ISSN 1683-5603

International Journal of Statistical Sciences Vol. 3, 2004, pp 1–20 © 2004 Dept. of Statistics, Univ. of Rajshahi, Bangladesh

Simultaneous Estimation of Regression Parameters With Spherically Symmetric Errors Under Possible Stochastic Constraints

S. M. Mehdi Tabatabaey

Department of Statistics School of Mathematical Sciences Ferdowsi University of Mashhad , Mashhad, Iran Email:tabatab@math.um.ac.ir

A. K. Md. E. Saleh

School of Mathematics and Statistics Carleton University, Ottawa, Canada K1S 5B6 Email: esaleh@math.carleton.ca

B. M. Golam Kibria

Department of Statistics Florida International University, Miami, FL 33199, USA Email: kibriag@fiu.edu

[Received January 24, 2003; Revised March 2, 2004; Accepted May 4, 2004]

Abstract

The problem of pre-test and shrinkage estimation of the parameters for the linear regression model with spherically symmetric errors under possible stochastic constraints are discussed in this paper. We consider five well known estimators, namely, unrestricted estimator (UE), restricted estimator (RE), preliminary test (PT) estimator, shrinkage estimator (SE) and positive rule (PR) shrinkage estimator. The bias and risk functions of the proposed estimators are analyzed under both the null and alternative hypotheses. Under the null hypothesis, the restricted estimator (RE) has the smallest risk followed by the pre-test or shrinkage estimators. However, the pre-test or shrinkage estimator (RE) when the parameter moves away from the subspace of the restrictions. The conditions of superiority of the proposed estimator for departure parameter are provided. It is evident that the positive rule shrinkage estimator utilizes both sample and non-sample information and performs uniformly better than UE and ordinary shrinkage estimator.

Keywords and Phrases: Bias; James and Stein Estimator; Preliminary Test; Risk; Restricted Estimators; Spherical Distribution; Students *t*; Stochastic Constraints; Superiority.

AMS Classification: 62J07; 62F03.

1 Introduction

The pre-test or shrinkage estimation under the general linear hypothesis (exact or nonstochastic) are available in literature. In rare cases we have exact prior information on the linear combination of parameters when estimating economic relations. Some uncertainties about the prior information are stochastic for many practical situation (see Theil and Goldberger, 1961 and Theil 1963). In that case non-stochastic constraint does not work. In most applied as well as theoretical research works, 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 and Zellner 1976). It happens particularly if the error distribution has heavier tails. For instance, some economic data may be generated by processes whose distribution have more kurtosis than the normal distribution. 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. Fama (1965), Blatberg and Gonedes 1974, Ullah and Zinde-Walsh 1984, Sutradhar and Ali, and Kibria 1996). Furthermore, normal distribution can not handle the dependent but uncorrelated responses which are often common in time series and econometric studies. The multivariate Student t distribution can overcome both the problems of outliers and dependent but uncorrelated data. Moreover, the multivariate normal distribution is a special case of multivariate Student t distribution for large error degrees of freedoms. Also the Cauchy distribution is a special case of multivariate Student t distribution for one degrees of freedom. In this paper we consider five estimators, namely, the unrestricted estimator, the restricted estimator, the preliminary test estimator, the shrinkage estimator and positive rule shrinkage estimator under the possible stochastic prior information. We also consider that errors follow a spherically symmetric distribution and provide the bias and the risk functions of the estimators. We compared the performance of the estimators based on the quadratic risk functions.

The organization of the paper is as follows. In Section 2, we discuss the model and the estimators. Section 3 contains the expressions of biases and risks. Section 4 deals with the relative performance of the estimators. Finally some concluding remarks are added in Section 5.

2 The Model and the Proposed Estimators

Consider the following linear regression model

$$y_1 = X\beta + e,\tag{1}$$

where y_1 is an $n \times 1$ vector of observations on the dependent variable, X is an $n \times p$ matrix of full rank $p, \beta = (\beta_1, \beta_2, \dots, \beta_p)'$ is an $p \times 1$ vector of regression parameters and e is an $n \times 1$ vector of errors, which is distributed as multivariate Student-t with the following probability density function (pdf)

$$f_1(e) = \frac{\Gamma\left(\frac{n+\nu}{2}\right)}{(\pi\nu)^{n/2}\Gamma(\nu/2)\sigma^n} \left(1 + \frac{e'e}{\nu\sigma^2}\right)^{-\frac{n+\nu}{2}}, \quad 0 < \nu, \ \sigma, < \infty, \quad -\infty < e_i < \infty.$$
(2)

Our primary object is to estimate β when the error distribution belongs to (2) and it is suspected that β may be restricted to the subspace defined by

$$r = R\beta + v, \tag{3}$$

where r is an $q \times 1$ vector of observations, R is an $q \times p$ matrix of known constants of full rank q, and v is an $q \times 1$ vector of errors, which is distributed according to the laws belonging to the class of spherical compound normal distributions. 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,

$$f_2(v) = \int_0^\infty f_q(v|\gamma)g(\gamma)d\gamma,\tag{4}$$

where $f_2(v)$ is the *pdf* of v, $f_q(v|\gamma)$ is the normal *pdf* with mean vector ψ and variancecovariance matrix $\gamma^2 \Omega$ ($\gamma > 0$) and $g(\gamma)$ is the inverted Gamma density with scale σ^2 and degrees of freedom ν , denoted by $IG(\nu, \sigma^2)$. Thus

$$f_2(v) = \frac{\Gamma\left(\frac{q+\nu}{2}\right)|\Omega|^{-\frac{1}{2}}}{(\pi\nu)^{q/2}\Gamma(\nu/2)\sigma^q} \left(1 + \frac{(v-\psi)'\Omega^{-1}(v-\psi)'}{\nu\sigma^2}\right)^{-\frac{q+\nu}{2}} , \quad 0 < \nu, \ \sigma, < \infty, \ (5)$$
$$-\infty < v_i < \infty.$$

with

$$E(v) = \psi$$
 and $E(vv') = \sigma_e^2 \Omega + \psi \psi', \ \sigma_e^2 = \frac{\nu}{\nu - 2} \sigma^2$

We assume that $v|\gamma$ and $e|\gamma$ are independent. We combine the sample and stochastic prior information to get the following linear statistical model

$$\begin{bmatrix} y_1 \\ r \end{bmatrix} = \begin{bmatrix} X \\ R \end{bmatrix} \beta + \begin{bmatrix} e \\ v \end{bmatrix}, \tag{6}$$

where

$$\begin{bmatrix} e \\ v \end{bmatrix} | \gamma \sim N_{n+q} \left\{ \begin{bmatrix} 0 \\ \psi \end{bmatrix}, \gamma^2 \begin{bmatrix} I_n & 0 \\ 0 & \Omega \end{bmatrix} \right\},$$
(7)

subject to condition

$$\left[-R, I_q\right] \left[\begin{array}{c} \beta\\ R\beta + \psi \end{array}\right] = \psi = 0. \tag{8}$$

Rewrite the model as

$$y = Z\phi + u,\tag{9}$$

subject to exact restriction

$$H\phi = \psi = 0, \tag{10}$$

where

$$y = \begin{bmatrix} y_1\\ \Omega^{-\frac{1}{2}}r \end{bmatrix}, \ Z = \begin{bmatrix} X & 0\\ 0 & \Omega^{-\frac{1}{2}} \end{bmatrix}, \ \phi = \begin{bmatrix} \beta\\ R\beta + \psi \end{bmatrix}, \ H = [-R, I_q]$$
(11)

and

$$u = \begin{bmatrix} e \\ \Omega^{-\frac{1}{2}}(v-\psi) \end{bmatrix} | \gamma \sim N_{n+q}(0, \gamma^2 I_{n+q}).$$
(12)

Then using (4), we have the pdf of u as,

$$f(u) = \frac{\Gamma\left(\frac{n+q+\nu}{2}\right)}{(\pi\nu)^{\frac{n+q+\nu}{2}}\Gamma(\nu/2)\sigma^{(n+q)}} \left(1 + \frac{u'u}{\nu\sigma^2}\right)^{-\frac{n+q+\nu}{2}} , \quad 0 < \nu, \ \sigma, < \infty,$$
(13)
$$-\infty < u_i < \infty.$$

For the full model the unrestricted least squares estimator (UE) of ϕ is given by

$$\hat{\phi}^{UE} = (Z'Z)^{-1}Z'y = \begin{bmatrix} \hat{\phi}_1^{UE} \\ \hat{\phi}_2^{UE} \end{bmatrix} = \begin{bmatrix} \hat{\beta}^{UE} \\ r \end{bmatrix},$$
(14)

where $\hat{\beta}^{UE} = (X'X)^{-1}X'y$ is the unrestricted estimator of β . The restricted least squares estimator (RE) of ϕ is given by

$$\hat{\phi}^{RE} = \hat{\phi}^{UE} - (Z'Z)^{-1}H'[H(Z'Z)^{-1}H']^{-1}\hat{\phi}^{UE} = \begin{bmatrix} \hat{\phi}_1^{RE} \\ \hat{\phi}_2^{RE} \end{bmatrix} = \begin{bmatrix} \hat{\beta}^{RE} \\ R\hat{\beta}^{RE} \end{bmatrix}, \quad (15)$$

where $\hat{\beta}^{RE}$ is the stochastic hypothesis restricted estimator of β and is given by

$$\hat{\beta}^{RE} = \hat{\beta}^{UE} - S^{-1}R'(RS^{-1}R' + \Omega)^{-1}(R\hat{\beta}^{UE} - r),$$

where S = X'X is the information matrix.

The estimator of ϕ in (14) is usually used in the case when there is no hypothesis information available on the vector of parameter of interest ϕ . On the other hand, the estimator of ϕ in (15) is useful in the presence of null hypothesis H_0 (10). Therefore, if we have hypothesis information on the parameter space, it is advantageous to use this additive information in the hope of obtaining better estimator (Ahmed 2002). Furthermore, it is well known that the RE performs better than the UE, when the restrictions hold but as the parameters, ϕ moves away from the subspace $H\phi = 0$, the RE becomes biased and inefficient while the performance of the UE remains stable. As a result, one may combine the UE and RE to obtain a better performance of the estimators in presence of the uncertain prior information $H\phi = 0$, which leads to the preliminary test least squares estimator (PT) and defined as

$$\hat{\phi}^{PT} = \hat{\phi}^{RE} I(\mathcal{L}_n \leq \mathcal{L}_{n,\alpha}) + \hat{\phi}^{UE} I(\mathcal{L}_n > \mathcal{L}_{n,\alpha}),$$

$$= \begin{pmatrix} \hat{\beta}^{PT} \\ R\hat{\beta}^{RE} I(\mathcal{L}_n \leq \mathcal{L}_{n,\alpha}) + rI(\mathcal{L}_n > \mathcal{L}_{n,\alpha}) \end{pmatrix}$$
(16)

where,

$$\hat{\beta}^{PT} = \hat{\beta}^{RE} I(\mathcal{L}_n \le \mathcal{L}_{n,\alpha}) + \hat{\beta}^{UE} I(\mathcal{L}_n > \mathcal{L}_{n,\alpha})$$

is the stochastic preliminary test least squares estimator,

$$\mathcal{L}_n = \frac{(R\hat{\beta}^{UE} - r)'(RS^{-1}R' + \Omega)^{-1}(R\hat{\beta}^{UE} - r)}{qS_e^2},$$

with

$$S_e^2 = \frac{(y - Z\hat{\phi}^{UE})'(y - Z\hat{\phi}^{UE})}{n - p} = \frac{(y_1 - X\hat{\beta}^{UE})'(y_1 - X\hat{\beta}^{UE})}{n - p},$$

 $\mathcal{L}_{n,\alpha}$ 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, n-p) degrees of freedom while under the alternative hypothesis, $H\phi \neq 0$, the pdf of \mathcal{L}_n is given by

$$g_{q,n-p}(\mathcal{L}_n, \Delta, \nu) = \sum_{r=0}^{\infty} \left(\frac{q}{n-p}\right)^{q/2+r} \frac{\Gamma\left(\frac{\nu}{2}+r\right)\Gamma\left(\frac{n-p+q}{2}+r\right)}{\Gamma(r+1)\Gamma(\frac{\nu}{2})\Gamma(\frac{q}{2}+r)\Gamma\left(\frac{n-p}{2}\right)} \\ \times \frac{\left(\frac{\Delta}{\nu-2}\right)^r}{\left(1+\frac{\Delta}{\nu-2}\right)^{\nu/2+r}} \frac{\mathcal{L}_n^{q/2+r-1}}{\left(1+\frac{q}{n-p}\mathcal{L}_n\right)^{\frac{n-p+q}{2}+r}},$$
(17)

where

$$\Delta = \frac{\psi'(RC^{-1}R' + \Omega)^{-1}\psi}{\sigma_e^2},\tag{18}$$

is the departure parameter from the null hypothesis. It is important to remark that $\hat{\phi}^{PT}$ is bounded and performs better than $\hat{\phi}^{RE}$ in some part of the parameter space.

The preliminary test approach estimation has been pioneered by Bancroft (1944), followed by Han and Bancroft (1968), Saleh and Sen (1978), Giles (1991), Kibria and Saleh (1993), Benda (1996) and very recently Ahmed (2002) among others. They have assumed that the disturbances of the model are normally distributed.

Note that, the preliminary test estimator (PT) has two characteristics: (1) it produces only two values, the unrestricted estimator and the restricted estimator, (2) it depends heavily on the level of significance of the preliminary test (PT). What about the intermediate value between $\hat{\phi}^{UE}$ and $\hat{\phi}^{RE}$. To overcome this shortcoming, we consider the Stein-type estimator. The Stein-type shrinkage estimator (SE) of ϕ is defined as

$$\hat{\phi}^{SE} = \hat{\phi}^{UE} - d\mathcal{L}_n^{-1}(\hat{\phi}^{UE} - \hat{\phi}^{RE}),$$

$$= \begin{pmatrix} \hat{\beta}^{SE} \\ r - d\mathcal{L}_n^{-1}(r - R\hat{\beta}^{RE}) \end{pmatrix}$$
(19)

where,

$$\hat{\beta}^{SE} = \hat{\beta}^{UE} - d\mathcal{L}_n^{-1}(\hat{\beta}^{UE} - \hat{\beta}^{RE})$$

is the stochastic shrinkage estimator,

$$d = \frac{(q-2)(n-p)}{q(n-p+2)}, \text{ and } q \ge 3.$$

The SE in (19) will provide uniform improvement over $\hat{\phi}^{UE}$, however it is not a convex combination of $\hat{\phi}^{UE}$ and $\hat{\phi}^{RE}$. Both (16) and (19) involve the statistic \mathcal{L}_n which adjusts the estimator for departure from H_0 . For large value of \mathcal{L}_n both (16) and (19) yield $\hat{\beta}^{UE}$, while for small value of \mathcal{L}_n their performance is different. The SE has the disadvantage that it has strange behavior for small \mathcal{L}_n . Also the shrinkage factor $(1 - d\mathcal{L}_n^{-1})$ becomes negative for $\mathcal{L}_n < d$. This encourage one to find an alternative estimator. Hence, we define a better estimator, namely, the positive-rule shrinkage estimator (PR) of β as follows:

$$\hat{\phi}^{PR} = \hat{\phi}^{SE} - (1 - d\mathcal{L}_n^{-1})I(\mathcal{L}_n \leq d)(\hat{\phi}^{UE} - \hat{\phi}^{RE}),$$

$$= \begin{pmatrix} \hat{\beta}^{PR} \\ r - d\mathcal{L}_n^{-1}(r - R\hat{\beta}^{RE}) - (1 - d\mathcal{L}_n^{-1})I(\mathcal{L}_n \leq d)(r - R\hat{\beta}^{RE}) \end{pmatrix} \quad (20)$$

where

$$\hat{\phi}^{PR} = \hat{\beta}^{SE} - (1 - d\mathcal{L}_n^{-1})I(\mathcal{L}_n \le d)(\hat{\beta}^{UE} - \hat{\beta}^{RE})$$

is the stochastic positive rule shrinkage estimator.

The PR estimator in (20) will provide uniform improvement over $\hat{\phi}^{UE}$ and it is a convex combination of $\hat{\phi}^{UE}$ and $\hat{\phi}^{RE}$. The properties of stein-type estimators have been analyzed under normal assumption by various researchers such as, James and Stein (1961), Judge and Bock (1978) and Shalabh (1995) to mention a few. The positive part shrinkage estimator has been considered under the normal assumption by Ohtani (1993) and Adkins and Hill (1990) among others.

Since, we are interested only in the marginal estimate of β under the stochastic information, in the following section we will provide the bias and risk functions of the proposed estimators.

3 Bias and Risks of the Estimators

In this section we give the expressions for the bias and quadratic risk of the estimators $\hat{\beta}^{UE}$, $\hat{\beta}^{RE}$, $\hat{\beta}^{PT}$, $\hat{\beta}^{SE}$, and $\hat{\beta}^{PR}$. In the following subsection we will discuss about the biases of the estimators.

3.1 Biases of the Estimators

Note that the biases of the proposed estimators are routinely derived following Judge and Bock (1978, Chapter 10). Therefore, we omit all derivation, instead, we present the expressions for the biases of the estimators in the following theorem.

Theorem 3.1 Bias of the of the unrestricted estimator (UE), restricted estimator (RE), preliminary test estimator (PT), shrinkage estimator (SE) and the positive-rule shrinkage estimator (PR) under the possible stochastic constraints are given respectively

 $B(\hat{\beta}^{UE}) = E(\hat{\beta}^{UE} - \beta) = 0$

$$B(\hat{\beta}^{RE}) = E(\hat{\beta}^{RE} - \beta) = -\eta$$

$$B(\hat{\beta}^{PT}) = E(\hat{\beta}^{PT} - \beta) = -\eta G_{q+2,n-p}^{(2)}(x,\Delta)$$

$$B(\hat{\beta}^{SE}) = E(\hat{\beta}^{SE} - \beta) = -qd\eta E^{(2)}[\chi_{q+2}^{-2}(\Delta)]$$
 and

$$B(\hat{\beta}^{PR}) = E(\hat{\beta}^{PR} - \beta) = \eta \left\{ \frac{qd}{q+2} E^{(2)} \left[F_{q+2,n-p}^{-1}(\Delta) I \left(F_{q+2,n-p}(\Delta) \le \frac{qd}{q+2} \right) \right] - \frac{qd}{q+2} E^{(2)} [F_{q+2,n-p}^{-1}(\Delta)] - G_{q+2,n-p}^{(2)}(x,\Delta) \right\},$$
(21)

where $\eta = S^{-1}R'(RS^{-1}R' + \Omega)^{-1}\psi$ and

$$G_{q+2i,n-p}^{(j)}(x;\Delta) = \sum_{r=0}^{\infty} \frac{\Gamma\left(\frac{\nu}{2}+r+j-2\right)}{\Gamma(r+1)\Gamma\left(\frac{\nu}{2}+j-2\right)} \frac{\left(\frac{\Delta}{\nu-2}\right)^{r}}{\left(1+\frac{\Delta}{\nu-2}\right)^{\nu/2+r+j-2}} I_{x}\left\{\frac{1}{2}(q+2i)+r,\frac{n-p}{2}\right\},(22)$$

where $I\{.\}$ is the incomplete beta function and

$$x = \frac{qF_{\alpha}}{n - p + qF_{\alpha}}.$$

Also

$$E^{(j)}[\chi_{q+s}^{-2}(\Delta)] = \sum_{r=0}^{\infty} \frac{\Gamma\left(\frac{\nu}{2} + r + j - 2\right) \left(\frac{\Delta}{\nu-2}\right)^r (q+s-2+2r)^{-1}}{\Gamma(r+1)\Gamma\left(\frac{\nu}{2} + j - 2\right) \left(1 + \frac{\Delta}{\nu-2}\right)^{\frac{\nu}{2} + r + j - 2}}$$
$$E^{(j)}[\chi_{q+s}^{-2}(0)] = (q+s-2)^{-1}, \quad j=1,2$$
(23)

and

$$E^{(j)}\left[F_{q+s,n-p}^{-1}(\Delta)I\left(F_{q+s,n-p}(\Delta) \leq \frac{qd}{q+s}\right)\right]$$

$$= \sum_{r=0}^{\infty} \frac{\Gamma\left(\frac{\nu}{2}+r+j-2\right)\left(\frac{\Delta}{\nu-2}\right)^{r}(q+s)I_{x^{*}}\left[\frac{1}{2}(q+s-2+2r),\frac{1}{2}(n-p+2)\right]}{\Gamma(r+1)\Gamma\left(\frac{\nu}{2}+j-2\right)\left(1+\frac{\Delta}{\nu-2}\right)^{\frac{\nu}{2}+r+j-2}(q+s-2+2r)}$$

$$E^{(j)}\left[F_{q+s,n-p}^{-1}(0)I\left(F_{q+s,n-p}(0) \leq \frac{qd}{q+s}\right)\right]$$

$$= (q+s)(q+s-2)^{-1}I_{x^{*}}\left[\frac{1}{2}(q+s-2),\frac{1}{2}(n-p+2)\right], \quad j=1,2$$
(24)

where

$$x^* = \frac{qd}{n - p + qd}.$$

For $\alpha = 0$, the bias of $\hat{\beta}^{PT}$ coincides with that of the restricted estimator, $\hat{\beta}^{RE}$, while for $\alpha = 1$, it coincides with that of $\hat{\beta}^{UE}$, the unrestricted estimator β . Also as the departure parameter $\Delta \to \infty$, $B(\hat{\beta}^{UE}) = B(\hat{\beta}^{PT}) = B(\hat{\beta}^{SE}) = B(\hat{\beta}^{PR}) = 0$, while the $B(\hat{\beta}^{RE})$ becomes unbounded. However, under H_0 all the estimators are unbiased.

Now we compare the biases under the alternative hypothesis. In order to present a clear cut picture of the biases, we transform them in convenient quadratic (scaler) form. Let $\hat{\beta}^*$ be any estimator of β and W be the positive semidefinite matrix, then the quadratic bias function is defined as $QB(\hat{\beta}^*) = B(\hat{\beta}^*)'WB(\hat{\beta}^*)$. The quadratic bias functions of the estimators can be expressed by the following theorem.

Theorem 3.2 Quadratic bias of the of the unrestricted estimator (UE), restricted estimator (RE), preliminary test estimator (PT), shrinkage estimator (SE) and the positive-rule shrinkage estimator (PR) are given respectively

$$QB(\hat{\beta}^{UE}) = 0$$

$$QB(\hat{\beta}^{RE}) = \sigma_{e}^{2}\psi'D\psi$$

$$QB(\hat{\beta}^{PT}) = \sigma_{e}^{2}\psi'D\psi \left[G_{q+2,n-p}^{(2)}(x,\Delta)\right]^{2}$$

$$QB(\hat{\beta}^{SE}) = \sigma_{e}^{2}\psi'D\psi \left[qdE^{(2)}[\chi_{q+2}^{-2}(\Delta)]\right]^{2}$$

$$QB(\hat{\beta}^{PR}) = \sigma_{e}^{2}\psi'D\psi \left\{\frac{qd}{q+2}E^{(2)}\left[F_{q+2,n-p}^{-1}(\Delta)I\left(F_{q+2,n-p}(\Delta) \leq \frac{qd}{q+2}\right)\right]$$

$$- \frac{qd}{q+2}E^{(2)}[F_{q+2,n-p}^{-1}(\Delta)] - G_{q+2,n-p}^{(2)}(x,\Delta)\right\}^{2}, \qquad (25)$$

where

$$D = [RS^{-1}R' + \Omega]^{-1}RS^{-1}WS^{-1}R'[RS^{-1}R' + \Omega]^{-1}.$$

The quadratic bias functions of all the estimators except $\hat{\beta}^{UE}$ depends upon the parameters only through the Δ ; thus the bias is a function of Δ . The bias of the PT depends on α , the size of the test. The magnitude of the RE increases without a bound and tends to ∞ as $\Delta \to \infty$. Since both $E(\chi_{q+2}^{-2}(\Delta))$ and $G_{q+2,n-p}^{(2)}(x,\Delta)$ are decreasing function of Δ , the quadratic bias of PT, SE and PR estimators start from 0 and increase to a point and then decrease gradually to 0 when $\Delta \to \infty$. However, the bias of PR estimator remain below the curve of SE and PT estimators. Based on the above analysis we may establish the following inequality

$$QB(\hat{\beta}^{UE}) \le QB(\hat{\beta}^{PR}) \le QB(\hat{\beta}^{SE}) \le QB(\hat{\beta}^{PT}) \le QB(\hat{\beta}^{RE}),$$

if

$$E[\chi_{q+2}^{-2}(\Delta)] \le \frac{1}{qd} G_{q+2,n-p}(l_{\alpha};\Delta)$$

otherwise

$$QB(\hat{\beta}^{UE}) \le QB(\hat{\beta}^{PR}) \le QB(\hat{\beta}^{PT}) \le QB(\hat{\beta}^{SE}) \le QB(\hat{\beta}^{RE}).$$

if

$$E[\chi_{q+2}^{-2}(\Delta)] \ge \frac{1}{qd} G_{q+2,n-p}(l_{\alpha};\Delta)$$

3.2 Risks Analysis of the Estimators

In this subsection we will present the quadratic risk function of the proposed estimators. Let $\hat{\beta}^*$ be any estimator of β and W be the positive semidefinite matrix, then the quadratic loss function is defined as

$$L(\hat{\beta}^*;\beta) = (\hat{\beta}^* - \beta)' W(\hat{\beta}^* - \beta).$$
(26)

9

The corresponding risk function of the estimator $\hat{\beta}^*$ is defined as

$$R(\hat{\beta}^*;\beta) = E(\hat{\beta}^* - \beta)' W(\hat{\beta}^* - \beta).$$
(27)

The quadratic risk functions of the proposed estimators are routinely derived following Judge and Bock (1978) lead to the following theorem.

Theorem 3.3: Risks of UE, RE, PT, SE and PR are given respectively

$$\begin{split} R(\hat{\beta}^{UE};W) &= \sigma_e^2 tr(S^{-1}W) \\ R(\hat{\beta}^{RE};W) &= \sigma_e^2 tr(S^{-1}W) - \sigma_e^2 tr(A) + \psi' D\psi \\ R(\hat{\beta}^{PT};W) &= \sigma_e^2 tr(S^{-1}W) - \sigma_e^2 tr(A)G_{q+2,n-p}^{(1)}(x,\Delta) \\ &+ \psi' D\psi \left[2G_{q+2,n-p}^{(2)}(x,\Delta) - G_{q+4,n-p}^{(2)}(x,\Delta) \right] \\ R(\hat{\beta}^{SE};W) &= \sigma_e^2 tr(S^{-1}W) - dq\sigma_e^2 tr(A) \left\{ 2E^{(1)}[\chi_{q+2}^{-2}(\Delta)] - (q-2)E^{(1)}[\chi_{q+2}^{-4}(\Delta)] \right\} \\ &+ qd\psi' D\psi \left\{ (q-2)E^{(2)}[\chi_{q+4}^{-4}(\Delta)] + 2 \left(E^{(2)}[\chi_{q+2}^{-2}(\Delta)] - E^{(2)}[\chi_{q+4}^{-2}(\Delta)] \right) \right\} \\ &= \sigma_e^2 tr(S^{-1}W) - dq\sigma_e^2 tr(A) \left\{ (q-2)E^{(1)}[\chi_{q+2}^{-4}(\Delta)] + \left[1 - \frac{(q+2)\psi' D\psi}{2\sigma_e^2 \Delta tr(A)} \right] 2\Delta E^{(2)}[\chi_{q+4}^{-4}(\Delta)] \right\} \\ R(\hat{\beta}^{PR};W) &= R(\hat{\beta}_n^{SE}(k);\beta) \\ &- \sigma_e^2 \left\{ tr(A)E^{(1)} \left[\left(1 - \frac{qd}{q+2}F_{q+2,m}^{-1}(\Delta) \right)^2 I \left(F_{q+2,n-p}(\Delta) < \frac{qd}{q+2} \right) \right] \right\} \\ &+ \frac{\psi' D\psi}{\sigma_e^2} E^{(2)} \left[\left(1 - \frac{qd}{q+4}F_{q+4,n-p}^{-1}(\Delta) - 1 \right) I \left(F_{q+2,n-p}(\Delta) < \frac{qd}{q+2} \right) \right], (28) \end{split}$$

where

$$A = (RS^{-1}R' + \Omega)^{-1}RS^{-1}WS^{-1}R'$$

and

$$D = [RS^{-1}R' + \Omega]^{-1}RS^{-1}WS^{-1}R'[RS^{-1}R' + \Omega]^{-1}.$$

Also,

$$E^{(j)}[\chi_{q+s}^{-4}(\Delta)] = \sum_{r=0}^{\infty} \frac{\Gamma\left(\frac{\nu}{2}+r+j-2\right) \left(\frac{\Delta}{\nu-2}\right)^r (q+s-2+2r)^{-1} (q+s-4+2r)^{-1}}{\Gamma(r+1)\Gamma\left(\frac{\nu}{2}+j-2\right) \left(1+\frac{\Delta}{\nu-2}\right)^{\frac{\nu}{2}+r+j-2}} E^{(j)}[\chi_{q+s}^{-4}(0)] = (q+s-2)^{-1} (q+s-4)^{-1}, \quad j=1,2$$
(29)

and

$$E^{(j)}\left[F_{q+s,n-p}^{-2}(\Delta)I\left(F_{q+s,n-p}(\Delta) \leq \frac{qd}{q+s}\right)\right]$$

$$= \sum_{r=0}^{\infty} \frac{\Gamma\left(\frac{\nu}{2}+r+j-2\right)\left(\frac{\Delta}{\nu-2}\right)^{r}(q+s)^{2}I_{x^{*}}\left[\frac{1}{2}(q+s-4+2r),\frac{1}{2}(n-p+4)\right]}{\Gamma(r+1)\Gamma\left(\frac{\nu}{2}+j-2\right)\left(1+\frac{\Delta}{\nu-2}\right)^{\frac{\nu}{2}+r+j-2}(n-p)(q+s-2+2r)(q+s-4+2r)}$$

$$E^{(j)}\left[F_{q+s,n-p}^{-2}(0)I\left(F_{q+s,n-p}(0) \leq \frac{qd}{q+s}\right)\right]$$

$$= \frac{(q+s)^{2}}{(n-p)(q+s-2)(q+s-4)}I_{x^{*}}\left[\frac{1}{2}(q+s-4),\frac{1}{2}(n-p+4)\right], \quad j=1,2$$
(30)

Based on the above information we consider the performance of the estimators in the following section.

4 Risk Analysis of the Estimators

In this section we will compare the performance of the proposed estimators in the light of quadratic risk function. For our convenience we assume that ν is known. We obtain from Anderson (1984, Theorem A.2.4, p.590) that

$$\lambda_p \le \frac{\psi' D\psi}{\psi' [RS^{-1}R' + \Omega]^{-1}\psi} \le \lambda_1, \quad \text{or}$$

$$\sigma_e^2 \Delta \lambda_p \le \psi' D\psi \le \sigma_e^2 \Delta \lambda_1, \quad (31)$$

where λ_1 and λ_p are the largest and the smallest characteristic roots of the matrix A.

4.1 Comparison of $\hat{\beta}^{UE}$ and $\hat{\beta}^{RE}$

First, we compare between $\hat{\beta}^{UE}$ and $\hat{\beta}^{RE}$. Using (28), the risk difference is,

$$R(\hat{\beta}^{UE};W) - R(\hat{\beta}^{RE};W) = \sigma_e^2 tr(A) - \psi' D\psi.$$
(32)

The difference in (32) will be non-negative whenever $\Delta \leq \frac{tr(A)}{\lambda_1}$. That is RE will dominate UE when $\Delta \leq \frac{tr(A)}{\lambda_1}$, otherwise UE will dominate RE when $\Delta \geq \frac{tr(A)}{\lambda_p}$. For W = S, and $\Omega = 0$, we note that $\hat{\beta}^{RE}$ performs better than $\hat{\beta}^{UE}$ in the interval [0, q] and worse outside this interval.

4.2 Comparison of $\hat{\beta}^{UE}$, $\hat{\beta}^{RE}$ and $\hat{\beta}^{PT}$

First we compare $\hat{\beta}^{PT}$ versus $\hat{\beta}^{UE}$. The risk difference is

$$R(\hat{\beta}^{UE}; W) - R(\hat{\beta}^{PT}; W) = \sigma_e^2 tr(A) G_{q+2,n-p}^{(1)}(x, \Delta)$$

$$- \psi' D\psi \left[2G_{q+2,n-p}^{(2)}(x,\Delta) - G_{q+4,n-p}^{(2)}(x,\Delta) \right].$$
(33)

The difference in (33) will be non-negative whenever

$$\Delta \le \frac{tr(A)G_{q+2,n-p}^{(1)}(x,\Delta)}{\lambda_1 \left[2G_{q+2,n-p}^{(2)}(x,\Delta) - G_{q+4,n-p}^{(2)}(x,\Delta) \right].}$$
(34)

Thus $\hat{\beta}^{PT}$ is superior to $\hat{\beta}^{UE}$ if $\Delta \in \left[0, \frac{tr(A)G_{q+2,n-p}^{(1)}(x,\Delta)}{\lambda_1 \left[2G_{q+2,n-p}^{(2)}(x,\Delta) - G_{q+4,n-p}^{(2)}(x,\Delta)\right]}\right)$ and $\hat{\beta}^{UE}$ performs better than $\hat{\beta}^{PT}$ if $\Delta \in \left[\frac{tr(A)G_{q+2,n-p}^{(1)}(x,\Delta)}{\lambda_p \left[2G_{q+2,n-p}^{(2)}(x,\Delta) - G_{q+4,n-p}^{(2)}(x,\Delta)\right]}, \infty\right)$ It follows from

(33) that under H_0 , $\hat{\beta}^{PT}$ is superior to $\hat{\beta}^{UE}$ for all $\alpha \in (0, 1)$. We can describe the graph of $R(\hat{\beta}^{PT}; W)$ as follows. It assumes a value of $\sigma_e^2 tr(S^{-1}W) - \sigma_e^2 tr(A)G_{q+2,n-p}^{(1)}(x,0)$ at $\Delta = 0$, then increase crossing the risk of $\hat{\beta}^{UE}$ to a maximum then drops gradually towards $\sigma_e^2 tr(S^{-1}W)$ as $\Delta \to \infty$.

Now we compare the risk between $\hat{\beta}^{RE}$ and $\hat{\beta}^{PT}$. Both are superior than $\hat{\beta}^{UE}$ under the null hypothesis. We note that

$$R(\hat{\beta}^{RE};W) - R(\hat{\beta}^{PT};W) = -\sigma_e^2 tr(A)[1 - G_{q+2,n-p}^{(1)}(x,\Delta)] + \psi' D\psi \left[1 - 2G_{q+2,n-p}^{(2)}(x,\Delta) + G_{q+4,n-p}^{(2)}(x,\Delta)\right] (35)$$

(1)

The difference in (35) will be non-positive whenever

$$\Delta \le \frac{tr(A)[1 - G_{q+2,n-p}^{(1)}(x,\Delta)]}{\lambda_1 \left[1 - 2G_{q+2,n-p}^{(2)}(x,\Delta) + G_{q+4,n-p}^{(2)}(x,\Delta)\right]}.$$
(36)

 $\begin{array}{l} \text{Thus } \hat{\beta}^{PT} \text{ is superior to } \hat{\beta}^{RE} \text{ if } \Delta \in \left[0, \frac{tr(A)[1-G_{q+2,n-p}^{(1)}(x,\Delta)]}{\lambda_1 \left[1-2G_{q+2,n-p}^{(2)}(x,\Delta)+G_{q+4,n-p}^{(2)}(x,\Delta)\right]}\right) \text{ and } \hat{\beta}^{RE} \\ \text{performs better than } \hat{\beta}^{PT} \text{ if } \Delta \in \left[\frac{tr(A)[1-G_{q+2,n-p}^{(1)}(x,\Delta)]}{\lambda_p \left[1-2G_{q+2,n-p}^{(2)}(x,\Delta)+G_{q+4,n-p}^{(2)}(x,\Delta)\right]}, \infty\right). \end{array}$

Comparison of $\hat{\beta}^{UE}$, $\hat{\beta}^{RE}$, $\hat{\beta}^{PT}$ and $\hat{\beta}^{SE}$ 4.3

Now we investigate the comparative statistical properties of the Stein-type estimator. First we compare between UE and SE. The risk difference is

$$R(\hat{\beta}^{UE};W) - R(\hat{\beta}^{SE};W) = dq\sigma_e^2 tr(A) \left\{ (q-2)E^{(1)}[\chi_{q+2}^{-4}(\Delta)] \right\}$$

+
$$\left[1 - \frac{(q+2)\psi'D\psi}{2\sigma_e^2\Delta tr(A)}\right](2\Delta)E^{(2)}[\chi_{q+4}^{-4}(\Delta)]\right\}.$$
 (37)

Using (31), the risk difference in (37) is positive for all \mathcal{F} such that

$$\mathcal{F} = \left\{ \mathcal{F} : \frac{tr(A)}{\lambda_1} \ge \frac{q+2}{2} \right\}.$$
(38)

Thus $\hat{\beta}^{SE}$ uniformly dominates $\hat{\beta}^{UE}$. Further, as $\Delta \to \infty$, the risk difference tends to 0 from below. Now we wish to compare $\hat{\beta}^{RE}$ and $\hat{\beta}^{SE}$. We have

$$R(\hat{\beta}^{SE};W) - R(\hat{\beta}^{RE};W) = \sigma_e^2 tr(A) - \psi' D\psi - dq \sigma_e^2 tr(A) \left\{ (q-2)E^{(1)}[\chi_{q+2}^{-4}(\Delta)] + \left[1 - \frac{(q+2)\psi' D\psi}{2\sigma_e^2 \Delta tr(A)} \right] (2\Delta)E^{(2)}[\chi_{q+4}^{-4}(\Delta)] \right\}.$$
(39)

From (39) we note that under H_0 , $R(\hat{\beta}^{SE}; W) \geq R(\hat{\beta}^{RE}; W)$. Thus $\hat{\beta}^{RE}$ performs better than $\hat{\beta}^{SE}$ under H_0 . However, ψ moves away from 0, $\psi' D\psi$ increases and the risk of $\hat{\beta}^{RE}$ becomes unbounded while the risk of $\hat{\beta}^{SE}$ remains below the risk of $\hat{\beta}^{UE}$ and merges with it as $\Delta \to \infty$. Thus $\hat{\beta}^{SE}$ dominates $\hat{\beta}^{RE}$ outside an interval around the origin.

Now we compare $\hat{\beta}^{PT}$ and $\hat{\beta}^{SE}$ under H_0 . We have

$$R(\hat{\beta}^{SE}; W) - R(\hat{\beta}^{PT}; W) = \sigma_e^2 tr(A) \left[G_{q+2,n-p}^{(1)}(x,0) - d \right].$$
(40)

The above difference is positive for all $\alpha \in (0,1)$ such that F_{α} satisfies the following inequality

$$\left\{\alpha: F_{\alpha} > \frac{q+2}{q} F_{q+2,n-p}^{-1}(d,0)\right\}.$$
(41)

Thus PT dominates SE when (41) satisfies, while SE dominates PT when F_{α} satisfies the following inequality

$$\left\{\alpha: F_{\alpha} < \frac{q+2}{q} F_{q+2,n-p}^{-1}(d,0)\right\}.$$
(42)

Thus it is clear that Stein-type shrinkage estimator, $\hat{\beta}^{SE}$ does not always dominate PT under H_0 . Whichever dominates depend on size of the critical level.

Under the alternative hypothesis the risk difference is

$$\begin{aligned} R(\hat{\beta}^{SE};W) - R(\hat{\beta}^{PT};W) &= -\sigma_e^2 tr(A) \left\{ qd \left[2E^{(1)}[\chi_{q+2}^{-2}(\Delta)] - (q-2)E^{(1)}[\chi_{q+2}^{-4}(\Delta)] \right] \\ &- G^{(1)}_{q+2,n-p}(x;\Delta) \right\} + \psi' D\psi \left\{ dq(q-2)E^{(2)}[\chi_{q+2}^{-4}(\Delta)] \right. \end{aligned}$$

International Journal of Statistical Sciences, Vol. 3, 2004

+
$$2qd\left[E^{(2)}[\chi_{q+2}^{-2}(\Delta)] - E^{(2)}[\chi_{q+4}^{-2}(\Delta)]\right]$$

- $\left[2G_{q+2,n-p}^{(2)}(x;\Delta) - G_{q+4,n-p}^{(2)}(x;\Delta)\right]\right\}.$ (43)

The risk difference in (43) is positive and therefore PT will dominate SE if

$$\Delta \ge \frac{tr(A)\left\{dq\left(2E^{(1)}[\chi_{q+2}^{-2}(\Delta)] - (q-2)E^{(1)}[\chi_{q+2}^{-4}(\Delta)]\right) - G^{(1)}_{q+2,n-p}(x;\Delta)\right\}}{\lambda_p \times f_1(\Delta,\alpha)}, \quad (44)$$

while SE will dominate PT whenever

$$\Delta \leq \frac{tr(A)\left\{dq\left(2E^{(1)}[\chi_{q+2}^{-2}(\Delta)] - (q-2)E^{(1)}[\chi_{q+2}^{-4}(\Delta)]\right) - G^{(1)}_{q+2,n-p}(x;\Delta)\right\}}{\lambda_1 \times f_1(\Delta,\alpha)}, \quad (45)$$

where

$$f_{1}(\Delta, \alpha) = dq(q-2)E^{(2)}[\chi_{q+2}^{-4}(\Delta)] + 2qd\left[E^{(2)}[\chi_{q+2}^{-2}(\Delta)] - E^{(2)}[\chi_{q+4}^{-2}(\Delta)]\right] - \left[2G_{q+2,n-p}^{(2)}(x;\Delta) - G_{q+4,n-p}^{(2)}(x;\Delta)\right].$$

Thus under alternative hypothesis, RE will dominate SE if

$$\Delta \leq \frac{tr(A) \left\{ dq \left(2E^{(1)}[\chi_{q+2}^{-2}(\Delta)] - (q-2)E^{(1)}[\chi_{q+2}^{-4}(\Delta)] \right) - 1 \right\}}{\lambda_1 \left\{ dq(q-2)E^{(2)}[\chi_{q+2}^{-4}(\Delta)] + 2qd \left[E^{(2)}[\chi_{q+2}^{-2}(\Delta)] - E^{(2)}[\chi_{q+4}^{-2}(\Delta)] \right] - 1 \right\}},$$

while SE will dominate RE if

$$\Delta \geq \frac{tr(A)\left\{dq\left(2E^{(1)}[\chi_{q+2}^{-2}(\Delta)] - (q-2)E^{(1)}[\chi_{q+2}^{-4}(\Delta)]\right) - 1\right\}}{\lambda_p\left\{dq(q-2)E^{(2)}[\chi_{q+2}^{-4}(\Delta)] + 2qd\left[E^{(2)}[\chi_{q+2}^{-2}(\Delta)] - E^{(2)}[\chi_{q+4}^{-2}(\Delta)]\right] - 1\right\}}.$$

4.4 Comparison of $\hat{\beta}^{UE}$, $\hat{\beta}^{RE}$, $\hat{\beta}^{PT}$, $\hat{\beta}^{SE}$, $\hat{\beta}^{PR}$

First we compare between $\hat{\beta}^{UE}$ and $\hat{\beta}^{PR}$. From (28) and (31) it is observed that

$$R(\hat{\beta}^{PR}; W) \le R(\hat{\beta}^{UE}; W), \quad \forall \Delta, \ q \ge 3.$$
(46)

This $\hat{\beta}^{PR}$ uniformly dominates $\hat{\beta}^{UE}$. Further the risk of $\hat{\beta}^{PR}$ remains below the risk of $\hat{\beta}^{UE}$ and merges with it when $\Delta \to \infty$. To compare $\hat{\beta}^{RE}$ and $\hat{\beta}^{PR}$, under null hypothesis, we have

$$R(\hat{\beta}^{PR};W) - R(\hat{\beta}^{RE};W) = \sigma_e^2 tr(A) \left\{ (1-d) - E^{(1)} \left[\left(1 - \frac{qd}{q+2} F_{q+2,n-p}^{-1}(0) \right)^2 \right] \right\} + \frac{1}{2} \left[\left(1 - \frac{qd}{q+2} F_{q+2,n-p}^{-1}(0) \right)^2 \right] + \frac{1}{2} \left[\left(1 - \frac{qd}{q+2} F_{q+2,n-p}^{-1}(0) \right)^2 \right] + \frac{1}{2} \left[\left(1 - \frac{qd}{q+2} F_{q+2,n-p}^{-1}(0) \right)^2 \right] + \frac{1}{2} \left[\left(1 - \frac{qd}{q+2} F_{q+2,n-p}^{-1}(0) \right)^2 \right] + \frac{1}{2} \left[\left(1 - \frac{qd}{q+2} F_{q+2,n-p}^{-1}(0) \right)^2 \right] + \frac{1}{2} \left[\left(1 - \frac{qd}{q+2} F_{q+2,n-p}^{-1}(0) \right)^2 \right] + \frac{1}{2} \left[\left(1 - \frac{qd}{q+2} F_{q+2,n-p}^{-1}(0) \right)^2 \right] + \frac{1}{2} \left[\left(1 - \frac{qd}{q+2} F_{q+2,n-p}^{-1}(0) \right)^2 \right] + \frac{1}{2} \left[\left(1 - \frac{qd}{q+2} F_{q+2,n-p}^{-1}(0) \right)^2 \right] + \frac{1}{2} \left[\left(1 - \frac{qd}{q+2} F_{q+2,n-p}^{-1}(0) \right)^2 \right] + \frac{1}{2} \left[\left(1 - \frac{qd}{q+2} F_{q+2,n-p}^{-1}(0) \right)^2 \right] + \frac{1}{2} \left[\left(1 - \frac{qd}{q+2} F_{q+2,n-p}^{-1}(0) \right)^2 \right] + \frac{1}{2} \left[\left(1 - \frac{qd}{q+2} F_{q+2,n-p}^{-1}(0) \right)^2 \right] + \frac{1}{2} \left[\left(1 - \frac{qd}{q+2} F_{q+2,n-p}^{-1}(0) \right)^2 \right] + \frac{1}{2} \left[\left(1 - \frac{qd}{q+2} F_{q+2,n-p}^{-1}(0) \right)^2 \right] + \frac{1}{2} \left[\left(1 - \frac{qd}{q+2} F_{q+2,n-p}^{-1}(0) \right)^2 \right] + \frac{1}{2} \left[\left(1 - \frac{qd}{q+2} F_{q+2,n-p}^{-1}(0) \right)^2 \right] + \frac{1}{2} \left[\left(1 - \frac{qd}{q+2} F_{q+2,n-p}^{-1}(0) \right)^2 \right] + \frac{1}{2} \left[\left(1 - \frac{qd}{q+2} F_{q+2,n-p}^{-1}(0) \right)^2 \right] + \frac{1}{2} \left[\left(1 - \frac{qd}{q+2} F_{q+2,n-p}^{-1}(0) \right)^2 \right] + \frac{1}{2} \left[\left(1 - \frac{qd}{q+2} F_{q+2,n-p}^{-1}(0) \right)^2 \right] + \frac{1}{2} \left[\left(1 - \frac{qd}{q+2} F_{q+2,n-p}^{-1}(0) \right)^2 \right] + \frac{1}{2} \left[\left(1 - \frac{qd}{q+2} F_{q+2,n-p}^{-1}(0) \right)^2 \right] + \frac{1}{2} \left[\left(1 - \frac{qd}{q+2} F_{q+2,n-p}^{-1}(0) \right)^2 \right] + \frac{1}{2} \left[\left(1 - \frac{qd}{q+2} F_{q+2,n-p}^{-1}(0) \right)^2 \right] + \frac{1}{2} \left[\left(1 - \frac{qd}{q+2} F_{q+2,n-p}^{-1}(0) \right)^2 \right] + \frac{1}{2} \left[\left(1 - \frac{qd}{q+2} F_{q+2,n-p}^{-1}(0) \right)^2 \right] + \frac{1}{2} \left[\left(1 - \frac{qd}{q+2} F_{q+2,n-p}^{-1}(0) \right)^2 \right] + \frac{1}{2} \left[\left(1 - \frac{qd}{q+2} F_{q+2,n-p}^{-1}(0) \right] + \frac{1}{2} \left[\left(1 - \frac{qd}{q+2} F_{q+2,n-p}^{-1}(0) \right] + \frac{1}{2} \left[\left(1 - \frac{qd}{q+2} F_{q+2,n-p}^{-1}(0) \right] + \frac{1}{2} \left[\left(1 - \frac{qd}{q+2} F_{q+2,n-p}^{-1}(0) \right] + \frac{1}{2} \left[\left(1 - \frac{qd}{q+2} F_{q+2,n-p}^{-1}(0) \right] + \frac{1$$

Tabatabaey, Saleh and Kibria: Estimation of Regression Parameters 15

$$\times I\left(F_{q+2,n-p}(0) \le \frac{dq}{q+2}\right)\right] \right\}.$$
(47)

Since

$$E^{(1)}\left[\left(1 - \frac{qd}{q+2}F_{q+2,n-p}^{-1}(0)\right)^2 I\left(F_{q+2,n-p}(0) \le \frac{dq}{q+2}\right)\right]$$

$$\leq E\left[\left(1 - \frac{qd}{q+2}F_{q+2,n-p}^{-1}(0)\right)^2\right] = 1 - d,$$

the difference in (47) is always positive. This $\hat{\beta}^{RE}$ performs better than $\hat{\beta}^{PR}$ under H_0 . However, ψ moves away from 0, $\psi' D \psi$ increases and the risk of $\hat{\beta}^{RE}$ becomes unbounded while the risk of $\hat{\beta}^{PR}$ remains below the risk of $\hat{\beta}^{UE}$ and merges with it as $\Delta \to \infty$. Thus $\hat{\beta}^{PR}$ dominates $\hat{\beta}^{RE}$ outside an interval around the origin.

Now we compare $\hat{\beta}^{PT}$ and $\hat{\beta}^{PR}$. Under H_0 , the risk difference is

$$R(\hat{\beta}^{PR}; W) - R(\hat{\beta}^{PT}; W) = \sigma_e^2 tr(A) \left\{ (G_{q+2,n-p}^{(1)}(x,0) - d - E^{(1)} \left[\left(1 - \frac{qd}{q+2} F_{q+2,n-p}^{-1}(0) \right)^2 \times I \left(F_{q+2,n-p}(0) \le \frac{dq}{q+2} \right) \right] \right\}.$$
(48)

The difference in (48) is always positive for all α satisfying the condition

$$\left\{\alpha: F_{\alpha} > \frac{q+2}{q} F_{q+2,n-p}^{-1}(d^*,0)\right\},\tag{49}$$

where $d^* = d + E^{(1)} \left(1 - \frac{qd}{q+2} F_{q+2,n-p}^{-1}(0) \right)^2 \times I \left(F_{q+2,n-p}(0) \leq \frac{dq}{q+2} \right)$. The risk of $\hat{\beta}^{PR}$ is smaller than that of the risk of $\hat{\beta}^{PT}$ when the critical value F_{α} satisfies the following condition

$$\left\{\alpha: F_{\alpha} < \frac{q+2}{q}F_{q+2,n-p}^{-1}(d^*,0)\right\}.$$
(50)

This leads to the conclusion that neither of the estimators, $\hat{\beta}^{PR}$ or $\hat{\beta}^{PT}$ uniformly dominate under H_0 . This is because, under H_0 , the PT reduces to RE.

Now we compare $\hat{\beta}^{PR}$ and $\hat{\beta}^{PT}$ under the alternative hypothesis. The risk difference is

$$\begin{aligned} R(\hat{\beta}^{PR};W) &- R(\hat{\beta}^{PT};W) \\ &= -\sigma_e^2 tr(A) \left\{ dq \left(2E^{(1)}[\chi_{q+2}^{-2}(\Delta)] - (q-2)E^{(1)}[\chi_{q+2}^{-4}(\Delta)] \right) \right. \end{aligned}$$

International Journal of Statistical Sciences, Vol. 3, 2004

$$+ E^{(1)} \left[\left(1 - \frac{qd}{q+2} F_{q+2,n-p}^{-1}(\Delta) \right)^2 I \left(F_{q+2,n-p}(\Delta) \le \frac{dq}{q+2} \right) \right] - G^{(1)}_{q+2,n-p}(x,\Delta) + \psi' D\psi \left\{ qd(q-2)E^{(2)}[\chi_{q+2}^{-4}(\Delta)] + 2qd \left(E^{(2)}[\chi_{q+2}^{-2}(\Delta)] - E^{(2)}[\chi_{q+2}^{-2}(\Delta)] \right) \right. - \left(2G^{(2)}_{q+2,n-p}(x,\Delta) - G^{(2)}_{q+4,n-p}(x,\Delta) \right) + E^{(2)} \left[\left(1 - \frac{qd}{q+2} F_{q+2,n-p}^{-1}(\Delta) \right)^2 I \left(F_{q+2,n-p}(\Delta) \le \frac{dq}{q+2} \right) \right] \\ - E^{(2)} \left[\left(\frac{qd}{q+2} F_{q+2,n-p}^{-1}(\Delta) - 1 \right) \times I \left(F_{q+2,n-p}(\Delta) \le \frac{dq}{q+2} \right) \right] \right\}.$$
(51)

The right hand side of (51) will be non-negative if

$$\Delta \ge \frac{f_2(\Delta, \alpha)}{\lambda_p \times f_3(\Delta, \alpha)},\tag{52}$$

where

$$f_{2}(\Delta, \alpha) = \sigma_{e}^{2} tr(A) \left\{ dq \left(2E^{(1)}[\chi_{q+2}^{-2}(\Delta)] - (q-2)E^{(1)}[\chi_{q+2}^{-4}(\Delta)] \right) + E^{(1)} \left[\left(1 - \frac{qd}{q+2}F_{q+2,n-p}^{-1}(\Delta) \right)^{2} I \left(F_{q+2,n-p}(\Delta) \le \frac{dq}{q+2} \right) \right] - G_{q+2,n-p}^{(1)}(x,\Delta) \right\}$$
(53)

and

$$f_{3}(\Delta, \alpha) = \left\{ qd(q-2)E^{(2)}[\chi_{q+2}^{-4}(\Delta)] + 2qd\left(E^{(2)}[\chi_{q+2}^{-2}(\Delta)] - E^{(2)}[\chi_{q+2}^{-2}(\Delta)]\right) - \left(2G^{(2)}_{q+2,n-p}(x,\Delta) - G^{(2)}_{q+4,n-p}(x,\Delta)\right) + E^{(2)}\left[\left(1 - \frac{qd}{q+2}F^{-1}_{q+2,n-p}(\Delta)\right)^{2}I\left(F_{q+2,n-p}(\Delta) \le \frac{dq}{q+2}\right)\right] - E^{(2)}\left[\left(\frac{qd}{q+2}F^{-1}_{q+2,n-p}(\Delta) - 1\right) \times I\left(F_{q+2,n-p}(\Delta) \le \frac{dq}{q+2}\right)\right]\right\}.(54)$$

Thus PR estimator will dominate PT estimator when (52) holds, while PT will dominate PR when

$$\Delta \le \frac{f_2(\Delta, \alpha)}{\lambda_1 \times f_3(\Delta, \alpha)}.$$
(55)

Finally we compare the risks of $\hat{\beta}^{PR}$ and $\hat{\beta}^{SE}$. The risk difference is given by

$$R(\hat{\beta}^{PR};W) - R(\hat{\beta}^{SE};W) = -\sigma_e^2 tr(A) E^{(1)} \left[\left(1 - \frac{qd}{q+2} F_{q+2,n-p}^{-1}(\Delta) \right)^2 I\left(F_{q+2,n-p}(\Delta) \le \frac{dq}{q+2} \right) \right]$$

16

Tabatabaey, Saleh and Kibria: Estimation of Regression Parameters 17

$$-\psi' D\psi E^{(2)} \left[\left(1 - \frac{qd}{q+4} F_{q+2,n-p}^{-1}(\Delta) \right)^2 I \left(F_{q+2,n-p}(\Delta) \le \frac{dq}{q+2} \right) \right] \\ - 2\psi' D\psi E^{(2)} \left[\left(\frac{qd}{q+2} F_{q+2,n-p}^{-1}(\Delta) - 1 \right) I \left(F_{q+2,n-p}(\Delta) \le \frac{dq}{q+2} \right) \right].$$
(56)

The right hand side of (56) is always negative since the expectation of a positive random variable is positive. Thus for all β , the risk of $\hat{\beta}^{PR}$ is smaller than that of $\hat{\beta}^{SE}$. Therefore, under the stochastic restriction, the positive rule shrinkage estimator (PR) not only confirms the inadmissibility of the shrinkage estimator (SE), but also demonstrates a simple superior estimator. Now, based on the above discussion we may state the following theorem.

Theorem 4.1: Under the null hypothesis and the inequalities (41), (42), (49), and (50) the dominance picture of the estimators is as follows

$$\hat{\beta}^{RE} \ge \hat{\beta}^{PT} \ge \hat{\beta}^{PR} \ge \hat{\beta}^{SE} \ge \hat{\beta}^{UE},\tag{57}$$

where the notations > means dominates in the sense of smaller risk. The position of preliminary test estimator may shift from "in between" $R(\hat{\beta}^{RE}; W)$ and $R(\hat{\beta}^{PR}; W)$ to "in between" $R(\hat{\beta}^{SE}; W)$ and $R(\hat{\beta}^{UR}; W)$. Thus the dominance picture under the H_0 may change as follows:

$$\hat{\beta}^{RE} \le \hat{\beta}^{PR} \le \hat{\beta}^{SE} \le \hat{\beta}^{PT} \le \hat{\beta}^{UE}.$$
(58)

The dominance pictures in (57) and (58) changes as ψ moves away from 0. We note that $\hat{\beta}^{UE}$ has constant risk $\sigma_e^2 tr(S^{-1}W)$ while the risk of $\hat{\beta}^{RE}$ depends on ψ and therefore, the risk of $\hat{\beta}^{RE}$ becomes unbounded as ψ moves always from 0. Also for $\Delta \to \infty$, the risk of $\hat{\beta}^{PT}$ and $\hat{\beta}^{PR}$ converge to the risk of $\hat{\beta}^{UE}$. For reasonable ψ near 0, the risk of $\hat{\beta}^{PT}$ is smaller than that of $\hat{\beta}^{PR}$ for $q \geq 3$. Thus neither $\hat{\beta}^{PT}$ nor $\hat{\beta}^{PR}$ dominates the other except they share common property that as $\Delta \to \infty$ the risk of both becomes $\sigma_e^2 tr(S^{-1}W)$. However the risk of $\hat{\beta}^{PR}$ is below the risk of $\hat{\beta}^{UE}$ while the risk of $\hat{\beta}^{PT}$ exceeds the risk of $\hat{\beta}^{UE}$ at some intermediate values of Δ depends on α .

5 Concluding Remarks

In this paper we have discussed some finite sample theory of five well known possible stochastic restricted estimators of β that are a combination of the sample and non sample information. The RE performs the best compare to other estimators in the neighborhood of the null hypothesis, however, it performs the worst when Δ moves away from its origin. We have demonstrated the superiority conditions of the estimators based on the quadratic risk function. We find that $\hat{\beta}^{SE}$ and $\hat{\beta}^{PR}$ are more efficient than $\hat{\beta}^{UE}$ in the whole parameter space. Note that the application of $\hat{\beta}^{PR}$ and $\hat{\beta}^{SE}$ is constrained by the requirement $q \geq 3$, while $\hat{\beta}^{PT}$ does not need such constraint. However, the choice of the level of significance of the test has a dramatic impact on the nature of the risk function for the PT estimator. Thus when $q \geq 3$, one would use $\hat{\beta}^{PR}$ otherwise $\hat{\beta}^{PT}$ with some optimum size α . Student's t distribution covers a class of symmetric and long tailed distributions for moderate degrees of freedom (ν). The distribution of the estimators are robust for $2 < \nu <$ moderate value of ν . However, when the degrees of freedom goes to infinity, the robustness property disappears and the properties of the estimators are induced by the normality assumptions. Therefore, we may conclude that the findings of this paper is also valid for normal distribution when degrees of freedoms is large.

It is noted that the stochastic estimators and the departure parameter depend on the covariance matrix Ω . It is advisable to consider the cases when requirements for knowledge of Ω are less demanding and more feasible. The prior information about the Ω may come either from the analysis of previous sample or from introspection which may be due to theory and casual observation. For more discussion about the possible solution for Ω , see Judge and Bock (1978).

Acknowledgments

The authors would like to thank the referee for his helpful comments. They also thankful to the Editor for his time on precessing this paper.

References

- Ahmed, S.E. (2002). Simultaneous estimation of coefficients of variation. Journal of Statistical Planning and Inference, 104, 31-51.
- [2] Anderson, T. W. (1984). An introduction to multivariate statistical analysis. Second Edition. John Wiley, NY.
- [3] Adkins, L. C. and Hill, R. C. (1990). The RLS positive part Stein estimator. American Journal of Agricultural Economics, 72, 727-730.
- [4] Bancroft, T. A. (1944). On biases in estimation due to use of preliminary tests of significance. Annals of Mathematics and Statistics, 15, 190-204.
- [5] Benda, N. (1996). Pre-test estimation and design in the linear model. J. Statistical Planning and Inference, 52, 225 -240.
- [6] Blattberg, R. C. and Gonedes, N. J. (1974). A comparison of the stable and Student t distributions as statistical models for stock prices. *Journal of Business*, 47, 224-280.

- [7] Fama, E. F. (1965). The behavior of stock market prices. *Journal of Business*, 38, 34-105.
- [8] Giles, A. J. (1991). Pretesting for Linear Restrictions in a Regression Model with Spherically Symmetric Distributions. J. Econometrics. 50, 377-398.
- [9] Gnanadesikan, R. (1977). Methods for Statistical Data Analysis of Multivariate Observations, New York: Wiley.
- [10] Han, C-P, and Bancroft, T. A. (1968). On Pooling means when variance is unknown. Journal of the American Statistical Association, 63, 1333-1342.
- [11] James, W. and Stein, C. (1961). Estimation with quadratic loss. Proceeding of the Fourth Barkeley Symposium on Mathematical Statistics and Probability. University of California Press, Berkeley, CA, 361-379.
- [12] 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.
- [13] Judge, G. C., Miyazaki, S. and Yancey, T. (1985). Minimax estimators for the location vectors of spherically symmetric densities. *Econometric Theory*, 1, 409-417.
- [14] Kibria, B. M. G. (1996). On shrinkage ridge regression estimators for restricted linear models with multivariate t disturbances. Students, 1 (3), 177-188.
- [15] Kibria, B. M. G. and Saleh, A. K. Md. E. (1993). Performance of shrinkage preliminary test estimator in regression analysis. *Jahangirnagar Review*. A 17, 133-148.
- [16] Ohtani, K. (1993). A comparison of the Stein-rule and positive part Stein-rule estimators in a misspecified linear regression models. *Econometric Theory*, 9, 668-679.
- [17] Saleh, A. K. Md. E. and Sen, P. K. (1978). Non-parametric estimation of location parametric after a preliminary test regression. *Annals of Statistics*, 6, 154-168.
- [18] Shalabh (1995). Performance of Stein-rule procedure for simultaneous prediction of actual and average values of study variable in linear regression models. *Bull. Internat. Statist. Inst.* 56, 1375-1390.
- [19] Sutradhar, B. C. and Ali, M. M. (1986). Estimation of the parameters of a regression model with a multivariate t error variable. *Communications and Statistics-Theory and Methods*, 15, 429 450.

- [20] Theil, H. (1963). On the use of incomplete prior information in regression analysis. Journal of the American Statistical Association, 58, 401-414.
- [21] Theil, H. and Goldberger, A. S. (1961). On pure and mixed statistical estimation in Economics. *International Economics Review*, 2, 65-78.
- [22] Zellner, A. (1976). Bayesian and non-Bayesian analysis of the regression model with multivariate Student t error terms, *Journal of the American Statistical As*sociation, 71, 400-405.