.* (1988) 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), Billah and Saleh (2000), and recently Kibria (2002) and Kibria and Saleh (2003) among others. In the following section, we will provide the biases and risks of the proposed estimators.

3. Biases and risks of the estimators

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

The biases of the proposed estimators are as follows:

$$B(\hat{\beta}_*^{PT}) = -R\eta G_{q+2,n-p}(l_1^*; \Delta) - kC^{-1}(k)\beta, \tag{3.1}$$

where $(*)$ stands for either of W, LR and LM tests and

$$\begin{aligned} l_1^W &= \left(\frac{n-p}{q+2}\right) \frac{\chi_\alpha^2(q)}{n}, \\ l_1^{LR} &= \left(\frac{n-p}{q+2}\right) \left(e^{\frac{\chi_\alpha^2(q)}{n}} - 1\right), \\ l_1^{LM} &= \left(\frac{n-p}{q+2}\right) \frac{\chi_\alpha^2(q)}{(n - \chi_\alpha^2(q))}, \end{aligned} \tag{3.2}$$

also $G_{q+2,n-p}(.; \Delta)$ 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*, however, for $\alpha = 0$, we do not reject the null hypothesis and the bias of the proposed estimators coincide with that of the *RRRE*. As $\Delta \to \infty$, $B(\hat{\beta}_W^{PT}(k)) = B(\hat{\beta}_{LR}^{PT}(k)) =$

$B(\hat{\beta}^{PT}_{LM}(k)) = B(\tilde{\beta}(k))$, 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

$$G_{q+2,n-p}(l_1^{LM};\Delta) \geq G_{q+2,n-p}(l_1^{LR};\Delta) \geq G_{q+2,n-p}(l_1^{W};\Delta). \tag{3.3}$$

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

THEOREM 1. *Under the null hypothesis the proposed estimators are biased and the amount of bias are same. However, under the alternative hypothesis the dominance picture of the proposed estimators is*

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

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

The risk for the $PTRRE$ based on the W, LR and LM are provided below:

$$\begin{aligned} R(\hat{\beta}^{PT}_{*}(k)) =& \sigma^2 tr(RC^{-1}R') - \sigma^2 tr(RAR')G_{q+2,n-p}(l_1^{*};\Delta) \\ &+ \eta' R' R\eta \times [2G_{q+2,n-p}(l_1^{*};\Delta) - G_{q+4,n-p}(l_2^{*};\Delta)] \\ &+ 2kG_{q+2,n-p}(l_1^{*};\Delta)\eta' R' C^{-1}(k)\beta + k^2\beta' C^{-2}(k)\beta, \end{aligned} \tag{3.4}$$

where $(*)$ stands for either of W, LR and LM tests and

$$\begin{aligned} l_2^{W} &= \left(\frac{n-p}{q+4}\right)\frac{\chi^2_\alpha(q)}{n}, \\ l_2^{LR} &= \left(\frac{n-p}{q+4}\right)\left(e^{\frac{\chi^2_\alpha(q)}{n}} - 1\right), \\ l_2^{LM} &= \left(\frac{n-p}{q+4}\right)\frac{\chi^2_\alpha(q)}{(n-\chi^2_\alpha(q))}, \end{aligned} \tag{3.5}$$

also $G_{q+4,n-p}(.;\Delta)$ 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

In this Section we will compare the performance of the proposed estimators by using risks. We note from (3.4) that for given α and known data, the risks depend on the departure parameter Δ and shrinkage 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 as a function of Δ

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

$$\begin{aligned} &\gamma_p \leq \frac{\eta' R' R\eta}{\eta' C\eta} \leq \gamma_1, \quad \text{or} \\ &\sigma^2\Delta\gamma_p \leq \eta' R' R\eta \leq \sigma^2\Delta\gamma_1, \end{aligned} \tag{4.1}$$

where γ_1 and γ_p are the largest and the smallest characteristic roots of the matrix $(R'RC^{-1})$.

Now we compare between $\hat{\beta}_W^{PT}(k)$ and $\hat{\beta}_{LR}^{PT}(k)$. The risk difference is:

$$R(\hat{\beta}_W^{PT}(k)) - R(\hat{\beta}_{LR}^{PT}(k)) = \sigma^2 tr(RAR')\psi - \eta' R' R\eta[2\psi - \psi^*] - 2k\eta' R' C^{-1}(k)\beta\psi, \tag{4.2}$$

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 (3.3) that both ψ and ψ^* are positive for all k, Δ and α.

The difference in (4.2) is non-negative (≥ 0), whenever

$$\Delta \leq \frac{tr(RAR') - 2k\sigma^{-2}\eta' R' C^{-1}(k)\beta}{\gamma_1\left(2 - \frac{\psi^*}{\psi}\right)} = \Delta_1(k,\alpha). \tag{4.3}$$

Thus, $\hat{\beta}_{LR}^{PT}(k)$ performs better than $\hat{\beta}_W^{PT}(k)$, when (4.3) holds. However, $\hat{\beta}_W^{PT}(k)$ performs better than $\hat{\beta}_{LR}^{PT}(k)$, whenever

$$\Delta > \frac{tr(RAR') - 2k\sigma^{-2}\eta' R' C^{-1}(k)\beta}{\gamma_p\left(2 - \frac{\psi^*}{\psi}\right)} = \Delta_2(k,\alpha). \tag{4.4}$$

Under the null hypothesis, the difference in (4.2) is always positive for all α, therefore, $\hat{\beta}_{LR}^{PT}(k)$ is superior to $\hat{\beta}_W^{PT}(k)$.

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

$$R(\hat{\beta}_{LR}^{PT}(k)) - R(\hat{\beta}_{LM}^{PT}(k)) = \sigma^2 tr(RAR')\psi_1 - \eta' R' R\eta[2\psi_1 - \psi_1^*] - 2k\eta' R' C^{-1}(k)\beta\psi_1, \tag{4.5}$$

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 (4.5) is non-negative (≥ 0), whenever

$$\Delta \leq \frac{tr(RAR') - 2k\sigma^{-2}\eta' R' C^{-1}(k)\beta}{\gamma_1\left(2 - \frac{\psi_1^*}{\psi_1}\right)} = \Delta_3(k,\alpha). \tag{4.6}$$

Thus, $\hat{\beta}_{LM}^{PT}(k)$ performs better than $\hat{\beta}_{LR}^{PT}(k)$ when (4.6) holds, otherwise $\hat{\beta}_{LR}^{PT}(k)$ performs better than $\hat{\beta}_{LM}^{PT}(k)$, whenever

$$\Delta > \frac{tr(RAR') - 2k\sigma^{-2}\eta' R' C^{-1}(k)\beta}{\gamma_p\left(2 - \frac{\psi_1^*}{\psi_1}\right)} = \Delta_4(k,\alpha). \tag{4.7}$$

Under the null hypothesis the difference in (4.5) 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

$\hat{\beta}_*^{PT}(k)$ ($*$ stands for either of W, LR and LM tests) as follows. At $\Delta = 0$, it assumes a value

$$\sigma^2 tr(RC^{-1}R') - \sigma^2 tr(RAR')G_{q+2,n-p}(l_1^*; 0) + k^2\beta' C^{-2}(k)\beta,$$

then increases from 0, crossing the risk of $\tilde{\beta}(k)$ to a maximum and then drops gradually towards the risk of $\tilde{\beta}(k)$ 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*

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

where $\geq$ denotes the dominance in the sense of having smaller risk.
Under the alternative hypothesis, the dominance picture of the proposed estimators is

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

in the interval

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

where $\Delta_{13}^*(k,\alpha) = \min\{\Delta_1(k,\alpha), \Delta_3(k,\alpha)\}$, also $\Delta_1(k,\alpha)$ and $\Delta_3(k,\alpha)$ are given in (4.3) *and* (4.6) *respectively, while*

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

in the interval

$$\Delta \in (\Delta_{24}^*(k,\alpha), \infty)\,,$$

where $\Delta_{24}^*(k,\alpha) = \max\{\Delta_2(k,\alpha), \Delta_4(k,\alpha)\}$, also $\Delta_2(k,\alpha)$ and $\Delta_4(k,\alpha)$ are given in (4.4) *and* (4.7) *respectively. For $k = 0$, the results in Theorem* 2 *will coincide with that of Billah and Saleh* (2000).

4.2. Performance based on k

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

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

where $\lambda_1, \lambda_2, ..., \lambda_p$ denote the eigen values of the matrix C. As C is symmetric we can write

$$RAR' = Q[\Lambda + kI_p]^{-1}\Lambda A^*\Lambda[\Lambda + kI_p]^{-1}Q', \tag{4.8}$$

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

$$tr(RC^{-1}R') = \sum_{i=1}^{p} \frac{\lambda_i}{(\lambda_i + k)^2} \quad \text{and} \quad tr(RAR') = \sum_{i=1}^{p} \frac{\lambda_i^2 a_{ii}^*}{(\lambda_i + k)^2}, \tag{4.9}$$

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

$$\beta' C^{-2}(k)\beta = \sum_{i=1}^{p} \frac{\alpha_i^2}{(\lambda_i + k)^2}, \quad \text{where} \quad \alpha = Q'\beta, \tag{4.10}$$

$$\eta' R' R \eta = \sum_{i=1}^{p} \frac{\lambda_i^2 \eta_i^{*2}}{(\lambda_i + k)^2} \quad \text{and} \quad \eta' R' C^{-1}(k)\beta = \sum_{i=1}^{p} \frac{\alpha_i \lambda_i {\eta_i}^*}{(\lambda_i + k)^2}, \tag{4.11}$$

where $\eta^* = \eta' Q$. Using equations (4.9) to (4.11), the risk difference in equation (4.2) can be expressed in terms of the eigen values as

$$R(\hat{\beta}_W^{PT}(k)) - R(\hat{\beta}_{LR}^{PT}(k)) = \sum_{i=1}^{p} \frac{\lambda_i}{(\lambda_i + k)^2} \left[\sigma^2 \psi a_{ii}^* \lambda_i - (2\psi - \psi^*)\lambda_i \eta_i^{*2} - 2\psi k \eta_i^* \alpha_i\right]. \tag{4.12}$$

The difference in (4.12) will be non-negative (≥ 0) if

$$k \leq \frac{\min\left[\sigma^2 \psi a_{ii}^* \lambda_i - (2\psi - \psi^*)\lambda_i \eta_i^{*2}\right]}{\max[2\psi \eta_i^* \alpha_i]} = k_1(\alpha, \Delta). \tag{4.13}$$

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

$$k > \frac{\max\left[\sigma^2 \psi a_{ii}^* \lambda_i - (2\psi - \psi^*)\lambda_i \eta_i^{*2}\right]}{\min[2\psi \eta_i^* \alpha_i]} = k_2(\alpha, \Delta). \tag{4.14}$$

Now we compare between $\hat{\beta}_{LR}^{PT}(k)$ and $\hat{\beta}_{LM}^{PT}(k)$ estimators. As before, the risk difference between $\hat{\beta}_{LR}^{PT}(k)$ and $\hat{\beta}_{LM}^{PT}(k)$ is:

$$R(\hat{\beta}_{LR}^{PT}(k)) - R(\hat{\beta}_{LM}^{PT}(k)) = \sum_{i=1}^{p} \frac{\lambda_i}{(\lambda_i + k)^2} \left[\sigma^2 \psi_1 a_{ii}^* \lambda_i - (2\psi_1 - \psi_1^*)\lambda_i \eta_i^{*2} - 2\psi_1 k \eta_i^* \alpha_i\right]. \tag{4.15}$$

The difference in (4.15) will be non-negative (≥ 0) if

$$k \leq \frac{\min\left[\sigma^2 \psi_1 a_{ii}^* \lambda_i - (2\psi_1 - \psi_1^*)\lambda_i \eta_i^{*2}\right]}{\max[2\psi_1 \eta_i^* \alpha_i]} = k_3(\alpha, \Delta). \tag{4.16}$$

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

$$k > \frac{\max\left[\sigma^2 \psi_1 a_{ii}^* \lambda_i - (2\psi_1 - \psi_1^*)\lambda_i \eta_i^{*2}\right]}{\min[2\psi_1 \eta_i^* \alpha_i]} = k_4(\alpha, \Delta). \tag{4.17}$$

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*

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

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* (4.13) *and* (4.16) *respectively, while*

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

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* (4.14) *and* (4.17) *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:*

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

in the interval,

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

while

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

in the interval,

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

5. Computed risk analysis

In this section we will provide some graphical representation of the proposed estimators. Note that, for given α, the risks of the estimators depends 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 risks of the proposed estimators, we consider $H'H = I$, $\beta'\beta$=1, and $h = 0$. Using these restrictions in (3.4), the risks of the proposed estimators become:

$$(5.1)\quad R(\hat{\beta}_{*}^{PT}(k)) = \frac{1}{(1+k)^2}\Big[\sigma^2(p - qG_{q+2,n-p}(l_1^*;\Delta)) + \Delta[2G_{q+2,n-p}(l_1^*;\Delta) - G_{q+4,n-p}(l_2^*;\Delta)] + 2k\Delta G_{q+2,n-p}(l_1^*;\Delta) + k^2\Big],$$

where $(*)$ stands for either of W, LR and LM tests and l_i^* $(i = 1, 2)$ stands for either of the critical values l_i^W, l_i^{LR} and l_i^{LM} of the tests.

Note 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 $\Delta \in (0, \frac{q}{2-\frac{\psi^*}{\psi}+2k}]$ otherwise $\hat{\beta}_{W}^{PT}(k)$ is superior to $\hat{\beta}_{LR}^{PT}(k)$ if $\Delta \in (\frac{q}{2-\frac{\psi^*}{\psi}+2k}, \ \infty)$. Now, $\hat{\beta}_{LM}^{PT}(k)$ is superior to $\hat{\beta}_{LR}^{PT}(k)$ if $\Delta \in (0, \ \frac{p}{2-\frac{\psi_1^*}{\psi_1}+2k})$, otherwise $\hat{\beta}_{LR}^{PT}(k)$ will be superior

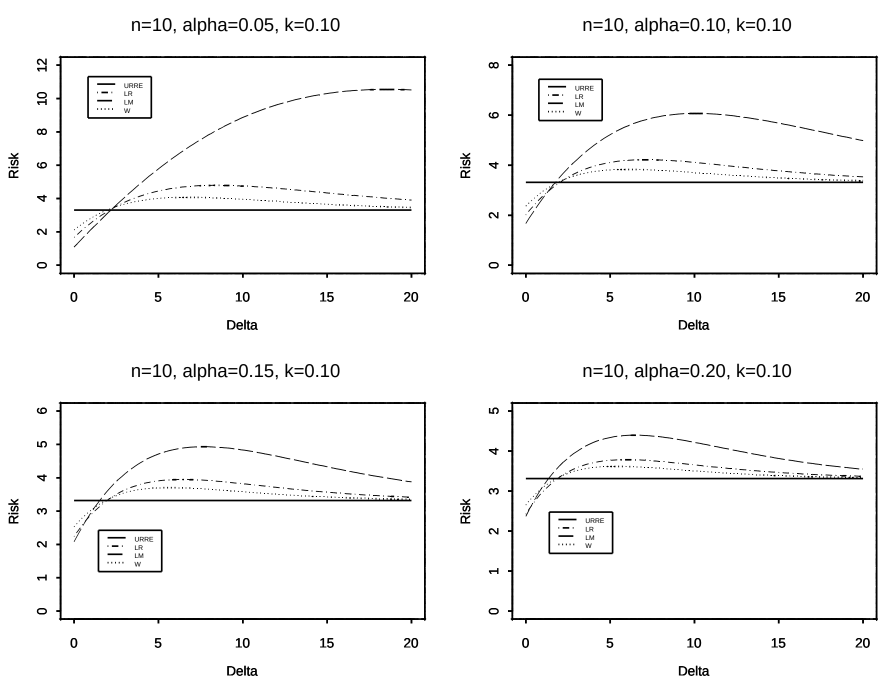

Figure 1. Risk function of the $PTRRE$ based on the W, LR and LM tests for different significance levels and fixed $n = 10$ and $k = 0.10$.

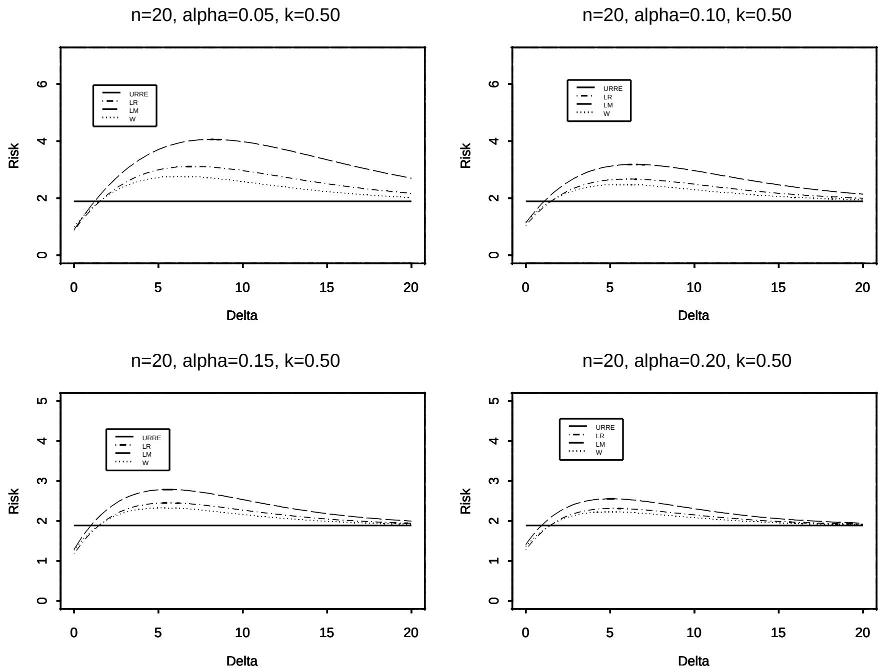

Figure 2. Risk function of the $PTRRE$ based on the W, LR and LM tests for different significance levels and fixed $n = 20$ and $k = 0.50$.

to $\hat{\beta}_{LM}^{PT}(k)$ if $\Delta \in [\frac{q}{2-\frac{\psi_1^*}{\psi_1}+2k},\ \infty)$. Similarly, for given Δ, $\hat{\beta}_{LR}^{PT}(k)$ is superior to $\hat{\beta}_{W}^{PT}(k)$ if $k \in (0, \frac{1}{2}(\frac{q}{\Delta} - (2 - \frac{\psi^*}{\psi}))]$ otherwise $\hat{\beta}_{W}^{PT}(k)$ is superior to $\hat{\beta}_{LR}(k)$. Similarly, $\hat{\beta}_{LM}^{PT}(k)$ is superior to $\hat{\beta}_{LR}^{PT}(k)$ if $k \in (0, \frac{1}{2}(\frac{q}{\Delta} - (2 - \frac{\psi_1^*}{\psi_1}))]$ otherwise $\hat{\beta}_{LR}^{PT}(k)$ will be 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 shrinkage parameter k. We have plotted the risk 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 and 2. The computation of the figures have been done by Splus software. From these figures we observe that the graphical analysis support the findings of the paper.

6. Efficiency analysis

In this section, we describe the relative efficiency of the proposed estimators for β. Accordingly, we provide max-min rule for the optimum choice of the level of significance for the $PTRRE$ of the null hypothesis. For a fixed value of $k(> 0)$, the relative efficiency of the $PTRRE$ compared to the $URRE$ is a function of α, and Δ. Let us denote this by

$$E(k, \alpha, \Delta) = \frac{R(\tilde{\beta}(k))}{R(\hat{\beta}_*^{PT}(k))} = [1 - h(k, \alpha, \Delta)]^{-1}, \tag{6.1}$$

where $(*)$ stands for either of W, LR and LM tests,

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

and

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

l_i^* $(i = 1, 2)$ stands for either of the critical values l_i^W, l_i^{LR} and l_i^{LM} of the tests. 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

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

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

$$\max_{0\leq\alpha\leq 1} E(k,\alpha,0) = E(k,1,0) = \left[1 - \frac{\sigma^2 tr(RAR')}{\sigma^2 tr(RC^{-1}R') + k^2\beta' C^{-2}(k)\beta}\right]^{-1}.$$

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 $0 \leq \Delta \leq \Delta_1(k,1)$, where $\Delta_1(k,1)$ is available from (4.3) for $\alpha = 1$. 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 $0 < \Delta < \Delta_1(k,1)$, $\hat{\beta}(k)$ 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_{Min} (minimum guaranteed efficiency) and consider the set $\mathcal{A} = \{\alpha | E(k,\alpha,\Delta) \geq E_{Min}\}$ 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

$$\max_{\alpha\in\mathcal{A}} \min_{\Delta} E(k,\alpha,\Delta) = E_{Min}. \tag{6.2}$$

The solution α^* for (6.2) gives the optimum choice of α and the value of $\Delta = \Delta_{min}(k)$ for which (6.2) 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.

$$\hat{k}(\alpha,\Delta) = \frac{\min\left[\sigma^2 a_{ii}^* \lambda_i G_{q+2,n-p}(l_1^*;\Delta) - \lambda_i \eta_i^{*2}\left\{2G_{q+2,n-p}(l_1^*;\Delta) - G_{q+4,n-p}(l_2^*;\Delta)\right\}\right]}{\max[2\eta_i^* \alpha_i G_{q+2,n-p}(l_1^*;\Delta)]}.$$

The above equation is obtained from the risk difference of $URRE$ and $PTRRE$ and based on the smaller risk criterion. We have not made any attempt for numerical computation of k. However, details discussion about the estimation procedure of k are avaliable in Gibbons (1981), McDonald and Galarneau (1975) and most recently Kibria (2003) among others.

For each estimator we can find the optimum significance level say α_*^W, α_*^{LR}, α_*^{LM} respectively, with minimum guaranteed efficiency E_{Min}. Then, we choose $\alpha_* = \min(\alpha_*^W, \alpha_*^{LR}, \alpha_*^{LM})$ as optimum level of significance. Note that our main goal is to choose the smallest level of significance (α) which gives the best estimator in the sense of highest efficiency. Imposing the restrictions on $X'X = I_p$, $H'H = I$, and $\beta'\beta = 1$, in equation (6.1), we obtain the minimum guaranteed efficency of the proposed estimators compared to $URRE$ (for $k = 0$, we obtain for $URLSE$ and the results will coincide with that of Billah and Saleh (2000)). Tables 1 and 2 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 $k = 0.10$ and 0.50 respectively. How can one use the table? For example, if $n = 10$, $p = 4$, $k = 0.10$, and the experimenter wishes to have an estimator with a minimum guaranteed efficiency of 0.80. 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 0.80007 and maximum efficiency 1.56396. Note that the size of α corresponding to the minimum guaranteed efficiency of 0.80 for $\hat{\beta}_{LR}^{PT}(k)$ and $\hat{\beta}_{LM}^{PT}(k)$ are 0.15 and 0.25 respectively. Therefore, we choose $\alpha_* = \min(0.05, 0.15, 0.25) = 0.05$, which corresponds to Wald test. Suppose, $n = 20$, $p = 4$, $k = 0.50$, and the experimenter wishes to have an estimator with a minimum guaranteed efficiency of 0.60. Now using Table 2, we select $\alpha = 0.05$, corresponding to $\hat{\beta}_W^{PT}(k)$, because such a choice of α would yield an estimator with a minimum guaranteed efficiency of 0.61809 and maximum efficiency of 1.89190. Note that the size of α corresponding to the minimum guaranteed efficiency of 0.60 for $\hat{\beta}_{LR}^{PT}(k)$ and $\hat{\beta}_{LM}^{PT}(k)$ are 0.10 and 0.15 respectively. Thus we choose, $\alpha_* = \min(0.05, 0.10, 0.15) = 0.05$, which again corresponds to Wald test. Therefore, from the application point of view it is evident that for all n, p, and k, the $PTRRE$ based Wald test will give the minimum guaranteed efficiency compared to $URRE$ among the three test procedures.

7. Summary and concluding remarks

In this paper we studied the effect of Wald, likelihood ratio and Lagrangian multiplier tests on the performance of the preliminary test ridge regression estimator for estimating the regression parameters when there exist a uncertain prior information in the parameter space. We have effectively determined some conditions on the departure parameter and the shrinkage 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

Table 1. Max & Min guaranteed efficiency of $PTRRE$s ($k = 0.10$).

		$n = 10$						
Test	α	5%	10%	15%	20%	25%	30%	50%
W	E_{Max}	1.56396	1.40452	1.31381	1.25158	1.20502	1.16843	1.07581
	E_{Min}	0.80007	0.85409	0.88635	0.90905	0.92626	0.93986	0.97410
	Δ_{Min}	6.69856	6.02871	5.64593	5.35885	5.07177	4.88038	4.40191
LR	E_{Max}	1.98363	1.65488	1.48185	1.37054	1.29167	1.23250	1.09514
	E_{Min}	0.67336	0.77081	0.82746	0.86606	0.89438	0.91608	0.96704
	Δ_{Min}	8.51675	7.08134	6.41148	5.93301	5.55024	5.26316	4.49761
LM	E_{Max}	3.43551	2.43850	1.93330	1.65251	1.47724	1.35867	1.12509
	E_{Min}	0.31104	0.55459	0.68753	0.77157	0.82902	0.87027	0.95597
	Δ_{Min}	19.25837	10.64593	8.25359	7.05742	6.33971	5.86124	4.66507
		$n = 15$						
W	E_{Max}	1.83172	1.58811	1.45108	1.35828	1.28972	1.23641	1.10432
	E_{Min}	0.74746	0.81245	0.85238	0.88099	0.90296	0.92052	0.96541
	Δ_{Min}	6.88995	6.12440	5.74163	5.35885	5.16746	4.97608	4.40191
LR	E_{Max}	2.21955	1.82231	1.60781	1.46859	1.36954	1.29508	1.12168
	E_{Min}	0.65577	0.74985	0.80693	0.84713	0.87744	0.90121	0.95945
	Δ_{Min}	8.13397	6.88995	6.22010	5.74163	5.45455	5.16746	4.40191
LM	E_{Max}	3.02054	2.29117	1.89905	1.65995	1.50006	1.38623	1.14484
	E_{Min}	0.48647	0.63996	0.73065	0.79255	0.83780	0.87223	0.95150
	Δ_{Min}	11.12440	8.37321	7.05742	6.33971	5.86124	5.50239	4.54545
		$n = 20$						
W	E_{Max}	1.99859	1.70170	1.53497	1.42272	1.34030	1.27661	1.12065
	E_{Min}	0.72092	0.79093	0.83455	0.86610	0.89050	0.91011	0.96066
	Δ_{Min}	6.98565	6.22010	5.74163	5.45455	5.16746	4.97608	4.40191
LR	E_{Max}	2.33411	1.90680	1.67227	1.51912	1.40984	1.32755	1.13557
	E_{Min}	0.64999	0.74169	0.79837	0.83890	0.86984	0.89438	0.95570
	Δ_{Min}	7.84689	6.69856	6.12440	5.74163	5.35885	5.16746	4.40191
LM	E_{Max}	2.89281	2.24537	1.88835	1.66382	1.50993	1.39823	1.15406
	E_{Min}	0.54036	0.66813	0.74594	0.80052	0.84143	0.87324	0.94959
	Δ_{Min}	9.56938	7.65550	6.69856	6.10048	5.62201	5.38278	4.54545
		$n = 30$						
W	E_{Max}	2.19088	1.83293	1.63124	1.49606	1.39742	1.32167	1.13854
	E_{Min}	0.69436	0.76904	0.81624	0.85069	0.87754	0.89923	0.95564
	Δ_{Min}	7.08134	6.22010	5.74163	5.45455	5.16746	4.97608	4.40191
LR	E_{Max}	2.44533	1.99128	1.73747	1.57050	1.45095	1.36076	1.14985
	E_{Min}	0.64583	0.73482	0.79081	0.83140	0.86278	0.88792	0.95201
	Δ_{Min}	7.65550	6.60287	6.02871	5.64593	5.35885	5.07177	4.40191
LM	E_{Max}	2.79492	2.20990	1.88024	1.66770	1.51903	1.40929	1.16290
	E_{Min}	0.58185	0.69065	0.75858	0.80735	0.84466	0.87424	0.94784
	Δ_{Min}	8.61244	7.17703	6.33971	5.86124	5.50239	5.26316	4.42584

Table 2. Max & Min guaranteed efficiency of $PTRRE$s ($k = 0.50$).

		$n=10$						
Test	α	5%	10%	15%	20%	25%	30%	50%
W	E_{Max}	1.51569	1.37315	1.29093	1.23405	1.19123	1.15742	1.07123
	E_{Min}	0.71448	0.78574	0.83035	0.86269	0.88774	0.90787	0.95977
	Δ_{Min}	6.12440	5.45455	5.07177	4.78469	4.59330	4.40191	3.92344
LR	E_{Max}	1.87925	1.59587	1.44259	1.34245	1.27074	1.21653	1.08929
	E_{Min}	0.56276	0.67760	0.75008	0.80210	0.84169	0.87287	0.94893
	Δ_{Min}	7.84689	6.50718	5.74163	5.35885	4.97608	4.68900	4.01914
LM	E_{Max}	3.02014	2.25530	1.83651	1.59378	1.43847	1.33169	1.11720
	E_{Min}	0.22079	0.43750	0.57873	0.67853	0.75214	0.80791	0.93209
	Δ_{Min}	18.42105	9.92823	7.65550	6.45933	5.74163	5.26316	4.18660
		$n=15$						
W	E_{Max}	1.74955	1.53706	1.41503	1.33135	1.26895	1.22012	1.09785
	E_{Min}	0.64950	0.73094	0.78387	0.82322	0.85428	0.87964	0.94657
	Δ_{Min}	6.31579	5.55024	5.16746	4.88038	4.59330	4.40191	3.92344
LR	E_{Max}	2.07654	1.74144	1.55445	1.43073	1.34155	1.27385	1.11402
	E_{Min}	0.54366	0.65239	0.72380	0.77678	0.81828	0.85180	0.93751
	Δ_{Min}	7.46411	6.22010	5.64593	5.16746	4.88038	4.68900	4.01914
LM	E_{Max}	2.71119	2.13546	1.80729	1.60031	1.45887	1.35664	1.13556
	E_{Min}	0.37214	0.52639	0.62932	0.70540	0.76428	0.81105	0.92551
	Δ_{Min}	10.40670	7.65550	6.45933	5.74163	5.26316	4.90431	4.06699
		$n=20$						
W	E_{Max}	1.89190	1.63684	1.48996	1.38955	1.31503	1.25697	1.11307
	E_{Min}	0.61809	0.70359	0.76018	0.80279	0.83677	0.86471	0.93941
	Δ_{Min}	6.41148	5.64593	5.16746	4.88038	4.68900	4.40191	3.92344
LR	E_{Max}	2.17058	1.81391	1.61111	1.47585	1.37795	1.30344	1.12694
	E_{Min}	0.53757	0.64278	0.71306	0.76596	0.80795	0.84223	0.93192
	Δ_{Min}	7.27273	6.12440	5.55024	5.16746	4.88038	4.59330	3.92344
LM	E_{Max}	2.61346	2.09784	1.79814	1.60370	1.46767	1.36747	1.14411
	E_{Min}	0.42389	0.55762	0.64792	0.71582	0.76934	0.81264	0.92271
	Δ_{Min}	8.85167	7.05742	6.10048	5.50239	5.14354	4.78469	4.06699
		$n=30$						
W	E_{Max}	2.05283	1.75059	1.57509	1.45530	1.36674	1.29809	1.12970
	E_{Min}	0.58753	0.67644	0.73635	0.78202	0.81879	0.84929	0.93189
	Δ_{Min}	6.50718	5.64593	5.26316	4.88038	4.68900	4.49761	3.92344
LR	E_{Max}	2.26080	1.88572	1.66800	1.52148	1.41492	1.33359	1.14020
	E_{Min}	0.53330	0.63483	0.70370	0.75624	0.79844	0.83329	0.92643
	Δ_{Min}	7.08134	6.02871	5.45455	5.07177	4.78469	4.59330	3.92344
LM	E_{Max}	2.53770	2.06856	1.79120	1.60710	1.47577	1.37746	1.15230
	E_{Min}	0.46571	0.58327	0.66357	0.72485	0.77391	0.81420	0.92018
	Δ_{Min}	7.89474	6.45933	5.74163	5.26316	4.90431	4.66507	3.94737

minimum guaranteed efficient estimators. The preliminary test ridge regression estimator 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. Finally, we recommend the practitioner to use the Wald test among these three test procedures when they consider the preliminary test ridge regression estimator for estimating the regression parameter β. Furthermore, the findings of this paper is also valid for $k = 0$, which agrees with that of Billah and Saleh (2000).

Acknowledgements

The authors are thankful to the referees for their valuable comments which improve the present version of the paper greatly.

References

Aitchison, J. and Silvey, D. (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, New York.

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. (2000). Performance of the PTE based on the conflicting W, LR and LM Tests in Regression model, *Advances on Methodological and Applied Aspects of Probability and Statistics*, (ed. Balakrishnan, N.), 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.

Han, C.-P. and Bancroft, T. A. (1968). On Pooling means 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 non-orthogonal 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.

Judge, G. G., Hill, R. C., Griffiths, W. F., Lutkepohl, H. and Lee, T.-C. (1988). *Introduction to the Theory and Practice of Econometrics*, Wiley, New York.

Kibria, B. M. G. (1996). On Preliminary test ridge regression estimators for linear restrictions in a regression model with non-normal disturbances, *Comm. Statist.*, **A25**, 2349–2369.

Kibria, B. M. G. (2002). Simultaneous shrinkage estimation of the mean of a multivariate normal distribution, *Journal of Statistical Research*, **36**, 29–44.

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

Kibria, B. M. G. and Saleh, A. K. Md. E. (1993). Performance of shrinkage preliminary test estimator in regression analysis, *Jahangirnagar Review*, **A17**, 133–148.

Kibria, B. M. G. and Saleh, A. K. Md. E. (2003). Estimation of the mean vector of a multivariate normal distribution under various test statistics, *Journal of Probability and Statistics*, to appear.

Lawless, J. F. (1978). Ridge and related estimation procedure, *Communications in Statistics*, **A7**, 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.

Rao, C. R. (1947). Large sample tests of statistical hypothesis concerning several parameters, *Proceedings of the Cambridge Philosophical Society*, Vol. 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. E. and Kibria, B. M. G. (1993). Performances of some new preliminary test ridge regression estimators and their properties, *Communications in Statistics-Theory and Methods*, **22**, 2747–2764.

Sarkar, N. (1992). A new estimator combining the ridge regression and the restricted least squares method of estimation, *Comm. Statist.*, **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.

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