Journal of the Korean Statistical Society (2004), **33**: *4, pp 411–433*

CONFLICT AMONG THE SHRINKAGE ESTIMATORS INDUCED BY W, LR AND LM TESTS UNDER A STUDENT'S t REGRESSION MODEL

B. M. GOLAM KIBRIA[1]

ABSTRACT

The shrinkage preliminary test ridge regression estimators (SPTRRE) based on Wald (W), Likelihood Ratio (LR) and Lagrangian Multiplier (LM) tests for estimating the regression parameters of the multiple linear regression model with multivariate Student's t error distribution are considered in this paper. The quadratic biases and risks of the proposed estimators are compared under both null and alternative hypotheses. It is observed that there is conflict among the three estimators with respect to their risks because of certain inequalities that exist among the test statistics. In the neighborhood of the restriction, the SPTRRE based on LM test has the smallest risk followed by the estimators based on LR and W tests. However, the SPTRRE based on W test performs the best followed by the LR and LM based estimators when the parameters move away from the subspace of the restrictions. Some tables for the maximum and minimum guaranteed efficiency of the proposed estimators have been given, which allow us to determine the optimum level of significance corresponding to the optimum estimator among proposed estimators. It is evident that in the choice of the smallest significance level to yield the best estimator the SPTRRE based on Wald test dominates the other two estimators.

AMS 2000 subject classifications. Primary 62J07; Secondary 62F03.
Keywords. Efficiency, Lagrangian multiplier, likelihood ratio test, ridge regression, risk, shrinkage estimation, Student's t, superiority, Wald test.

Received January 2004; accepted August 2004.

[1]Department of Statistics, Florida International University, Miami, FL 33199, U.S.A. (e-mail : kibriag@fiu.edu)

1. INTRODUCTION

Consider the linear regression model, $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}$, where $\mathbf{y} = (y_1, y_2, \ldots, y_n)'$ is an $n \times 1$ vector of observations on the dependent variable, $\mathbf{X}$ is an $n \times p$ matrix of full rank p, $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_p)'$ is an $p \times 1$ vector of parameters and $\mathbf{e} = (e_1, e_2, \ldots, e_n)'$ is an $n \times 1$ vector of errors, which is distributed according to the laws belonging to the class of spherical compound normal distributions with $E(\mathbf{e}) = \mathbf{0}$ and $E(\mathbf{ee}') = \sigma_e^2 \mathbf{I}_n$, where $\mathbf{I}_n$ is the n-dimensional identity matrix and σ_e^2 the common variance of e_i $(i = 1, 2, \ldots, n)$. This class is a subclass of the family of spherically symmetric distributions (SSDs) which can be expressed as a variance mixture of normal distributions, that is,

$$f(\mathbf{e}) = \int_0^\infty f(\mathbf{e}|\theta) g(\theta) d\theta, \tag{1.1}$$

where $f(\mathbf{e})$ is the probability density function (*pdf*) of $\mathbf{e}$, $f(\mathbf{e}|\theta)$ is the *pdf* of normal with mean vector $\mathbf{0}$ and variance-covariance matrix $\theta^2 \mathbf{I}_n$ and $g(\theta)$ is the *pdf* of θ with support $[0, \infty)$. In this case, $E(\theta^2) = \sigma_e^2$ and we write $\mathbf{e} \sim \text{SSD}(\{(\mathbf{0}, E(\theta^2)\mathbf{I}_n)\})$. Using (1.1), the multivariate Student's t distribution with mean vector, $E(\mathbf{e}) = \mathbf{0}$ and variance-covariance matrix, $E(\mathbf{ee}') = \nu(\nu - 2)^{-1} \sigma^2 \mathbf{I}_n = \sigma_e^2 \mathbf{I}_n$, $\nu > 2$ can be obtained if $g(\theta)$ be assumed to have an inverted gamma (IG) density with a scale parameter σ^2 and degrees of freedom ν.

In most applied as well as theoretical research work, the error terms in linear models are assumed to be normally and independently distributed. However, such assumptions may not be appropriate in many practical situation (for example, see Gnanadesikan, 1977; Zellner, 1976). It happens particularly if the error distributions have heavier tails. One can tackle such situation by using the well known t distribution as it has heavier tail than the normal distribution, specially for smaller degrees of freedom (*e.g.* Blattberg and Gonedes, 1974). For details readers are referred to Kelker (1970).

Our primary interest is to estimate the regression coefficients $\boldsymbol{\beta}$ when it is *a priori* suspected but not certain that $\boldsymbol{\beta}$ may be restricted to the subspace

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

where $\mathbf{H}$ is an $q \times p$ known matrix of full rank q $(< p)$ and $\mathbf{h}$ is an $q \times 1$ vector of known constants. The *unrestricted least squares estimator* (URLSE) of $\boldsymbol{\beta}$ is given by

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

where $\mathbf{C} = \mathbf{X}'\mathbf{X}$ is the information matrix. The corresponding unbiased estimator of σ_e^2 is $\tilde{\sigma}_e^2 = (\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}})'(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}})/m, \ m = n - p$.

Similarly, *the restricted least squares estimator* (RLSE) of $\boldsymbol{\beta}$ is given by

$$\hat{\boldsymbol{\beta}} = \tilde{\boldsymbol{\beta}} - \mathbf{C}^{-1}\mathbf{H}'(\mathbf{H}\mathbf{C}^{-1}\mathbf{H}')^{-1}(\mathbf{H}\tilde{\boldsymbol{\beta}} - \mathbf{h}) \tag{1.4}$$

and the corresponding estimator of σ_e^2 is $\hat{\sigma}_e^2 = (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})/(m + q)$, which is unbiased under the null hypothesis. Note that the restricted least squares estimator satisfies the condition $\mathbf{H}\hat{\boldsymbol{\beta}} = \mathbf{h}$.

Now we consider another estimator, called *shrinkage restricted least squares estimator* (SRLSE) of $\boldsymbol{\beta}$ as

$$\hat{\boldsymbol{\beta}}^{SR} = d\tilde{\boldsymbol{\beta}} + (1 - d)\hat{\boldsymbol{\beta}}, \tag{1.5}$$

where d is the coefficient of distrust and $0 \leq d \leq 1$. The value of d may completely be determined by the practitioner on the basis of his/her belief about the null hypothesis. For example, if the practitioner wishes to rely entirely on data and believes that the parameter space belongs to the restricted subspace, then he/she should use $d = 0$. The SRLSE is a modification of RLSE of $\boldsymbol{\beta}$. It yields smaller risk at and near the null hypothesis but performs poorly in the rest of the parameter space. However, SRLSE provides a wider range than the RLSE in which it dominates URLSE. This motivates us to replace the RLSE by SRLSE in the formulation of *shrinkage preliminary test least squares estimator* (SPTLSE) of $\boldsymbol{\beta}$ as

$$\hat{\boldsymbol{\beta}}^{SPT} = \hat{\boldsymbol{\beta}}^{SR} I(\mathcal{L}_n \leq \mathcal{L}_{n,\alpha}) + \tilde{\boldsymbol{\beta}} I(\mathcal{L}_n > \mathcal{L}_{n,\alpha}), \tag{1.6}$$

where $\mathcal{L}_n$ is the general test-statistic for testing the null hypothesis in (1.2), $\mathcal{L}_{n,\alpha}$ is the upper α-level critical value of $\mathcal{L}_n$ and $I(A)$ is the indicator function of the set A. Under the null hypothesis and normal theory, $\mathcal{L}_n$ follows a central F-distribution with (q, m) degrees of freedom, while under the alternative hypothesis it follows the non-central F-distribution with (q, m) degrees of freedom and non-centrality parameter $\Delta/2$, where

$$\Delta = \frac{(\mathbf{H}\boldsymbol{\beta} - \mathbf{h})'(\mathbf{H}\mathbf{C}^{-1}\mathbf{H}')^{-1}(\mathbf{H}\boldsymbol{\beta} - \mathbf{h})}{\sigma_e^2} \tag{1.7}$$

is the departure parameter from the null hypothesis. For $d = 0$, we obtain the preliminary test least squares estimator (PTLSE) of $\boldsymbol{\beta}$ as

$$\hat{\boldsymbol{\beta}}^{PT} = \hat{\boldsymbol{\beta}} I(\mathcal{L}_n \leq \mathcal{L}_{n,\alpha}) + \tilde{\boldsymbol{\beta}} I(\mathcal{L}_n > \mathcal{L}_{n,\alpha}).$$

The use of PTLSE is limited by the size of the PT as compared to the SPTLSE. It is interesting to note that the SPTLSE has a significant edge over the PTLSE with respect to the size of PT. The SPTLSE is a compromised estimator of $\boldsymbol{\beta}$ when it is suspected that $\mathbf{H}\boldsymbol{\beta} = \mathbf{h}$ holds and is a convex combination of $\tilde{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\beta}}^{SR}$. If H_0 is true then $\hat{\boldsymbol{\beta}}^{SPT} = \hat{\boldsymbol{\beta}}^{SR}$; otherwise $\hat{\boldsymbol{\beta}}^{SPT} = \tilde{\boldsymbol{\beta}}$. Also for $d = 0$, SPTLSE reduces to PTLSE. The preliminary test approach estimation has been pioneered by Bancroft (1944), followed by Bancroft (1964), Han and Bancroft (1968), Sen and Saleh (1987), Judge and Bock (1978), Saleh and Han (1990), Giles (1991), Kibria and Saleh (1993), Ahmed (2002) and recently Kibria and Saleh (2003a) among others.

It is observed from (1.3) that the usual least squares estimator (LSE) of $\boldsymbol{\beta}$ depends heavily on the characteristics of the information matrix $\mathbf{C} = \mathbf{X}'\mathbf{X}$. If the $\mathbf{C}$ matrix is ill-conditioned, the least squares estimator (LSE) produces unduly large sampling variances. Moreover, some of the regression coefficients may be statistically insignificant with wrong sign and meaningful statistical inference become difficult for the researcher. Hoerl and Kennard (1970) found that multicollinearity is a common problem in many fields of applications. To resolve this problem, they suggested to use $\mathbf{C}(k) = \mathbf{X}'\mathbf{X} + k\mathbf{I}_p$, $k \geq 0$ rather than $\mathbf{C}$ in the estimation of $\boldsymbol{\beta}$. The resulting estimator of $\boldsymbol{\beta}$ is known as the *unrestricted ridge regression estimator* (URRE) of $\boldsymbol{\beta}$ and defined as

$$\tilde{\boldsymbol{\beta}}(k) = \mathbf{R}(k)\tilde{\boldsymbol{\beta}}, \tag{1.8}$$

where $\mathbf{R}(k) = \left(\mathbf{I}_p + k\mathbf{C}^{-1}\right)^{-1}$ is the ridge or biasing parameter and $k \geq 0$ is the shrinkage parameter.

Based on the RLSE, Sarkar (1992) proposed the following *restricted ridge regression estimator* (RRRE) of $\boldsymbol{\beta}$,

$$\hat{\boldsymbol{\beta}}^{RE}(k) = \mathbf{R}(k)\hat{\boldsymbol{\beta}}. \tag{1.9}$$

Finally, based on the SPTLSE, we define the following *shrinkage preliminary test ridge regression estimator* (SPTRRE) of $\boldsymbol{\beta}$ as

$$\hat{\boldsymbol{\beta}}^{SPT}(k) = \hat{\boldsymbol{\beta}}^{SR}(k)I(\mathcal{L}_n \leq \mathcal{L}_{n,\alpha}) + \tilde{\boldsymbol{\beta}}(k)I(\mathcal{L}_n > \mathcal{L}_{n,\alpha}) = \mathbf{R}(k)\hat{\boldsymbol{\beta}}^{SPT}, \tag{1.10}$$

where $\hat{\boldsymbol{\beta}}^{SR}(k) = \mathbf{R}(k)\hat{\boldsymbol{\beta}}^{SR}$ is the *shrinkage restricted ridge regression estimator* (SRRRE) of $\boldsymbol{\beta}$ which is considered by Haq and Kibria (1996).

The ridge regression approach has been studied by Hoerl and Kennard (1970), Gibbons (1981), Sarkar (1992), Saleh and Kibria (1993), Aldrin (1997), Kibria and Ahmed (1997), Foucart (1999), Gunst (2000) and very recently Kibria and Saleh (2003b) to mention a few.

The main objective of this paper is to study the finite sample properties of the SPTRRE based on the Wald, the likelihood ratio and the Lagrangian multiplier tests. We assume a multiple linear regression model with Student's t disturbances for the estimation of regression coefficients in the model. The plan of this paper is as follows: In Section 2 we propose some shrinkage preliminary test ridge regression estimators. Section 3 contains the bias and the risk expressions of the estimators. In Section 4 we discuss the relative performance of the estimators. The maximum and minimum guaranteed efficiency are discussed in Section 5. Finally, concluding remarks have been presented in Section 6.

2. PROPOSED ESTIMATORS OF $\boldsymbol{\beta}$

In order to define the shrinkage preliminary least squares test estimators of $\boldsymbol{\beta}$, we consider three well-known test-statistics for testing $H_0 : \mathbf{H}\boldsymbol{\beta} = \mathbf{h}$ against $H_A : \mathbf{H}\boldsymbol{\beta} \neq \mathbf{h}$ with Student's t-error, namely (i) the Wald (W) test (ii) the likelihood ratio (LR) test and (iii) the Lagrangian multiplier (LM) test and they are given respectively

$$\begin{aligned} \mathcal{L}_W &= \lambda(n)\frac{nq}{m}F, \\ \mathcal{L}_{LR} &= n\log\left(1+\frac{q}{m}F\right), \\ \mathcal{L}_{LM} &= \lambda(n)^{-1}\left(\frac{nqF}{m+qF}\right), \end{aligned} \tag{2.1}$$

where

$$\lambda(n) = \frac{n+\nu}{n+\nu+2}, \quad 0 < \lambda(n) < 1$$

and

$$F = \frac{(\mathbf{X}\tilde{\boldsymbol{\beta}}-\mathbf{h})'\left(\mathbf{H}\mathbf{C}^{-1}\mathbf{H}'\right)^{-1}(\mathbf{X}\tilde{\boldsymbol{\beta}}-\mathbf{h})/q}{(\mathbf{y}-\mathbf{X}\tilde{\boldsymbol{\beta}})'(\mathbf{y}-\mathbf{X}\tilde{\boldsymbol{\beta}})/m}$$

is the test statistic for testing the null hypothesis (1.2) and follows a central F-distribution with (q, m) degrees of freedom under H_0 (see Zellner, 1976; King, 1980). For details, we refer to Ullah and Zinde-Walsh (1984). Note that if n is large then $\lambda(n)$ is close to 1 and the results in (2.1) also hold for normal regression model.

Ullah and Zinde-Walsh (1984) showed that for test statistics in (2.1), the following inequalities hold:

$$\begin{aligned}
&\mathcal{L}_W < \mathcal{L}_{LR} < \mathcal{L}_{LM}, && \text{if} \;\; \omega < \omega_1(\lambda(n)), \\
&\mathcal{L}_{LR} \le \mathcal{L}_W < \mathcal{L}_{LM}, && \text{if} \;\; \omega_1(\lambda(n)) \le \omega < \omega_2(\lambda(n)), \\
&\mathcal{L}_W \ge \mathcal{L}_{LM} > \mathcal{L}_{LR}, && \text{if} \;\; \omega_2(\lambda(n)) \le \omega < \omega_3(\lambda(n)), \\
&\mathcal{L}_W > \mathcal{L}_{LR} \ge \mathcal{L}_{LM}, && \text{if} \;\; \omega > \omega_3(\lambda(n)),
\end{aligned} \tag{2.2}$$

where $\omega = \mathcal{L}_W/n$, and $\omega_1(\lambda(n))$, $\omega_2(\lambda(n))$ and $\omega_3(\lambda(n))$ are respectively the unique positive roots (depending on $\lambda(n)$) of the equations

$$\begin{aligned}
\lambda(n)\omega - \log(1+\omega) &= 0, \\
\lambda(n)^2\omega - \omega(1+\omega)^{-1} &= 0, \\
\lambda(n)\log(1+\omega) - \omega(1+\omega)^{-1} &= 0.
\end{aligned}$$

It follows from (2.2) that the size of the Wald test can be greater or less than the LR test and LR test can be greater or less than the LM test depending on the solution for ω and the value of $\lambda(n)$.

The exact sampling distribution of the three test statistics can be complicated. Thus in practice the critical regions of the tests are commonly based on asymptotic approaches (see Kennedy, 1998, Chapter 2; Evans and Savin, 1982). It is known that the asymptotic distributions of the three tests are approximated by the chi-squared distribution with q degrees of freedom. Let the α level critical value of the distribution be $\chi^2_q(\alpha)$ as the first approximation. This choice of critical value for three tests leads to conflict as in the case of finite sample inference. The inequalities of statistics given in (2.2) will occur if

$$\begin{aligned}
&\text{either} \;\; \mathcal{L}_{LR} < \chi^2_q(\alpha) < \mathcal{L}_{LM} && \text{or} \;\; \mathcal{L}_W < \chi^2_q(\alpha) < \mathcal{L}_{LR}, \\
&\text{either} \;\; \mathcal{L}_W < \chi^2_q(\alpha) < \mathcal{L}_{LM} && \text{or} \;\; \mathcal{L}_{LR} < \chi^2_q(\alpha) < \mathcal{L}_W, \\
&\text{either} \;\; \mathcal{L}_{LM} < \chi^2_q(\alpha) < \mathcal{L}_W && \text{or} \;\; \mathcal{L}_{LR} < \chi^2_q(\alpha) < \mathcal{L}_{LM}, \\
&\text{either} \;\; \mathcal{L}_{LR} < \chi^2_q(\alpha) < \mathcal{L}_W && \text{or} \;\; \mathcal{L}_{LM} < \chi^2_q(\alpha) < \mathcal{L}_{LR},
\end{aligned} \tag{2.3}$$

respectively.

For the normal error case, Evans and Savin (1982) showed that on using the $\chi^2_q(\alpha)$ critical value there are two characteristics. First, they will differ with respect to their sizes and powers in small samples and there may be conflict between their conclusions. Second, when the sizes of the tests are corrected to be the same, the power are approximately the same and there may be no any confliction. For the Student's t error case, Ullah and Zinde-Walsh (1984) showed that the inequalities in (2.2) are complicated, and the relationship among the sizes of these test and the possibility of conflict is quite different from that of the normal case. For details we refer Ullah and Zinde-Walsh (1984).

For excellent references and for various researches on W, LR and LM tests, readers are referred to Savin (1976), Berndt and Savin (1977), Evans and Savin (1982), Ullah and Zinde-Walsh (1984), Billah and Saleh (2000) and most recently Kibria and Saleh (2003b) among others.

Based on the above considerations, we propose the following SPTRRE's based on W, LR and LM tests, which are given below,

$$\hat{\beta}_*^{SPT}(k) = \hat{\beta}^{SR}(k)I(\mathcal{L}_* \leq \chi_q^2(\alpha)) + \tilde{\beta}(k)I(\mathcal{L}_* > \chi_q^2(\alpha)), \tag{2.4}$$

where $*$ stands for either of W, LR and LM tests. Our objective is to compare the three estimators which have been developed based on the three test procedures. Therefore, for our convenience we consider the simpler critical value, $\chi_q^2(\alpha)$ in the formulation of the preliminary test estimators, which also induces the conflict in the inference procedures. In the following section, we provide the biases and risks of the proposed estimators.

3. BIASES AND RISKS

Let $\hat{\boldsymbol{\delta}}$ be an estimator of $\boldsymbol{\delta}$ and $\mathbf{W}$ be a positive semi-definite (psd) matrix, and consider the following quadratic loss function

$$L(\hat{\boldsymbol{\delta}}, \boldsymbol{\delta}) = n(\hat{\boldsymbol{\delta}} - \boldsymbol{\delta})'\mathbf{W}(\hat{\boldsymbol{\delta}} - \boldsymbol{\delta}) = n \operatorname{tr}\left\{\mathbf{W}(\hat{\boldsymbol{\delta}} - \boldsymbol{\delta})(\hat{\boldsymbol{\delta}} - \boldsymbol{\delta})'\right\},$$

where tr($\mathbf{A}$)=trace of the matrix $\mathbf{A}$. Then the risk of $\hat{\boldsymbol{\delta}}$ is given by

$$R(\hat{\boldsymbol{\delta}}, \boldsymbol{\delta}) = nE\left[\operatorname{tr}\left\{\mathbf{W}(\hat{\boldsymbol{\delta}} - \boldsymbol{\delta})(\hat{\boldsymbol{\delta}} - \boldsymbol{\delta})'\right\}\right].$$

The biases of the proposed estimators are

$$\mathbf{B}(\hat{\beta}_*^{SPT}(k)) = -\left\{(1-d)\mathbf{R}(k)\boldsymbol{\eta}G_{q+2,n-p}^{(2)}(l^*;\Delta) + k\mathbf{C}^{-1}(k)\beta\right\}, \tag{3.1}$$

where $\boldsymbol{\eta} = \mathbf{C}^{-1}\mathbf{H}'(\mathbf{H}\mathbf{C}^{-1}\mathbf{H}')^{-1}(\mathbf{H}\beta - \mathbf{h})$, $*$ stands for either of W, LR and LM tests, and l_* stands for any of the followings:

$$l^W = \frac{\chi_q^2(\alpha)}{n\lambda(n) + \chi_q^2(\alpha)}, \quad l^{LR} = 1 - e^{-\chi_q^2(\alpha)/n} \text{ and } l^{LM} = \frac{\lambda(n)\chi_q^2(\alpha)}{n}, \tag{3.2}$$

satisfying $l^W \le l^{LR} \le l^{LM}$. Further, we consider the following function

$$G^{(j)}_{q+2i,n-p}(l^*;\Delta) = \sum_{r=0}^{\infty} \frac{\Gamma(\nu/2+r+j-2)}{\Gamma(r+1)\Gamma(\nu/2+j-2)} \times \frac{\{\Delta/(\nu-2)\}^r}{\{1+\Delta/(\nu-2)\}^{\nu/2+r+j-2}} \times I_{l^*}\left\{\frac{1}{2}(q+2i)+r, \frac{m}{2}\right\}, \tag{3.3}$$

where $I\{\cdot\}$ is the incomplete beta function and l^* denotes any of the quantities, l^W, l^{LR} and l^{LM}. For $i=0$ and $j=2$, $G^{(2)}_{q,n-p}(l^*;\Delta)$ is the power function of the tests. Note that for $\alpha=1$, the biases and risks of the three estimators coincide with those of URRE. However, for $\alpha=0$, the biases and risks of the proposed estimators coincide with those of the RRRE. Since $l^{LM} \ge l^{LR} \ge l^W$ for all α, p and n, it follows that

$$G^{(j)}_{q+2i,n-p}(l^{LM};\Delta) \ge G^{(j)}_{q+2i,n-p}(l^{LR};\Delta) \ge G^{(j)}_{q+2i,n-p}(l^{W};\Delta);\ i,j=1,2. \tag{3.4}$$

The risk expressions for $\hat{\beta}^{SPT}_W(k)$, $\hat{\beta}^{SPT}_{LR}(k)$ and $\hat{\beta}^{SPT}_{LM}(k)$ are

$$\begin{aligned} &R(\hat{\beta}^{SPT}_*(k)) \\ &= \sigma_e^2 \mathrm{tr}\left\{\mathbf{R}(k)\mathbf{C}^{-1}\mathbf{R}(k)'\right\} - (1-d^2)\sigma_e^2 \mathrm{tr}\left\{\mathbf{R}(k)\mathbf{A}\mathbf{R}(k)'\right\} G^{(1)}_{q+2,n-p}(l^*;\Delta) \\ &\quad + (1-d)\boldsymbol{\eta}'\mathbf{R}(k)'\mathbf{R}(k)\boldsymbol{\eta}\left\{2G^{(2)}_{q+2,n-p}(l^*;\Delta) - (1+d)G^{(2)}_{q+4,n-p}(l^*;\Delta)\right\} \\ &\quad + 2(1-d)kG^{(2)}_{q+2,n-p}(l^*;\Delta)\boldsymbol{\eta}'\mathbf{R}(k)'\mathbf{C}^{-1}(k)\boldsymbol{\beta} + k^2\boldsymbol{\beta}'\mathbf{C}^{-2}(k)\boldsymbol{\beta}, \end{aligned} \tag{3.5}$$

where $\mathbf{A} = \mathbf{C}^{-1}\mathbf{H}'(\mathbf{H}\mathbf{C}^{-1}\mathbf{H}')^{-1}\mathbf{H}\mathbf{C}^{-1}$ and $*$ stands for either of W, LR and LM tests, l^* stands for any of the quantities in (3.2) and $G^{(i)}_{q+2i,n-p}(\cdot)$ $(i=1,2)$ are available from (3.3).

Now we assume that $\mathbf{P}$ be an orthogonal matrix so that

$$\mathbf{P}'\mathbf{C}\mathbf{P} = \boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_p),$$

where $\lambda_1, \lambda_2, \ldots, \lambda_p$ denote the eigenvalues of the matrix $\mathbf{C}$. Since $\mathbf{C}$ is symmetric, we can write

$$\mathbf{R}(k)\mathbf{A}\mathbf{R}(k)' = \mathbf{P}\left(\boldsymbol{\Lambda}+k\mathbf{I}_p\right)^{-1}\boldsymbol{\Lambda}\mathbf{A}^*\boldsymbol{\Lambda}\left(\boldsymbol{\Lambda}+k\mathbf{I}_p\right)^{-1}\mathbf{P}', \tag{3.6}$$

where $\mathbf{P}'\mathbf{A}\mathbf{P} = \mathbf{A}^*$. Now without loss of generality we assume that $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_p > 0$, and we can write,

$$\mathrm{tr}\left\{\mathbf{R}(k)\mathbf{C}^{-1}\mathbf{R}(k)'\right\} = \sum_{i=1}^{p} \frac{\lambda_i}{(\lambda_i+k)^2} \text{ and } \mathrm{tr}\left\{\mathbf{R}(k)\mathbf{A}\mathbf{R}(k)'\right\} = \sum_{i=1}^{p} \frac{\lambda_i^2 a^*_{ii}}{(\lambda_i+k)^2}, \tag{3.7}$$

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

$$\boldsymbol{\beta}'\mathbf{C}^{-2}(k)\boldsymbol{\beta} = \sum_{i=1}^{p} \frac{\beta_i^{*2}}{(\lambda_i + k)^2}, \tag{3.8}$$

where $\boldsymbol{\beta}^* = \mathbf{P}'\boldsymbol{\beta}$, and

$$\boldsymbol{\eta}'\mathbf{R}(k)'\mathbf{R}(k)\boldsymbol{\eta} = \sum_{i=1}^{p} \frac{\lambda_i^2 \eta_i^{*2}}{(\lambda_i + k)^2} \text{ and } \boldsymbol{\eta}'\mathbf{R}(k)'\mathbf{C}^{-1}(k)\boldsymbol{\beta} = \sum_{i=1}^{p} \frac{\lambda_i \eta_i{}^* \beta_i^*}{(\lambda_i + k)^2}, \tag{3.9}$$

where $\boldsymbol{\eta}^* = \mathbf{P}'\boldsymbol{\eta}$. Also, from Anderson (1984, Theorem A.2.4, p. 590) it follows that

$$\gamma_p \leq \frac{\boldsymbol{\eta}'\mathbf{R}(k)'\mathbf{R}(k)\boldsymbol{\eta}}{\boldsymbol{\eta}'\mathbf{C}\boldsymbol{\eta}} \leq \gamma_1, \tag{3.10}$$

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

Now we are ready to compare the performance of the proposed estimators in the following section.

4. PERFORMANCE OF THE PROPOSED ESTIMATORS

In this section we compare the performance of the proposed estimators based on quadratic bias and risk criterion. We assume that ν is known. First we compare the performance of the estimators based on the bias function.

4.1. Bias comparisons

In order to present a clear-cut picture of various bias functions, we transform them in scalar (quadratic) form by defining,

$$\text{QB}(\hat{\boldsymbol{\beta}}_*^{SPT}(k)) = \mathbf{B}(\hat{\boldsymbol{\beta}}_*^{SPT}(k))'\mathbf{B}(\hat{\boldsymbol{\beta}}_*^{SPT}(k)),$$

where $*$ stands for either of W, LR and LM tests. The quadratic biases for proposed estimators are

$$\begin{aligned}
&\text{QB}(\hat{\boldsymbol{\beta}}_*^{SPT}(k)) \\
&= (1-d)^2\boldsymbol{\eta}'\mathbf{R}(k)'\mathbf{R}(k)\boldsymbol{\eta}\left\{G_{q+2,n-p}^{(2)}(l^*;\Delta)\right\}^2 \\
&\quad + 2(1-d)k\boldsymbol{\eta}'\mathbf{R}(k)'\mathbf{C}^{-1}(k)\boldsymbol{\beta}G_{q+2,n-p}^{(2)}(l^*;\Delta) + k^2\boldsymbol{\beta}'\mathbf{C}^{-2}(k)\boldsymbol{\beta},
\end{aligned} \tag{4.1}$$

where l_* stands for either of l_W, l_{LR} and l_{LM}. If $\alpha = 1$, $\mathrm{QB}(\hat{\beta}_*^{SPT}(k))$ reduces to $\mathrm{QB}(\tilde{\beta}(k))$ and for $\alpha = 0$, $\mathrm{QB}(\hat{\beta}_*^{SPT}(k))$ reduces to $\mathrm{QB}(\hat{\beta}^{SR}(k))$.

We consider the difference, $\mathrm{QB}(\hat{\beta}_{LR}^{SPT}(k)) - \mathrm{QB}(\hat{\beta}_W^{SPT}(k))$, which is non-negative for any value of $\alpha \in (0,1)$, $d \in (0,1)$, Δ and k. Thus, $\hat{\beta}_W^{SPT}(k)$ has smaller quadratic bias than $\hat{\beta}_{LR}^{SPT}(k)$. Similarly, $\mathrm{QB}(\hat{\beta}_{LM}^{SPT}(k)) - \mathrm{QB}(\hat{\beta}_{LR}^{SPT}(k)) \geq 0$ and $\hat{\beta}_{LR}^{SPT}(k)$ has smaller quadratic bias than $\hat{\beta}_{LM}^{SPT}(k)$. Thus, we may order the quadratic bias functions as

$$\mathrm{QB}(\hat{\beta}_W^{SPT}(k)) \leq \mathrm{QB}(\hat{\beta}_{LR}^{SPT}(k)) \leq \mathrm{QB}(\hat{\beta}_{LM}^{SPT}(k)). \tag{4.2}$$

For $k = 0$ in (4.2), we obtain the order of the quadratic biases for the corresponding shrinkage preliminary test least squares estimators (SPTLSE) based on LM, LR and W tests. For $d = 0$ in (4.2), we obtain the order of the quadratic biases for the corresponding preliminary test ridge regression estimators (PTRRE) based on LM, LR and W tests. For $k = 0$ and $d = 0$ in (4.2), we obtain the order of the quadratic biases for the corresponding preliminary test least squares estimators (PTLSE) based on LM, LR and W tests.

4.2. Risk comparisons

We note from (3.5) that for given α and known data, the risks 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 provide them in the following two subsections.

4.2.1. Performance as a function of Δ. In this subsection, we compare the performance of the proposed estimators as a function of Δ. First, we compare between $\hat{\beta}_W^{SPT}(k)$ and $\hat{\beta}_{LR}^{SPT}(k)$. Using (3.5) the risk difference, $R(\hat{\beta}_W^{SPT}(k)) - R(\hat{\beta}_{LR}^{SPT}(k))$ is non-negative (≥ 0), whenever

$$\begin{aligned}\Delta &\leq \frac{\left[(1+d)\mathrm{tr}\left\{\mathbf{R}(k)\mathbf{A}\mathbf{R}(k)'\right\}\mathcal{A} - 2k\sigma_e^{-2}\boldsymbol{\eta}'\mathbf{R}(k)'\mathbf{C}^{-1}(k)\boldsymbol{\beta}\mathcal{B}\right]}{\gamma_1\left\{2\mathcal{B} - (1+d)\mathcal{E}\right\}} \\ &= \Delta_1(k,d,\alpha),\end{aligned} \tag{4.3}$$

where

$$\begin{aligned}\mathcal{A} &= G_{q+2,n-p}^{(1)}(l^{LR};\Delta) - G_{q+2,n-p}^{(1)}(l^{W};\Delta),\\ \mathcal{B} &= G_{q+2,n-p}^{(2)}(l^{LR};\Delta) - G_{q+2,n-p}^{(2)}(l^{W};\Delta),\\ \mathcal{E} &= G_{q+4,n-p}^{(2)}(l^{LR};\Delta) - G_{q+4,n-p}^{(2)}(l^{W};\Delta).\end{aligned}$$

Note from (3.4) that $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{E}$, are positive for all k, Δ and $\alpha \in (0,1)$.

Thus, $\hat{\beta}_{LR}^{SPT}(k)$ performs better than $\hat{\beta}_{W}^{SPT}(k)$, when (4.3) holds. Using equations (3.7) to (3.9), $\Delta_1(k,d,\alpha)$ can be expressed as

$$\Delta_1(k,d,\alpha) = \frac{1}{\gamma_1\left\{2\mathcal{B} - (1+d)\mathcal{E}\right\}} \sum_{i=1}^{p} \left\{ \frac{(1+d)\lambda_i^2 a_{ii}^* \mathcal{A} - 2k\sigma_e^{-2}\lambda_i \eta_i^* \beta_i^* \mathcal{B}}{(\lambda_i+k)^2} \right\}, \quad (4.4)$$

which depends on the eigenvalues, bias vectors $\boldsymbol{\eta}^*$ and $\boldsymbol{\beta}^*$. However, $\hat{\beta}_{W}^{SPT}(k)$ performs better than $\hat{\beta}_{LR}^{SPT}(k)$, whenever

$$\begin{aligned}\Delta &> \frac{1}{\gamma_p\left\{2\mathcal{B} - (1+d)\mathcal{E}\right\}} \sum_{i=1}^{p} \left\{ \frac{(1+d)\lambda_i^2 a_{ii}^* \mathcal{A} - 2k\sigma_e^{-2}\lambda_i \eta_i^* \beta_i^* \mathcal{B}}{(\lambda_i+k)^2} \right\} \\ &= \Delta_2(k,d,\alpha). \end{aligned} \quad (4.5)$$

Under the null hypothesis the difference, $R(\hat{\beta}_{W}^{SPT}(k)) - R(\hat{\beta}_{LR}^{SPT}(k))$ is always positive for all $\alpha \in (0,1)$ and $d \in (0,1)$, therefore, $\hat{\beta}_{LR}^{SPT}(k)$ is superior to $\hat{\beta}_{W}^{SPT}(k)$.

Now we compare between $\hat{\beta}_{LR}^{SPT}(k)$ and $\hat{\beta}_{LM}^{SPT}(k)$. Using (3.5) the risk difference, $R(\hat{\beta}_{LR}^{SPT}(k)) - R(\hat{\beta}_{LM}^{SPT}(k))$ is non-negative (≥ 0), whenever

$$\begin{aligned}\Delta &\leq \frac{1}{\gamma_1\left\{2\mathcal{B}^* - (1+d)\mathcal{E}^*\right\}} \sum_{i=1}^{p} \left\{ \frac{(1+d)\lambda_i^2 a_{ii}^* \mathcal{A}^* - 2k\sigma_e^{-2}\lambda_i \eta_i^* \beta_i^* \mathcal{B}^*}{(\lambda_i+k)^2} \right\} \\ &= \Delta_3(k,d,\alpha), \end{aligned} \quad (4.6)$$

where

$$\begin{aligned}\mathcal{A}^* &= G_{q+2,n-p}^{(1)}(l^{LM};\Delta) - G_{q+2,n-p}^{(1)}(l^{LR};\Delta), \\ \mathcal{B}^* &= G_{q+2,n-p}^{(2)}(l^{LM};\Delta) - G_{q+2,n-p}^{(2)}(l^{LR};\Delta), \\ \mathcal{E}^* &= G_{q+4,n-p}^{(2)}(l^{LM};\Delta) - G_{q+4,n-p}^{(2)}(l^{LR};\Delta).\end{aligned}$$

Note from (3.4) that $\mathcal{A}^*$, $\mathcal{B}^*$ and $\mathcal{E}^*$ are positive for all k, Δ and $\alpha \in (0,1)$.

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

$$\begin{aligned}\Delta &> \frac{1}{\gamma_p\left\{2\mathcal{B}^* - (1+d)\mathcal{E}^*\right\}} \sum_{i=1}^{p} \left\{ \frac{(1+d)\lambda_i^2 a_{ii}^* \mathcal{A}^* - 2k\sigma_e^{-2}\lambda_i \eta_i^* \beta_i^* \mathcal{B}^*}{(\lambda_i+k)^2} \right\} \\ &= \Delta_4(k,d,\alpha). \end{aligned} \quad (4.7)$$

Under the null hypothesis the difference, $R(\hat{\beta}_{LR}^{SPT}(k)) - R(\hat{\beta}_{LM}^{SPT}(k))$ is always positive for all $\alpha \in (0,1)$, therefore $\hat{\beta}_{LM}^{PT}(k)$ is superior to $\hat{\beta}_{LR}^{PT}(k)$. Now we describe the graph of $R(\hat{\beta}_*^{SPT}(k))$ as follows. At $\Delta = 0$, it assumes a value

$$\sigma_e^2 \mathrm{tr}\{\mathbf{R}(k)\mathbf{C}^{-1}\mathbf{R}(k)'\} - \sigma_e^2(1-d^2)\mathrm{tr}\{\mathbf{R}(k)\mathbf{A}\mathbf{R}(k)'\}G_{q+2,n-p}(l^*;0) + k^2\boldsymbol{\beta}'\mathbf{C}^{-2}(k)\boldsymbol{\beta},$$

then increases from 0, crossing the risk of URRE to a maximum and then drops gradually towards the risk of URRE as $\Delta \to \infty$.

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

THEOREM 4.1. *The dominance picture of the proposed estimators is*

$$R(\hat{\beta}_{LM}^{SPT}(k)) \leq R(\hat{\beta}_{LR}^{SPT}(k)) \leq R(\hat{\beta}_{W}^{SPT}(k)),$$

in the interval

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

where $\Delta_{13}^*(k,d,\alpha) = \min\{\Delta_1(k,d,\alpha), \Delta_3(k,d,\alpha)\}$, *also* $\Delta_1(k,d,\alpha)$ *and* $\Delta_3(k,d,\alpha)$ *are given in (4.4) and (4.6) respectively, while*

$$R(\hat{\beta}_{W}^{SPT}(k)) \leq R(\hat{\beta}_{LR}^{SPT}(k)) \leq R(\hat{\beta}_{LM}^{SPT}(k)),$$

in the interval

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

where $\Delta_{24}^*(k,d,\alpha) = \max\{\Delta_2(k,d,\alpha), \Delta_4(k,d,\alpha)\}$, *also* $\Delta_2(k,d,\alpha)$ *and* $\Delta_4(k,d,\alpha)$ *are given in (4.5) and (4.7) respectively. For* $k = 0$, *we obtain the corresponding dominance picture for the SPTLSE's. For* $d = 0$, *we obtain the corresponding dominance picture for the PTRRE's. For* $k = 0$ *and* $d = 0$, *we obtain the corresponding dominance picture for the PTLSE's.*

4.2.2. Risk comparison as a function of k. In this subsection, we compare the performance of the proposed estimators as a function of shrinkage parameter k. First we compare between $\hat{\beta}_{W}^{SPT}(k)$ and $\hat{\beta}_{LR}^{SPT}(k)$. Thus, using equations (3.7) to (3.9) the risk difference, $R(\hat{\beta}_{W}^{SPT}(k)) - R(\hat{\beta}_{LR}^{SPT}(k))$ will be non-negative (≥ 0) if

$$\begin{aligned} k &\leq \min_i \left[\sigma_e^2(1+d)a_{ii}^*\lambda_i\mathcal{A} - \{2\mathcal{B} - (1+d)\mathcal{E}\}\lambda_i\eta_i^{*2}\right]\left\{\max_i(2\eta_i^*\beta_i^*\mathcal{B})\right\}^{-1} \\ &= k_1(\alpha,d,\Delta), \end{aligned} \tag{4.8}$$

which depends on the eigenvalues, bias vectors $\boldsymbol{\eta}^*$ and $\boldsymbol{\beta}^*$. Thus, $\hat{\beta}_{LR}^{SPT}(k)$ will dominate $\hat{\beta}_{W}^{SPT}(k)$ if $0 \leq k \leq k_1(\alpha,d,\Delta)$, while $\hat{\beta}_{W}^{SPT}(k)$ will dominate $\hat{\beta}_{LR}^{SPT}(k)$ whenever

$$\begin{aligned} k &> \max_i \left[\sigma_e^2(1+d)a_{ii}^*\lambda_i\mathcal{A} - \{2\mathcal{B} - (1+d)\mathcal{E}\}\lambda_i\eta_i^{*2}\right]\left\{\min_i(2\eta_i^*\beta_i^*\mathcal{B})\right\}^{-1} \\ &= k_2(\alpha,d,\Delta). \end{aligned} \tag{4.9}$$

Now we compare between $\hat{\beta}_{LR}^{SPT}(k)$ and $\hat{\beta}_{LM}^{SPT}(k)$. Similarly, the risk difference, $R(\hat{\beta}_{LR}^{SPT}(k)) - R(\hat{\beta}_{LM}^{SPT}(k))$ will be non-negative (≥ 0) if

$$
\begin{aligned}
k &\leq \min_i \left[\sigma_e^2(1+d)a_{ii}^* \lambda_i \mathcal{A}^* - \{2\mathcal{B}^* - (1+d)\mathcal{E}^*\} \lambda_i \eta_i^{2*}\right] \left\{\max_i (2\eta_i^* \beta_i^* \mathcal{B}^*)\right\}^{-1} \\
&= k_3(\alpha, d, \Delta). \qquad (4.10)
\end{aligned}
$$

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

$$
\begin{aligned}
k &> \max_i \left[\sigma_e^2(1+d)a_{ii}^* \lambda_i \mathcal{A}^* - \{2\mathcal{B}^* - (1+d)\mathcal{E}^*\} \lambda_i {\eta_i^*}^2\right] \left\{\min_i (2\eta_i^* \beta_i^* \mathcal{B}^*)\right\}^{-1} \\
&= k_4(\alpha, d, \Delta). \qquad (4.11)
\end{aligned}
$$

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

THEOREM 4.2. *The dominance picture of the proposed estimators is*

$$R(\hat{\beta}_{LM}^{SPT}(k)) \leq R(\hat{\beta}_{LR}^{SPT}(k)) \leq R(\hat{\beta}_{W}^{SPT}(k)),$$

in the interval,

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

where $k_{13}^*(\alpha, d, \Delta) = \min\{k_1(\alpha, d, \Delta), k_3(\alpha, d, \Delta)\}$, *also* $k_1(\alpha, d, \Delta)$ *and* $k_3(\alpha, d, \Delta)$ *are given in (4.8) and (4.10) respectively, while*

$$R(\hat{\beta}_{W}^{SPT}(k)) \leq R(\hat{\beta}_{LR}^{SPT}(k)) \leq R(\hat{\beta}_{LM}^{SPT}(k)),$$

in the interval

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

where $k_{24}^*(\alpha, d, \Delta) = \max\{k_2(\alpha, d, \Delta), k_4(\alpha, d, \Delta)\}$, *also* $k_2(\alpha, d, \Delta)$ *and* $k_4(\alpha, d, \Delta)$ *are given in (4.9) and (4.11) respectively.*

Now we consider the conditions on Δ and k simultaneously and state the following theorem:

THEOREM 4.3. *The dominance picture of the proposed estimators is*

$$R(\hat{\beta}_{LM}^{SPT}(k)) \leq R(\hat{\beta}_{LR}^{SPT}(k)) \leq R(\hat{\beta}_{W}^{SPT}(k)),$$

in the interval,

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

while

$$R(\hat{\beta}_W^{SPT}(k)) \le R(\hat{\beta}_{LR}^{SPT}(k)) \le R(\hat{\beta}_{LM}^{SPT}(k)),$$

in the interval,

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

We have plotted the risk functions versus Δ for different n, p, q, α and k and presented them in Figures 1 and 2. From these figures we observed that for small values of Δ, the SPTRRE based on LM test has the smallest risk followed by the LR and W tests. However the situation becomes reverse when the restrictions are more inaccurate. Therefore, the graphical representation support the findings of the paper.

5. Relative Efficiency

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 SPTRRE for testing the null hypothesis (1.2). For a fixed value of k (> 0), the relative efficiency of the SPTRRE $(\hat{\beta}_*^{SPT}(k))$ compared to the URRE $(\tilde{\beta}(k))$ is a function of α and Δ. Let us denote this by

$$E(k, d, \alpha, \Delta) = \frac{R(\tilde{\beta}(k))}{R(\hat{\beta}_*^{SPT}(k))} = \{1 - h(k, d, \alpha, \Delta)\}^{-1}, \tag{5.1}$$

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

$$h(k, d, \alpha, \Delta) = \frac{g(k, d, \alpha, \Delta)}{\sum_{i=1}^p (\lambda_i + k)^{-2} \left(\lambda_i \sigma_e^2 + k^2 \beta_i^{*2}\right)},$$

and

$$g(k, d, \alpha, \Delta) = \sum_{i=1}^p (\lambda_i + k)^{-2} \Big[\sigma_e^2 (1 - d^2) \lambda_i^2 a_{ii}^* - (1 - d) \lambda_i^2 \eta_i^{*2} \Big\{2G_{q+2,n-p}^{(2)}(l^*; \Delta) - (1 + d) G_{q+4,n-p}^{(2)}(l^*; \Delta)\Big\} - 2(1 - d) k \lambda_i \eta_i^* \beta_i^* G_{q+2,n-p}^{(2)}(l^*; \Delta)\Big].$$

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

$$E_{\max}(k, d, \alpha, 0) = \left\{1 - \frac{\sigma_e^2 (1 - d^2) \sum_{i=1}^p (\lambda_i + k)^{-2} (\lambda_i^2 a_{ii}^*) G_{q+2,n-p}^{(1)}(l^*; 0)}{\sum_{i=1}^p (\lambda_i + k)^{-2} \left(\lambda_i \sigma_e^2 + k^2 \beta_i^{*2}\right)}\right\}^{-1}.$$

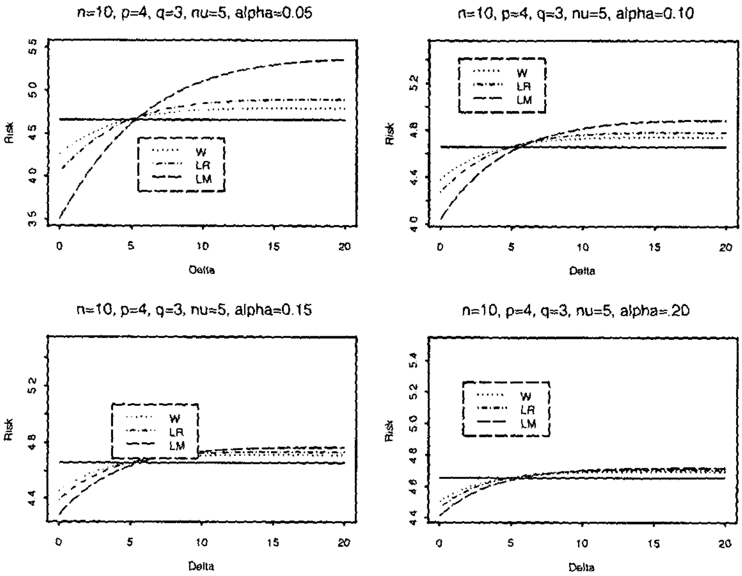

FIGURE 1 *Risk functions of the SPTRRE's based on the W, LR and LM tests for different significance levels and fixed k=0.2 and d=0.5*

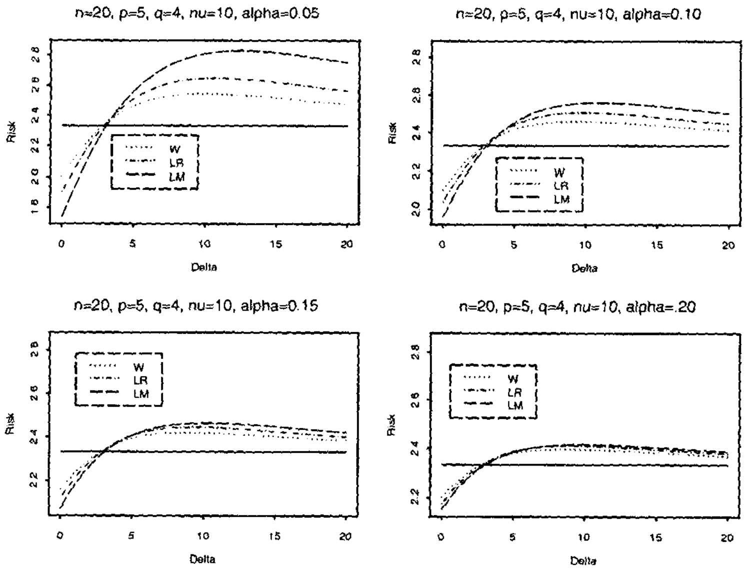

FIGURE 2 *Risk functions of the SPTRRE's based on the W, LR and LM tests for different significance levels and fixed k=0.7 and d=0.5*

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

$$\max_{0\le\alpha\le1} E(k,d,\alpha,0)=E(k,d,0,0)=\left\{1-\frac{\sigma_e^2(1-d^2)\sum_{i=1}^{p}(\lambda_i+k)^{-2}(\lambda_i^2 a_{ii}^*)}{\sum_{i=1}^{p}(\lambda_i+k)^{-2}\left(\lambda_i\sigma_e^2+k^2\beta_i^{*2}\right)}\right\}^{-1}$$

The value $E(k,d,\alpha,0)$ decreases as α increases. On the other hand, for $\alpha\neq0$, the graphs of Δ versus $E(k,d,0,\Delta)$ and Δ versus $E(k,d,1,\Delta)$ intersect in the range $0\le\Delta\le\Delta_1(k,d,0)$, where

$$\Delta_1(k,d,0) = \frac{1}{(1-d)\gamma_1}\left[\sum_{i=1}^{p}(\lambda_i+k)^{-2}\left\{(1+d)\lambda_i^2 a_{ii}^* - \sigma_e^{-2}2k\lambda_i\eta_i^*\beta_i^* G_{q+2,n-p}^{(2)}(l^*;\Delta)\right\}\right].$$

Thus in order to choose an estimator with optimum relative efficiency, we adopt the following rule for given k values. If $0<\Delta<\Delta_1(k,d,0)$, the SRRE would be chosen since $E(k,d,0,\Delta)$ is the largest in this interval. However, in general Δ is unknown and may not lie in this 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 $\mathcal{S}=\{\alpha\,|E(k,d,\alpha,\Delta)\ge E_{\min}\}$ and choose an estimator which maximizes $E(k,d,\alpha,\Delta)$ for all $\alpha\in\mathcal{S}$ and $\Delta\in[0,\infty)$. Thus we solve the following equation:

$$\max_{\alpha\in\mathcal{S}}\min_{\Delta} E(k,d,\alpha,\Delta) = E_{\min}. \tag{5.2}$$

The solution α^* for (5.2) gives the optimum choice of α and the value of $\Delta=\Delta_{\min}(k)$ for which (5.2) 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:

$$\hat{k}(\alpha,d,\Delta)=\min_i f(\alpha,\lambda_i,\eta_i)\left[\max_i\left\{2\eta_i^*\beta_i^* G_{q+2,n-p}^{(2)}(l^*;\Delta)\right\}\right]^{-1},$$

where

$$f(\alpha,\lambda_i,\eta_i) = \left[\sigma_e^2(1+d)a_{ii}^*\lambda_i G_{q+2,n-p}^{(1)}(l^*;\Delta) - \lambda_i\eta_i^{2*}\left\{2G_{q+2,n-p}^{(2)}(l^*;\Delta) - (1+d)G_{q+4,n-p}^{(2)}(l^*;\Delta)\right\}\right].$$

The above equation is obtained from the difference of $R(\tilde{\beta}(k))$ and $R(\hat{\beta}_*^{SPT}(k))$ and based on the smaller risk criterion. Since

$$2G^{(2)}_{q+2,n-p}(l^*;\Delta) - (1+d)G^{(2)}_{q+4,n-p}(l^{**};\Delta) \\ \leq 2G^{(2)}_{q+2,n-p}(l^*;\Delta) - G^{(2)}_{q+4,n-p}(l^{**};\Delta),$$

we obtain that the range of k for which the risk of the SPTRRE is less than that of the URRE, is therefore greater than the range of k for which the risk of the PTRRE is less than that of the URRE. Hence the proposed SPTRRE has wider range than the usual preliminary test ridge regression estimator (PTRRE) in which it dominates the URRE, for any value of ν, d and α.

Table 1 provides the values of the maximum and minimum guaranteed relative efficiencies and recommended corresponding size of α of the proposed estimators for $p = 4$, $q = 3$, $d = 0.5$ and $n = 10(5)30$, $\nu = 5, 10$ and $k = 0.20$. Table 2 gives the values of the maximum and minimum guaranteed relative efficiencies and recommended corresponding size of α of the PTRRE for $p = 4$, $q = 3$ and $n = 10(5)30$, $\nu = 5, 10$ and $k = 0.20$. From Tables 1 and 2, it is observed that under the restrictions $\Delta = 0$, the following inequalities (in the sense of higher guaranteed maximum efficiency) hold for any n, p, k and ν

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

which support the findings of the paper. It is also noted that the gain of efficiencies are higher for all proposed estimators for small values of k (related tables have been deleted to save the space of the paper). From Tables 1 and 2 we also note that the minimum guaranteed efficiencies of SPTRRE's based on W, LR and LM tests are higher than those of PTRRE's for all n, p, k and ν.

How can one use the table? For instance, if $n = 15$, $p = 4$, $q = 3$, $\nu = 10$, $k = 0.20$, and the experimenter believes that $d = 0.5$ and wishes to have an estimator with a minimum guaranteed efficiency of at least 0.95. Now using Table 1, we recommend him/her to select $\alpha = 0.025$, corresponding to $\hat{\beta}^{SPT}_{W}(k)$, because such a choice of α would yield an estimator with a minimum guaranteed efficiency of 0.95460 and maximum efficiency 1.20588. Note that the size of α corresponding to the minimum guaranteed efficiency of 0.95 for $\hat{\beta}^{SPT}_{LR}(k)$ and $\hat{\beta}^{SPT}_{LM}(k)$ are 0.05 and 0.10 respectively. Thus we choose $\alpha^* = \min(0.025, 0.05, 0.10) = 0.025$, which corresponds to Wald test. If we look at the maximum efficiency vector, (1.20588, 1.20462, 1.21668), we choose either Wald or LM based estimator. However, in the sense of smaller significant level, we choose Wald test. Again, if $n = 15$, $p = 4$, $q =$

TABLE 1 *Max & Min guaranteed efficiency of SPTRRE's (p=4, k=0.2, d=0.5)*

Test		*Efficiency for the following values of* α							
		2.5%	5%	10%	15%	2.5%	5%	10%	15%
		$n=10$ $\nu=5$				$n=10$ $\nu=10$			
W	$E_{\max}$	1.15798	1.11513	1.08773	1.06797	1.15098	1.10886	1.08231	1.06335
	$E_{\min}$	0.97853	0.98645	0.99046	0.99295	0.96636	0.97862	0.98489	0.98882
	$\Delta_{\min}$	21.54605	21.05263	20.72368	20.55921	13.32237	12.66447	12.17105	11.84211
LR	$E_{\max}$	1.24080	1.17828	1.13369	1.10076	1.24021	1.17786	1.13339	1.10054
	$E_{\min}$	0.95219	0.97318	0.98285	0.98831	0.92418	0.95656	0.97188	0.98069
	$\Delta_{\min}$	22.03947	21.38158	20.88816	20.55921	14.80263	13.81579	12.99342	12.50000
LM	$E_{\max}$	1.32889	1.29032	1.22633	1.16722	1.32908	1.29967	1.24052	1.18125
	$E_{\min}$	0.86505	0.93199	0.96459	0.97992	0.78590	0.88073	0.93537	0.96280
	$\Delta_{\min}$	26.71053	25.32895	24.63816	24.17763	18.65132	16.80921	15.42763	14.50658
		$n=15$ $\nu=5$				$n=15$ $\nu=10$			
W	$E_{\max}$	1.21099	1.15988	1.12429	1.09745	1.20588	1.15481	1.11968	1.09340
	$E_{\min}$	0.97198	0.98241	0.98764	0.99088	0.95460	0.97113	0.97959	0.98491
	$\Delta_{\min}$	21.05263	20.55921	20.23026	20.06579	13.15789	12.33553	11.84211	11.51316
LR	$E_{\max}$	1.26450	1.20511	1.15865	1.12239	1.26383	1.20462	1.15828	1.12212
	$E_{\min}$	0.95461	0.97412	0.98310	0.98824	0.92552	0.95647	0.97119	0.97982
	$\Delta_{\min}$	22.03947	21.38158	21.05263	20.72368	14.30921	13.15789	12.50000	12.00658
LM	$E_{\max}$	1.31711	1.26659	1.20942	1.15953	1.31874	1.27232	1.21668	1.16662
	$E_{\min}$	0.91767	0.95987	0.97710	0.98567	0.86539	0.93138	0.96017	0.97494
	$\Delta_{\min}$	25.32895	24.40789	23.94737	23.71711	16.57895	14.96711	14.04605	13.35526
		$n=20$ $\nu=5$				$n=20$ $\nu=10$			
W	$E_{\max}$	1.23574	1.18264	1.14357	1.11328	1.23205	1.17877	1.13996	1.11006
	$E_{\min}$	0.96917	0.98071	0.98647	0.99003	0.94925	0.96778	0.97726	0.98320
	$\Delta_{\min}$	21.05263	20.55921	20.23026	19.90132	12.99342	12.33553	11.84211	11.51316
LR	$E_{\max}$	1.27353	1.21646	1.16981	1.13243	1.27284	1.21594	1.16942	1.13213
	$E_{\min}$	0.95635	0.97476	0.98328	0.98822	0.92777	0.95722	0.97133	0.97970
	$\Delta_{\min}$	21.87500	21.38158	20.88816	20.55921	13.98026	12.82895	12.33553	11.84211
LM	$E_{\max}$	1.30994	1.25791	1.20405	1.15756	1.31117	1.26149	1.20846	1.16191
	$E_{\min}$	0.93736	0.96824	0.98091	0.98751	0.89756	0.94645	0.96748	0.97868
	$\Delta_{\min}$	24.86842	24.17763	23.94737	23.48684	15.65789	14.27632	13.35526	12.89474
		$n=30$ $\nu=5$				$n=30$ $\nu=10$			
W	$E_{\max}$	1.25805	1.20468	1.16281	1.12928	1.25579	1.20222	1.16046	1.12717
	$E_{\min}$	0.96664	0.97920	0.98544	0.98928	0.94416	0.96463	0.97507	0.98162
	$\Delta_{\min}$	21.05263	20.55921	20.23026	19.90132	12.99342	12.17105	11.67763	11.34868
LR	$E_{\max}$	1.28106	1.22655	1.18010	1.14191	1.28035	1.22600	1.17968	1.14159
	$E_{\min}$	0.95862	0.97563	0.98360	0.98830	0.93083	0.95836	0.97171	0.97974
	$\Delta_{\min}$	21.71053	21.21711	20.72368	20.39474	13.48684	12.66447	12.00658	11.67763
LM	$E_{\max}$	1.30316	1.25118	1.20039	1.15678	1.30361	1.25277	1.20242	1.15883
	$E_{\min}$	0.95195	0.97420	0.98368	0.98889	0.92151	0.95702	0.97269	0.98143
	$\Delta_{\min}$	24.40789	23.94737	23.71711	23.25658	14.73684	13.58553	12.89474	12.43421

TABLE 2 *Max & Min guaranteed efficiency of PTRRE's (p=4, q=3, k=0.2)*

Test		*Efficiency for the following values of* α							
		2.5%	5%	10%	15%	2.5%	5%	10%	15%
		$n = 10$	$\nu = 5$			$n = 10$	$\nu = 10$		
W	$E_{\max}$	1.6877	1.4312	1.2982	1.2146	1.6458	1.4025	1.2772	1.1987
	$E_{\min}$	0.9012	0.9375	0.9563	0.9682	0.8830	0.9259	0.9483	0.9625
	$\Delta_{\min}$	11.6776	11.1842	10.8553	10.5263	8.7171	8.0592	7.7303	7.4013
LR	$E_{\max}$	2.1649	1.6615	1.4223	1.2845	2.1599	1.6593	1.4212	1.2838
	$E_{\min}$	0.8358	0.9071	0.9403	0.9594	0.7983	0.8840	0.9251	0.9490
	$\Delta_{\min}$	12.8290	11.8421	11.3487	11.0197	9.8684	8.7171	8.2237	7.7303
LM	$E_{\max}$	3.3802	2.2374	1.6848	1.4115	3.4990	2.3661	1.7626	1.4590
	$E_{\min}$	0.6444	0.8443	0.9168	0.9505	0.5463	0.7910	0.8877	0.9337
	$\Delta_{\min}$	16.8092	14.2763	13.1250	12.6645	14.2763	10.5921	9.4408	8.7500
		$n = 15$	$\nu = 5$			$n = 15$	$\nu = 10$		
W	$E_{\max}$	2.0599	1.6600	1.4516	1.3219	2.0179	1.6311	1.4306	1.3063
	$E_{\min}$	0.8773	0.9219	0.9453	0.9601	0.8518	0.9052	0.9336	0.9517
	$\Delta_{\min}$	11.5132	11.0197	10.6908	10.5263	8.5526	8.0592	7.5658	7.4013
LR	$E_{\max}$	2.4620	1.8588	1.5558	1.3778	2.4549	1.8556	1.5541	1.3768
	$E_{\min}$	0.8343	0.9022	0.9353	0.9550	0.7958	0.8778	0.9188	0.9435
	$\Delta_{\min}$	12.3355	11.5132	11.0197	10.8553	9.3750	8.3882	7.8947	7.5658
LM	$E_{\max}$	3.1280	2.1948	1.7161	1.4558	3.1913	2.2565	1.7577	1.4834
	$E_{\min}$	0.7682	0.8824	0.9298	0.9548	0.7100	0.8505	0.9109	0.9430
	$\Delta_{\min}$	14.5066	13.1250	12.6645	12.2040	11.0526	9.4408	8.5197	8.0592
		$n = 20$	$\nu = 5$			$n = 20$	$\nu = 10$		
W	$E_{\max}$	2.2788	1.7966	1.5419	1.3840	2.2427	1.7715	1.5239	1.3708
	$E_{\min}$	0.8663	0.9146	0.9401	0.9564	0.8369	0.8952	0.9265	0.9465
	$\Delta_{\min}$	11.5132	11.0197	10.6908	10.3618	8.5526	7.8947	7.5658	7.4013
LR	$E_{\max}$	2.6015	1.9593	1.6256	1.4272	2.5933	1.9555	1.6236	1.4260
	$E_{\min}$	0.8356	0.9009	0.9335	0.9533	0.7971	0.8763	0.9167	0.9415
	$\Delta_{\min}$	12.0066	11.3487	11.0197	10.6908	9.0461	8.2237	7.7303	7.5658
LM	$E_{\max}$	3.0517	2.1927	1.7379	1.4810	3.0877	2.2294	1.7640	1.4990
	$E_{\min}$	0.8049	0.8945	0.9342	0.9563	0.7593	0.8690	0.9187	0.9463
	$\Delta_{\min}$	13.8158	12.8947	12.4342	12.204	10.1316	8.9803	8.2895	8.0592
		$n = 30$	$\nu = 5$			$n = 30$	$\nu = 10$		
W	$E_{\max}$	2.5112	1.9455	1.6397	1.4503	2.4848	1.9272	1.6268	1.4410
	$E_{\min}$	0.8564	0.9080	0.9355	0.9530	0.8230	0.8859	0.9199	0.9417
	$\Delta_{\min}$	11.5132	10.8553	10.5263	10.3618	8.5526	7.8947	7.5658	7.2368
LR	$E_{\max}$	2.7328	2.0597	1.6969	1.4782	2.7234	2.0554	1.6946	1.4768
	$E_{\min}$	0.8380	0.9004	0.9322	0.9519	0.7996	0.8756	0.9152	0.9397
	$\Delta_{\min}$	11.8421	11.1842	10.8553	10.6908	8.8816	8.0592	7.7303	7.4013
LM	$E_{\max}$	3.0010	2.2018	1.7649	1.5097	3.0138	2.2181	1.7774	1.5187
	$E_{\min}$	0.8314	0.9037	0.9377	0.9574	0.7950	0.8829	0.9247	0.9489
	$\Delta_{\min}$	13.3553	12.6645	12.2040	11.9737	9.4408	8.5197	8.0592	7.8290

3, $\nu = 10$, $k = 0.20$, and the experimenter believes that $d = 0$ and wishes to have an estimator with a minimum guaranteed efficiency of at least 0.90. Now using Table 2, we recommend him/her to select $\alpha = 0.05$, corresponding to $\hat{\beta}_W^{SPT}(k)$, because such a choice of α would yield an estimator with a minimum guaranteed efficiency of 0.9052 and maximum efficiency 1.6311. The size of α corresponding to the minimum guaranteed efficiency of 0.90 for $\hat{\beta}_{LR}^{SPT}(k)$ and $\hat{\beta}_{LM}^{SPT}(k)$ are 0.10 and 0.10 respectively. Thus we choose $\alpha^* = \min(0.05, 0.10, 0.10) = 0.05$, which is corresponds to Wald test. If we look at the maximum efficiency vector, (1.6311, 1.5541, 1.7577), we choose either Wald or LM based estimators. But, in the sense of smaller significant level, we choose Wald test. Therefore, using both SPTRRE and PTRRE, we obtain the smallest significance level corresponding to Wald test in the sense having highest minimum guaranteed efficiency. We note that the minimum guaranteed efficiencies of SPTRRE are always larger then the corresponding PTRRE. Therefore, from the application point of view, it is advisable to use SPTRRE based on Wald test among three test procedures.

6. Concluding Remarks

In this paper we studied the effect of W, LR and LM tests on the performance of the SPTRRE for the regression parameters when there exists a uncertain prior information in the parameter space. In literature, it is known that these test statistics satisfy the inequalities in (2.2). Thus there may exist conflict in the resulting test conclusions when certain fixed critical value is chosen. We have effectively determined some conditions on the departure parameter and the ridge parameter for the superiority of the proposed estimators. Note that the superiority of the proposed estimators depends on data and the information about the hypothesis. We have also discussed the method of choosing optimum level of significance to obtain minimum guaranteed efficient estimator. Under the restriction, the SPTRRE based on LM test has the smallest risk followed by the estimators based on LR and W tests. However, the SPTRRE 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 SPTRRE based on W test is found to perform the best with the choice of the smallest level of significance to yield the best estimator in the sense of the highest minimum guaranteed efficiency. We note that the SPTRRE provides the higher minimum guaranteed efficiency compared to PTRRE. The most significant feature of the results of our example is that the optimum choice of the level of significance becomes the traditional choice by the W test. The results of this paper also hold for SPTLSE, PTRRE and PTLSE based on W, LR and LM tests. We may expect that the analysis which have been done in this paper will be useful when the underlying distribution is normal for large ν and Cauchy for $\nu = 1$. Finally, based on the findings of this paper, we recommend the practitioner to use Wald test among three test procedures, when they consider the SPTRRE for estimating the regression parameters $\boldsymbol{\beta}$.

ACKNOWLEDGEMENTS

The author would like to thank two referees for their careful readings and useful comments which have improved the presentation of the paper greatly. He is also thankful to the Dean of College of Arts and Sciences of Florida International University for providing him some research facilities.

REFERENCES

AHMED, S. E. (2002). "Simultaneous estimation of coefficients of variation", *Journal of Statistical Planning and Inference*, **104**, 31–51.

ALDRIN, M. (1997). "Length modified ridge regression", *Computational Statistics and Data Analysis*, **25**, 377–398.

ANDERSON, T. W. (1984). *An Introduction to Multivariate Statistical Analysis*, 2nd ed., John Wiley & Sons, New York.

BANCROFT, T. A. (1944). "On biases in estimation due to the use of preliminary tests of significance", *The Annals of Mathematical 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. R. AND SAVIN, N. E. (1977). "Conflict among criteria for testing hypotheses in the multivariate linear regression model", *Econometrica*, **45**, 1263–1277.

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", In *Advances on Methodological and Applied Aspects of Probability and Statistics* (N. Balakrishnan, ed.), Gordon and Breach Science Publishers.

BLATTBERG, R. C. AND GONEDES, N. J. (1974). "A comparison of the stable and Student distributions as statistical models for stock prices", *Journal of Business*, **47**, 244–280.

EVANS, G. B. A. AND SAVIN, N. E. (1982). "Conflict among the criteria revisited; The W, LR and LM tests", *Econometrica*, **50**, 737–748.

FOUCART, T. (1999). "Stability of the inverse correlation matrix. Partial ridge regression", *Journal of Statistical Planning and Inference*, **77**, 141–154.

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

GILES, J. A. (1991). "Pre-testing for linear restrictions in a regression model with spherically symmetric disturbances", *Journal of Econometrics*, **50**, 377–398.

GNANADESIKAN, R. (1977). *Methods for Statistical Data Analysis of Multivariate Observations*, John Wiley & Sons, New York.

GUNST, R. (2000). "Classical studies that revolutionized the practice of regression analysis", *Technometrics*, **42**, 62–64.

HAN, C-P. AND BANCROFT, T. A. (1968). "On pooling means when variance is unknown", *Journal of the American Statistical Association*, **63**, 1333–1342.

HAQ, M. S. AND KIBRIA, B. M. G. (1996). "A shrinkage estimator for the restricted linear regression model : Ridge regression approach", *Journal of Applied Statistical Science*, **3**, 301–316.

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.

KELKER, D. (1970). "Distribution theory of spherical distributions and a location-scale parameter generalization", *Sankhyā*, **A32**, 419–438.

KENNEDY, P. (1998). *A Guide to Econometrics*, 4th ed., The MIT Press, Cambridge, Massachusetts.

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

KIBRIA, B. M. G. AND AHMED, S. E. (1997). "Shrinkage estimation for the multicollinear observations in a regression model with multivariate t disturbances", *Journal of Statistical Research*, **31**, 83–102.

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. (2003a). "Estimation of the mean vector of a multivariate normal distribution under various test statistics", *Journal of Probability and Statistical Science*, **1**, 141–155.

KIBRIA, B. M. G. AND SALEH, A. K. MD. E. (2003b). "Effect of W, LR, and LM tests on the performance of preliminary test ridge regression estimators", *Journal of the Japan Statistical Society*, **33**, 119–136.

KING, M. L. (1980). "Robust tests for spherical symmetry and their application to least squares regression", *The Annals of Statistics*, **8**, 1265–1271.

SALEH, A. K. MD. E. AND HAN, C-P. (1990). "Shrinkage estimation in regression analysis", *Estadistica*, **42**, 40–63.

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 methods of estimation", *Communications in Statistics-Theory and Methods*, **21**, 1987–2000.

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

SEN, P. K. AND SALEH, A. K. MD. E. (1987). "On preliminary test and shrinkage M-estimation in linear models", *The Annals of Statistics*, **15**, 1580–1592.

ULLAH, A. AND ZINDE-WALSH, V. (1984). "On the robustness of LM, LR and W tests in regression models", *Econometrica*, **52**, 1055–1066.

ZELLNER, A. (1976). "Bayesian and non-Bayesian analysis of the regression model with multivariate Student-t error terms", *Journal of the American Statistical Association*, **71**, 400–405.