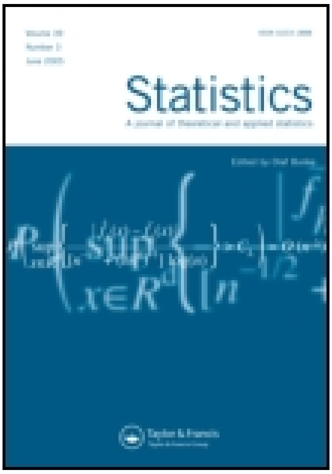

Statistics: A Journal of Theoretical and Applied Statistics
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/gsta20

Regression model with elliptically contoured errors

M. Arashi [a], A. K.Md.E. Saleh [b] & S. M.M. Tabatabaey [c]

[a] Faculty of Mathematics, Shahrood University of Technology, PO Box 316-3619995161, Shahrood, Iran

[b] School of Mathematics and Statistics, Carleton University, 1125 Colonel by drive, Ottawa, Ontario, Canada, K1S5B6

[c] Department of Statistics, School of Mathematical Sciences, Ferdowsi University of Mashhad, Mashhad, Iran

Published online: 21 Jun 2012.

To cite this article: M. Arashi, A. K.Md.E. Saleh & S. M.M. Tabatabaey (2013) Regression model with elliptically contoured errors, Statistics: A Journal of Theoretical and Applied Statistics, 47:6, 1266-1284, DOI: 10.1080/02331888.2012.694442

To link to this article: http://dx.doi.org/10.1080/02331888.2012.694442

Statistics, 2013
Vol. 47, No. 6, 1266–1284, http://dx.doi.org/10.1080/02331888.2012.694442

Regression model with elliptically contoured errors

M. Arashi[a]*, A.K.Md.E. Saleh[b] and S.M.M. Tabatabaey[c]

[a]*Faculty of Mathematics, Shahrood University of Technology, PO Box 316-3619995161, Shahrood, Iran;* [b]*School of Mathematics and Statistics, Carleton University, 1125 Colonel by drive, Ottawa, Ontario, Canada K1S5B6;* [c]*Department of Statistics, School of Mathematical Sciences, Ferdowsi University of Mashhad, Mashhad, Iran*

(*Received 14 February 2011; final version received 1 March 2012*)

For the regression model $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\epsilon}$ where the errors follow the elliptically contoured distribution, we consider the least squares, restricted least squares , preliminary test, Stein-type shrinkage and positive-rule shrinkage estimators for the regression parameters, $\boldsymbol{\beta}$.

We compare the quadratic risks of the estimators to determine the relative dominance properties of the five estimators.

Keywords: elliptically contoured distribution; multivariate student's *t*; positive-rule shrinkage estimator; preliminary test estimator; signed measure; Stein-type shrinkage estimator

1. Introduction

The most important model belonging to the class of general linear hypotheses is the *multiple regression* model. The general purpose of *multiple regression* is to learn more about the relationship between several independent or predictor variables and a dependent or criterion variable.

Consider the multiple regression model

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\epsilon}, \tag{1}$$

where $\boldsymbol{y}$ is an n-vector of responses, $\boldsymbol{X}$ is an $n \times p$ non-stochastic design matrix with a full column rank p, $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)'$ is p-vector of regression coefficients and $\boldsymbol{\epsilon} = (\epsilon_1, \dots, \epsilon_n)'$ is the n-vector of random errors distributed according to the law belonging to the class of elliptically contoured distributions (ECDs), $\mathcal{E}_n(\boldsymbol{0}, \sigma^2\boldsymbol{V}, \psi)$ for $\sigma \in \mathbb{R}^+$ and un-structured known matrix $\boldsymbol{V} \in S(n)$, where $S(n)$ denotes the set of all positive definite matrices of order $(n \times n)$ with the following characteristic function:

$$\phi_{\boldsymbol{\epsilon}}(\boldsymbol{t}) = \psi(\sigma^2\boldsymbol{t}'\boldsymbol{V}\boldsymbol{t}) \tag{2}$$

for some functions $\psi : [0, \infty) \to \mathbb{R}$ say characteristic generator [1].

*Corresponding author. Email: m_arashi_stat@yahoo.com

If ϵ has a density, then it is of the form

$$f(e) \propto |\sigma^2 \mathbf{V}|^{-1/2} g\left(\frac{1}{\sigma^2}\boldsymbol{\epsilon}' \boldsymbol{V}^{-1}\boldsymbol{\epsilon}\right), \tag{3}$$

where $g(.)$ is a non-negative function over $\mathbb{R}^+$ such that $f(.)$ is a density function w.r.t (with respect to) a σ-finite measure μ on $\mathbb{R}^p$. In this case, notation $\boldsymbol{\epsilon} \sim \mathcal{E}_n(\mathbf{0}, \sigma^2\boldsymbol{V}, g)$ would probably be used.

It is sometimes difficult to have complete analysis of the regression model with ECD errors of the type (2) or (3). To overcome such difficulties, one may consider any of the three sub-classes of ECDs, namely,

(i) scale mixture of normal distributions,
(ii) Laplace class of mixture of normal distributions, and
(iii) signed measure mixture of normal distributions.

General formula for the above mixture of distributions is given by

$$f_{\boldsymbol{\epsilon}}(\boldsymbol{x}) = \int_0^\infty \mathcal{W}(t)\phi_{\mathcal{N}_n(\mathbf{0},t^{-1}\sigma^2\boldsymbol{V})}(\boldsymbol{x})\,\mathrm{d}t, \tag{4}$$

where $\phi_{\mathcal{N}_n(\mathbf{0},t^{-1}\sigma^2\boldsymbol{V})}(.)$ is the pdf (probability density function) of $\mathcal{N}_n(\mathbf{0}, t^{-1}\sigma^2\boldsymbol{V})$.

(a) If

$$\mathcal{W}(\tau) = 2\left(\Gamma\left(\frac{\gamma}{2}\right)\right)^{-1}\left(\frac{\gamma\sigma^2}{2}\right)^{\gamma/2}\tau^{-(\gamma+1)}\,\mathrm{e}^{-\gamma\sigma^2/2\tau^2}, \quad 0 < \gamma, \sigma^2, \tau < \infty, \tag{5}$$

then we have

$$f(\boldsymbol{\epsilon}) = \frac{\Gamma((n+\gamma)/2)|\boldsymbol{V}|^{-1/2}}{(\pi\gamma)^{n/2}\Gamma(\gamma/2)\sigma^n}\left(1 + \frac{\boldsymbol{\epsilon}'\boldsymbol{V}^{-1}\boldsymbol{\epsilon}}{\gamma\sigma^2}\right)^{-(1/2)(n+\gamma)}, \tag{6}$$

where $E(\boldsymbol{\epsilon}) = \mathbf{0}$ and $E(\boldsymbol{\epsilon}\boldsymbol{\epsilon}') = (n\gamma\sigma^2/(\gamma-2))\boldsymbol{V} = \sigma_e^2$ for $\gamma > 2$.

(b) Chu [2] considered

$$\mathcal{W}(t) = (2\pi)^{n/2}|\sigma^2\boldsymbol{V}|^{1/2}t^{-p/2}\mathcal{L}^{-1}[f(s)], \tag{7}$$

$\mathcal{L}^{-1}[f(s)]$ denotes the inverse Laplace transform of $f(s)$ with $s = [\boldsymbol{x}'(\sigma^2\boldsymbol{V})^{-1}\boldsymbol{x}/2]$. For some examples of $f(.)$ and $\mathcal{W}(.)$, see Arashi *et al.* (2010).

The inverse Laplace transform of $f(.)$ exists provided that the following conditions are satisfied:
(i) $f(t)$ is differentiable when t is sufficiently large.
(ii) $f(t) = o(t^{-m})$ as $t \to \infty$, $m > 1$.
Although, it is rather difficult to derive the inverse Laplace transform of some functions, we are able to handle it for many density generators of elliptical densities. We refer the readers to Debnath and Batta [3] for more specific details.

The mean of $\boldsymbol{\epsilon}$ is the zero-vector and the covariance-matrix of $\boldsymbol{\epsilon}$ is

$$\begin{aligned}\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}} = \mathrm{Cov}(\boldsymbol{\epsilon}) &= \int_0^\infty \mathrm{Cov}(\boldsymbol{\epsilon}|t)\mathcal{W}(t)\,\mathrm{d}t \\ &= \int_0^\infty \mathcal{W}(t)\,\mathrm{Cov}\{\mathcal{N}_p(\mathbf{0}, t^{-1}\sigma^2\boldsymbol{V})\}\,\mathrm{d}t \\ &= \left(\int_0^\infty t^{-1}\mathcal{W}(t)\,\mathrm{d}t\right)\sigma^2\boldsymbol{V}, \end{aligned} \tag{8}$$

provided the above integral exists.

Comparing the models (3) and (4), since $\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}} = \mathrm{Cov}(\boldsymbol{\epsilon}) = -2\psi'(0)\sigma^2\boldsymbol{V}$, using Equation (6) we can conclude that

$$-2\psi'(0) = \int_0^\infty t^{-1}\mathcal{W}(t)\,\mathrm{d}t.$$

Now suppose that $\boldsymbol{X} \sim \mathcal{E}_n(\boldsymbol{\mu}, \boldsymbol{V}, g)$. Then it is important to point out that since $\int_{\boldsymbol{x}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = 1$, using Fubini's theorem we have

$$\begin{aligned}
1 &= \int_{\boldsymbol{x}}\int_0^\infty \mathcal{W}(t)\phi_{\mathcal{N}_n(\boldsymbol{\mu},t^{-1}\boldsymbol{V})}(\boldsymbol{x})\,\mathrm{d}t\,\mathrm{d}\boldsymbol{x}\\
&= \int_0^\infty \mathcal{W}(t)\int_{\boldsymbol{x}} \phi_{\mathcal{N}_n(\boldsymbol{\mu},t^{-1}\boldsymbol{V})}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t\\
&= \int_0^\infty \mathcal{W}(t)\,\mathrm{d}t.
\end{aligned}$$

Thus for non-negative function $\mathcal{W}(.)$, it is a density. For a non-negative function $\mathcal{W}(.)$, the elliptical models can be interpreted as a scale mixture of normal distributions.

(c) Srivastava and Bilodeau [4] considered the signed measure, $\mathcal{W}(t)$ such that

(i)

$$\int_0^\infty t^{-1}\mathcal{W}^+(\mathrm{d}t) < \infty,$$

(ii)

$$\int_0^\infty t^{-1}\mathcal{W}^-(\mathrm{d}t) < \infty, \tag{9}$$

where $\mathcal{W}^+ - \mathcal{W}^-$ is the Jordan decomposition of $\mathcal{W}$ in positive and negative parts. Note that from (i)–(ii) of Equation (9),

$$\int_0^\infty t^{-1}\mathcal{W}(\mathrm{d}t) < \infty \tag{10}$$

and thus, $\mathrm{Cov}(\boldsymbol{\epsilon})$ exists under the subclass defined above.

This subclass contains the subclass defined by (b).

Remark 1.1 Regarding the above classifications, we should take the following notes:

(1) In all the above classes, we have

$$\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}} = -2\psi'(0)\sigma^2\boldsymbol{V} = \left(\int_0^\infty t^{-1}\mathcal{W}(t)\,\mathrm{d}t\right)\sigma^2\boldsymbol{V},$$

resulting in $-2\psi'(0) = \int_0^\infty t^{-1}\mathcal{W}(t)\,\mathrm{d}t$.

(2) The subclass (a) is neither contained in subclass (b) nor in the subclass (c). However, subclass (b) is contained in the subclass(c). Thus, all the implications about the subclass (c) can be used for the subclass (b).

(3) For the subclass (c), we can assure that $-2\psi'(0) = \int_0^\infty t^{-1}\mathcal{W}(t)\,\mathrm{d}t$ exists. However, it may not exist for the subclass (b).

Some of the well-known members of the class of ECDs are the multivariate normal, Kotz Type, Pearson Types II and VII, multivariate Student's *t*, multivariate Cauchy, Logistic, Bessel and generalized slash distributions. Dating back to Kelker [5], there are many known results concerning ECDs, in particular the mathematical properties and its application to statistical inference. These results have been put forward by Cambanis *et al.* [6], Muirhead [7], Fang *et al.* [1] and Gupta and Varga [8] among others.

The object of the paper is the estimation of the regression parameters, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)'$ when it is suspected that $\boldsymbol{\beta}$ may belong to the sub-space defined by $\boldsymbol{H}\boldsymbol{\beta} = \boldsymbol{h}$ where $\boldsymbol{H}$ is a $q \times p$ matrix of constants and $\boldsymbol{h}$ is a q-vector of known constants with focus on the Stein-type estimators of $\boldsymbol{\beta}$ in addition to preliminary test estimator (PTE) based on the error distributions belonging to the subclass(c) which includes subclass(b) described earlier.

The recent book of Saleh [9] dealing with the proposed estimators (Chapter 7) presents an overview on the topic under normal as well as non-parametric theory covering many standard statistical models. Tabatabaey [10] and Arashi and Tabatabaey [11] cover the theory with spherically symmetric distribution of errors developing many interesting calculations. For some systematic work that have been done so far in the context of Stein-type estimations see Srivastava and Bilodeau [4] and Arashi and Tabatabaey [12] under elliptical symmetry.

We organize our paper as follows: Section 2, contains the estimation and the test of hypothesis along with proposed estimators of $\boldsymbol{\beta}$, Section 3 deals with the bias, risk and MSE expressions of the proposed estimators while the analysis of the risks and comparisons are presented in Section 4. Concluding remarks are presented in Section 5.

2. Estimation and test of hypothesis

In this section, we present the estimate of $\boldsymbol{\beta}$ and σ^2 under least-square (LS) theory. Furthermore, we discuss the problem of testing the general linear hypothesis, $\boldsymbol{H}\boldsymbol{\beta} = \boldsymbol{h}$. The test of this hypothesis covers many special cases considered in practical situations.

Using standard conditions, it is well known that the generalized LS (GLS) estimator of $\boldsymbol{\beta}$ is

$$\tilde{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y} = \boldsymbol{C}^{-1}\boldsymbol{X}'\boldsymbol{y}, \quad \boldsymbol{C} = \boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}. \tag{11}$$

Under elliptical assumptions, its distribution is $\mathcal{E}_p(\boldsymbol{\beta}, \sigma^2\boldsymbol{C}^{-1}, g)$. Similarly, the estimate of the σ^2 is

$$\tilde{\sigma}^2 = \frac{1}{n}(\boldsymbol{y} - \boldsymbol{X}\tilde{\boldsymbol{\beta}})'\boldsymbol{V}^{-1}(\boldsymbol{y} - \boldsymbol{X}\tilde{\boldsymbol{\beta}}). \tag{12}$$

It is easy to show that

$$S^2 = \frac{(\boldsymbol{y} - \boldsymbol{X}\tilde{\boldsymbol{\beta}})'\boldsymbol{V}^{-1}(\boldsymbol{y} - \boldsymbol{X}\tilde{\boldsymbol{\beta}})}{n-p} \tag{13}$$

is an unbiased estimator of $\sigma_\epsilon^2 = -2\psi'(0)\sigma^2$.

For a test of $\boldsymbol{H}\boldsymbol{\beta} = \boldsymbol{h}$ (where $q < p$), we first consider the restricted estimator given by

$$\hat{\boldsymbol{\beta}} = \tilde{\boldsymbol{\beta}} - \boldsymbol{C}\boldsymbol{H}'\boldsymbol{V}_1(\boldsymbol{H}\tilde{\boldsymbol{\beta}} - \boldsymbol{h}), \quad \boldsymbol{V}_1 = (\boldsymbol{H}\boldsymbol{C}\boldsymbol{H}')^{-1}. \tag{14}$$

It can be directly verified that $\hat{\boldsymbol{\beta}} \sim \mathcal{E}_p(\boldsymbol{\beta} - \boldsymbol{C}\boldsymbol{H}'\boldsymbol{V}_1(\boldsymbol{H}\boldsymbol{\beta} - \boldsymbol{h}), \sigma^2\boldsymbol{V}_2, g)$ for $\boldsymbol{V}_2 = \boldsymbol{C}(\boldsymbol{I}_p - \boldsymbol{H}'\boldsymbol{V}_1\boldsymbol{H}\boldsymbol{C})$. Therefore, we get

$$\begin{aligned} E(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) &= -\boldsymbol{C}\boldsymbol{H}'\boldsymbol{V}_1(\boldsymbol{H}\boldsymbol{\beta} - \boldsymbol{h}) \\ &= -\boldsymbol{\delta} \quad \text{(say)}, \\ E(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})'(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) &= -2\sigma^2\Psi'(0)\mathbf{tr}(\boldsymbol{V}_2) + \boldsymbol{\delta}'\boldsymbol{\delta} \\ &= \sigma_\epsilon^2\mathbf{tr}(\boldsymbol{V}_2) + \boldsymbol{\delta}'\boldsymbol{\delta}. \end{aligned} \tag{15}$$

Similarly, under $\boldsymbol{H\beta} = \boldsymbol{h}$, the following estimator is unbiased for σ_{ϵ}^2.

$$S^{*2} = \frac{(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})'\boldsymbol{V}^{-1}(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})}{n - p + q}, \tag{16}$$

from the LS theory.

Now we consider the linear hypothesis $\boldsymbol{H\beta} = \boldsymbol{h}$ and obtain the likelihood ratio test (LRT) statistic for the null hypothesis $H_0 : \boldsymbol{H\beta} = \boldsymbol{h}$ as well as its non-null distribution. The LRT statistic is a consequence of the results of Anderson *et al.* [13].

THEOREM 2.1 *Let*

$$w = \{\boldsymbol{\beta} : \boldsymbol{\beta} \in \mathbb{R}^p, \boldsymbol{H\beta} = \boldsymbol{h}, \sigma > 0, \boldsymbol{V} \in S(n)\}$$

and

$$\Omega = \{\boldsymbol{\beta} : \boldsymbol{\beta} \in \mathbb{R}^p, \sigma > 0, \boldsymbol{V} \in S(n)\}.$$

Moreover, suppose $y^{n/2}g(y)$ has a finite positive maximum y_g. Then under the assumptions of model (1) *the LRT for testing the null-hypothesis $H_0 : \boldsymbol{H\beta} = \boldsymbol{h}$ is given by*

$$\mathcal{L}_n = \frac{(\boldsymbol{H}\tilde{\boldsymbol{\beta}} - \boldsymbol{h})'\boldsymbol{V}_1(\boldsymbol{H}\tilde{\boldsymbol{\beta}} - \boldsymbol{h})}{qS^2}, \tag{17}$$

and has the following generalized non-central F-distribution with pdf

$$\boldsymbol{g}^*_{q,m}(\mathcal{L}_n) = \sum_{r \geq 0} \frac{(q/m)^{(1/2)(q+2r)}\mathcal{L}_n^{(1/2)(q+2r-2)}K_r^{(0)}(\Delta_*^2)}{r!\boldsymbol{B}((q+2r)/2, m/2)(1 + (q/m)\mathcal{L}_n)^{(1/2)(q+m+2r)}}, \tag{18}$$

where $m = n - p$, $\Delta_*^2 = \theta/\sigma_{\epsilon}^2$ for $\theta = (\boldsymbol{H\beta} - \boldsymbol{h})'\boldsymbol{V}_1(\boldsymbol{H\beta} - \boldsymbol{h})$, and the mixing distribution becomes

$$K_r^{(0)}(\Delta_*^2) = [-2\psi'(0)]^r \left(\frac{\Delta_*^2}{2}\right)^r \int_0^{\infty} \frac{t^r}{r!} \mathrm{e}^{(-t\Delta_*^2[-2\psi'(0)])/2}\mathcal{W}(t)\,\mathrm{d}t. \tag{19}$$

For the proof see the appendix.

COROLLARY 2.1 *Under H_0, the pdf of $\mathcal{L}_n$ is given by*

$$\boldsymbol{g}^*_{q,m}(\mathcal{L}_n) = \frac{(q/m)^{q/2}\mathcal{L}_n^{q/2-1}}{\boldsymbol{B}(q/2, m/2)(1 + (q/m)\mathcal{L}_n)^{(1/2)(q+m)}}, \tag{20}$$

which is the central F-distribution with (q, m) d.f. For $W(\tau)$ given by Equation (5), *we get the results produced by Tabatabaey* [10].

Now, consider the calculations of the probability of that $\mathcal{L}_n \leq F_{\alpha}$, which gives the power function of the test as

$$\boldsymbol{G}^*_{q,m}(F_{\alpha}; \Delta_*^2) = \sum_{r \geq 0} \frac{1}{r!}K_r^{(0)}(\Delta_*^2)I_x\left[\frac{1}{2}(q + 2r), \frac{m}{2}\right], \tag{21}$$

where $x = qF_{\alpha}/(m + qF_{\alpha})$ and $I_x(a, b)$ is the incomplete Beta-function,

$$I_x(a, b) = \frac{1}{\boldsymbol{B}(a, b)} \int_0^x u^{a-1}(1 - u)^{b-1}\,\mathrm{d}u. \tag{22}$$

The function (21) stands for the power function at α-level of significance and may be called the generalized non-central F-distribution cdf (cumulative distribution function) of the statistic $\mathcal{L}_n$.

Similarly, the cdf of a generalized non-central chi-square distribution with γ d.f. may be written as

$$\mathcal{H}_{\gamma}^{*}(x;\delta^2) = \sum_{r\geq 0} \frac{1}{r!} K_r^{(0)}(\Delta_*^2)\mathcal{H}_{\gamma+2r}(x;0), \tag{23}$$

where $\mathcal{H}_{\gamma+2r}(x;0)$ is the cdf of Chi-square distribution with $\gamma + 2r$ d.f.

Following Judge and Bock [14], we may write the following lemma where $E_{\mathcal{N}}$ stands for the expectation w.r.t. normal theory. About new insights, Lemma 2.1 gives possible extension for evaluating specific moments under elliptical assumption. As it can be seen from the result of Section 3, it is always necessary to obtain the expectation of a measurable function for risk functions. To ease the understanding of what Lemma 2.1 proposes, consider a random elliptical variable $\boldsymbol{X} \sim \mathcal{E}_q(\boldsymbol{\mu}, \boldsymbol{V}, g)$, then for a Borel measurable function h, we have

$$E(h(\boldsymbol{X})) = \int_0^{\infty} E_{\mathcal{N}}(h(\boldsymbol{X}))\mathcal{W}(t)\,\mathrm{d}t,$$

where $E_{\mathcal{N}}(h(\boldsymbol{X}))$ shows taking the expectation of h under the model $\mathcal{N}_q(\boldsymbol{\mu}, \boldsymbol{V})$.

For the proof of relevant expectations in Lemma 2.1 under normality assumption see Judge and Bock [14] and Saleh [9].

LEMMA 2.1 *If* $\boldsymbol{w} \sim \mathcal{E}_q(\boldsymbol{\eta}, \sigma^2\boldsymbol{I}_q, g)$, *and* Φ *is a measurable function, then*

(i)

$$\begin{aligned} E[\boldsymbol{w}\Phi(\boldsymbol{w}'\boldsymbol{w})] &= \int_0^{\infty} E_{\mathcal{N}}[\boldsymbol{w}\Phi(\boldsymbol{w}'\boldsymbol{w})|t]\ \mathcal{W}(t)\,\mathrm{d}t \\ &= \boldsymbol{\eta}\int_0^{\infty} E_{\mathcal{N}}[\Phi(\chi_{q+2}^2(\Delta_t^2))|t]\mathcal{W}(t)\,\mathrm{d}t, \end{aligned} \tag{24}$$

(ii)

$$\begin{aligned} E[(\boldsymbol{w}\boldsymbol{w}')\Phi(\boldsymbol{w}'\boldsymbol{w})] = \boldsymbol{I}_q &\int_0^{\infty} E_{\mathcal{N}}[\Phi(\chi_{q+2}^2(\Delta_t^2))|t]\mathcal{W}(t)\,\mathrm{d}t \\ &+ \boldsymbol{\eta}'\boldsymbol{\eta}\int_0^{\infty} E_{\mathcal{N}}[\Phi(\chi_{q+4}^2(\Delta_t^2))|t]\mathcal{W}(t)\,\mathrm{d}t, \end{aligned} \tag{25}$$

(iii)

$$\begin{aligned} E[(\boldsymbol{w}'\boldsymbol{A}\boldsymbol{w})\Phi(\boldsymbol{w}'\boldsymbol{w})] = \boldsymbol{tr}(\boldsymbol{A}) &\int_0^{\infty} E_{\mathcal{N}}[\Phi(\chi_{q+2}^2(\Delta_t^2))|t]\mathcal{W}(t)\,\mathrm{d}t \\ &+ (\boldsymbol{\eta}\boldsymbol{A}'\boldsymbol{\eta})\int_0^{\infty} E_{\mathcal{N}}[\Phi(\chi_{q+2}^2(\Delta_t^2))|t]\mathcal{W}(t)\,\mathrm{d}t, \end{aligned} \tag{26}$$

where $\Delta_t^2 = t\boldsymbol{\eta}'\boldsymbol{\eta}, \boldsymbol{A} \in S(q)$.

Further,

(i)

$$E^{(2-h)}[\chi_{q+s}^{*-2}(\Delta_*^2)] = \sum_{r\geq 0} \frac{1}{r!} K_r^{(h)}(\Delta_*^2)(q+s-2+2r)^{-1}, \tag{27}$$

(ii)

$$E^{(2-h)}[\chi_{q+s}^{*^{-4}}(\Delta_*^2)] = \sum_{r\geq 0} \frac{1}{r!} K_r^{(h)}(\Delta_*^2)(q+s-2+2r)^{-1}(q+s-4+2r)^{-1}, \tag{28}$$

and for $h = 0, 1$

$$K_r^{(h)}(\Delta_*^2) = [-2\psi'(0)]^r \left(\frac{\Delta_*^2}{2}\right)^r \int_0^\infty \frac{(t^{-1})^{-r+h}}{r!} \, \mathrm{e}^{(-t\Delta_*^2[-2\psi'(0)])/2} \mathcal{W}(t) \, \mathrm{d}t, \tag{29}$$

If $\Phi(\boldsymbol{w}'\boldsymbol{w}) = I(\boldsymbol{w}'\boldsymbol{w} < c)$, *where* $I(A)$ *is the indicator function of the set* A, *then*

(i)

$$\begin{aligned} E[\boldsymbol{w}I(\boldsymbol{w}'\boldsymbol{w} < c)] &= \boldsymbol{\eta} \int_0^\infty E_{\mathcal{N}}[I(\chi_{q+2}^2(\Delta_*^2) < c)|t] \mathcal{W}(t) \, \mathrm{d}t \\ &= \boldsymbol{\eta} \mathcal{H}_{q+2}^*(c, \Delta_*^2), \end{aligned} \tag{30}$$

(ii)

$$E[\boldsymbol{w}\boldsymbol{w}'I(\boldsymbol{w}'\boldsymbol{w} < c)] = \boldsymbol{I}_q \mathcal{H}_{q+2}^*(c, \Delta_*^2) + \boldsymbol{\eta}\boldsymbol{\eta}' \mathcal{H}_{q+4}^*(c, \Delta_*^2), \tag{31}$$

(iii)

$$E[\boldsymbol{w}'\boldsymbol{A}\boldsymbol{w}I(\boldsymbol{w}'\boldsymbol{w} < c)] = \mathrm{tr}(\boldsymbol{A}) \mathcal{H}_{q+2}^*(c, \Delta_*^2) + \boldsymbol{\eta}'\boldsymbol{A}\boldsymbol{\eta} \mathcal{H}_{q+4}^*(c, \Delta_*^2), \tag{32}$$

Further,

(i)

$$E\left[\boldsymbol{w}I\left(\frac{\boldsymbol{w}'\boldsymbol{w}}{qS^2} < c\right)\right] = \boldsymbol{\eta} G_{q+2,m}^{(2)}\left(\frac{q}{q+2}c; \Delta_*^2\right), \quad m = n-p, \tag{33}$$

(ii)

$$E\left[\boldsymbol{w}\boldsymbol{w}'I\left(\frac{\boldsymbol{w}'\boldsymbol{w}}{qS^2} < c\right)\right] = \boldsymbol{I}_q G_{q+2,m}^{(2)}\left(\frac{q}{q+2}c; \Delta_*^2\right) + \boldsymbol{\eta}\boldsymbol{\eta}' G_{q+4,m}^{(2)}\left(\frac{q}{q+2}c; \Delta_*^2\right), \tag{34}$$

(iii)

$$E\left[\boldsymbol{w}'\boldsymbol{A}\boldsymbol{w}I\left(\frac{w'w}{qS^2} < c\right)\right] = \mathrm{tr}(\boldsymbol{A}) G_{q+2,m}^{(2)}\left(\frac{q}{q+2}c; \Delta_*^2\right) + \boldsymbol{\eta}'\boldsymbol{A}\boldsymbol{\eta} G_{q+4,m}^{(2)}\left(\frac{q}{q+2}c; \Delta_*^2\right). \tag{35}$$

Further,

(i)

$$G_{q+2i,n-p}^{(2-h)}(x', \Delta_*^2) = \sum_{r=0}^\infty K_r^{(h)}(\Delta_*^2) I_{x'}\left[\frac{q+2i}{2} + r, \frac{n-p}{2}\right], \tag{36}$$

$$x' = \frac{qF_\alpha}{n-p+qF_\alpha}, \tag{37}$$

(ii)

$$
\begin{aligned}
&E^{(2-h)}\left[F_{q+s,n-p}^{-j}(\Delta_*^2) I\left(F_{q+s,n-p}(\Delta_*^2) < \frac{qd}{q+s}\right)\right] \\
&\quad = \sum_{r=0}^{\infty} K_r^{(h)}(\Delta_*^2)\left(\frac{q+s}{n-p}\right)^j \frac{B((q+s+2r-2j)/2,(n-p+2j)/2)}{B((q+s+2r)/2,(n-p)/2)} \\
&\quad \times I_x\left[\frac{q+s+2r-2j}{2}, \frac{n-p+2j}{2}\right], \qquad (38)\\
&x = \frac{qd}{n-p+qd}.
\end{aligned}
$$

In many practical situations, along with the model one may suspect that $\boldsymbol{\beta}$ belongs to the subspace defined by $\boldsymbol{H}\boldsymbol{\beta} = \boldsymbol{h}$. In such situation, one combines the estimate of $\boldsymbol{\beta}$ and the test-statistic to obtain three or more estimators as in Saleh [9], in addition to the unrestricted and the restricted estimators of $\boldsymbol{\beta}$. Now, we proceed to define three more estimators of $\boldsymbol{\beta}$ combining the unrestricted, restricted and the test-statistic $\mathcal{L}_n$ as in Saleh [9]. First, we consider the PTE of $\boldsymbol{\beta}$ which is a convex combination of $\tilde{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\beta}}$:

$$
\hat{\boldsymbol{\beta}}^{\mathrm{PT}} = \tilde{\boldsymbol{\beta}} I(\mathcal{L}_n \geq F_\alpha) + \hat{\boldsymbol{\beta}} I(\mathcal{L}_n < F_\alpha), \qquad (39)
$$

where $I(A)$ is the indicator function of the set A and F_α is the upper αth percentile of the central F-distribution with (q, m) d.f.

The PTE has the disadvantage that it depends on $\alpha (0 < \alpha < 1)$, the level of significance and also it yields the extreme results, namely $\hat{\boldsymbol{\beta}}$ and $\tilde{\boldsymbol{\beta}}$ depending on the outcome of the test. Therefore, we define Stein-type shrinkage estimator (SE) of β, as

$$
\begin{aligned}
\hat{\boldsymbol{\beta}}^{\mathrm{S}} &= \hat{\boldsymbol{\beta}} + (1 - d\mathcal{L}_n^{-1})(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}) \\
&= \tilde{\boldsymbol{\beta}} - d\mathcal{L}_n^{-1}(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}), \qquad (40)
\end{aligned}
$$

where

$$
d = \frac{(q-2)(n-p)}{q(n-p+2)} \quad \text{and} \quad q \geq 3. \qquad (41)
$$

The SE has the disadvantage that it has a strange behaviour for small values of $\mathcal{L}_n$. Also, the shrinkage factor $(1 - d\mathcal{L}_n^{-1})$ becomes negative for $\mathcal{L}_n < d$. Hence, we define a better estimator by positive-rule shrinkage estimator (PRSE) of β as

$$
\begin{aligned}
\hat{\boldsymbol{\beta}}^{\mathrm{S+}} &= \hat{\boldsymbol{\beta}} + (1 - d\mathcal{L}_n^{-1}) I[\mathcal{L}_n > d](\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}) \\
&= \hat{\boldsymbol{\beta}}^{\mathrm{S}} - (1 - d\mathcal{L}_n^{-1}) I[\mathcal{L}_n \leq d](\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}). \qquad (42)
\end{aligned}
$$

Note that this estimator is also a convex combination of $\hat{\boldsymbol{\beta}}$ and $\tilde{\boldsymbol{\beta}}$.

The biases, the quadratic risks and MSE-matrices of the estimators are given in the following section and the dominance properties are studied in Section 4.

3. Bias and quadratic risk of the estimators

Consider for a given non-singular matrix $\boldsymbol{W} \in S(p)$, the *weighted quadratic error loss function* of the form

$$\boldsymbol{L}(\boldsymbol{\beta}^*; \boldsymbol{\beta}) = (\boldsymbol{\beta}^* - \boldsymbol{\beta})'\boldsymbol{W}(\boldsymbol{\beta}^* - \boldsymbol{\beta}), \tag{43}$$

where $\boldsymbol{\beta}^*$ is any estimator of $\boldsymbol{\beta}$. Then the weighted quadratic risk function associated with Equation (43) is defined as

$$\boldsymbol{R}(\boldsymbol{\beta}^*; \boldsymbol{\beta}) = E[(\boldsymbol{\beta}^* - \boldsymbol{\beta})'\boldsymbol{W}(\boldsymbol{\beta}^* - \boldsymbol{\beta})]. \tag{44}$$

In this section, we determine the biases, and using the risk function (44), evaluate the quadratic risks and MSE matrices of the five different estimators under study. First, we consider bias expressions of the estimators.

Directly

$$\boldsymbol{b}_1 = E[\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}] = \boldsymbol{0}. \tag{45}$$

Also

$$\boldsymbol{b}_2 = E[\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}] = -\boldsymbol{\delta}. \tag{46}$$

Using Lemma 2.1, we have

$$\begin{aligned}
\boldsymbol{b}_3 &= E(\hat{\boldsymbol{\beta}}^{\mathrm{PT}} - \boldsymbol{\beta}) \\
&= E[\tilde{\boldsymbol{\beta}} - I(\mathcal{L}_n \le F_\alpha)(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}) - \boldsymbol{\beta}] \\
&= E[\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}] - E[I(\mathcal{L}_n \le F_\alpha)(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}})] \\
&= -\boldsymbol{C}\boldsymbol{H}'\boldsymbol{V}_1^{1/2}\, E[I(\mathcal{L}_n \le F_\alpha)\boldsymbol{V}_1^{1/2}(\boldsymbol{H}\tilde{\boldsymbol{\beta}} - \boldsymbol{h})] \\
&= -\boldsymbol{\delta} G^{(2)}_{q+2,m}(F_\alpha; \Delta_*^2).
\end{aligned} \tag{47}$$

Applying Lemma 2.1, we obtain

$$\begin{aligned}
\boldsymbol{b}_4 &= E(\hat{\boldsymbol{\beta}}^{\mathrm{S}} - \boldsymbol{\beta}) \\
&= E[\tilde{\boldsymbol{\beta}} - d\mathcal{L}_n^{-1}(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}) - \boldsymbol{\beta}] \\
&= -d\boldsymbol{C}^{-1}\boldsymbol{H}'\boldsymbol{V}_1^{1/2}E[\mathcal{L}_n^{-1}\boldsymbol{V}_1^{1/2}(\boldsymbol{H}\tilde{\boldsymbol{\beta}} - \boldsymbol{h})] \\
&= -dq\boldsymbol{\delta} E^{(2)}[\chi^{*^{-2}}_{q+2}(\Delta_*^2)],
\end{aligned} \tag{48}$$

Finally, by making use of Lemma 2.1 once more, we obtain

$$\begin{aligned}
\boldsymbol{b}_5 &= E(\hat{\boldsymbol{\beta}}^{\mathrm{S+}} - \boldsymbol{\beta}) \\
&= E(\hat{\boldsymbol{\beta}}^{\mathrm{S}} - \boldsymbol{\beta}) - E[I(\mathcal{L}_n \le d)(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}})] + dE[\mathcal{L}_n^{-1}I(\mathcal{L}_n \le d)(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}})] \\
&= -dq\boldsymbol{\delta} E^{(2)}[\chi^{*^{-4}}_{q+2}(\Delta_*^2)] + \boldsymbol{\delta} G^{(2)}_{q+2,m}(d; \Delta_*^2) \\
&\quad + \frac{qd}{q+2}\boldsymbol{\delta} E^{(2)}\left[F^{-1}_{q+2,m}(\Delta_*^2) I\left(F_{q+2,m}(\Delta_*^2) \le \frac{qd}{q+2}\right)\right].
\end{aligned} \tag{49}$$

Note that as the non-centrality parameter $\Delta_*^2 \to \infty$, $\boldsymbol{b}_1 = \boldsymbol{b}_3 = \boldsymbol{b}_4 = \boldsymbol{b}_5 = \boldsymbol{0}$ while $\boldsymbol{b}_2$ becomes unbounded. However, under $H_0 : \boldsymbol{H}\boldsymbol{\beta} = \boldsymbol{h}$, because $\boldsymbol{\delta} = \boldsymbol{0}$, $\boldsymbol{b}_1 = \boldsymbol{b}_2 = \boldsymbol{b}_3 = \boldsymbol{b}_4 = \boldsymbol{b}_5 = \boldsymbol{0}$.

For the risks of the estimators, considering quadratic risk function given by (3.2), we obtain

$$\begin{aligned} \boldsymbol{R}(\tilde{\boldsymbol{\beta}};\boldsymbol{\beta}) &= E[(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})'\boldsymbol{W}(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})] \\ &= \sigma_{\epsilon}^2\,\mathrm{tr}(\boldsymbol{C}^{-1}\boldsymbol{W}). \end{aligned} \tag{50}$$

Also

$$\begin{aligned} \boldsymbol{R}(\hat{\boldsymbol{\beta}};\boldsymbol{\beta}) &= E[(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})'\boldsymbol{W}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})] \\ &= \sigma_{\epsilon}^2\,\mathrm{tr}(\boldsymbol{V}_2\boldsymbol{W}) + \boldsymbol{\delta}'\boldsymbol{W}\boldsymbol{\delta} \\ &= R(\tilde{\boldsymbol{\beta}};\boldsymbol{\beta}) - \sigma_{\epsilon}^2\,\mathrm{tr}[\boldsymbol{W}\boldsymbol{C}\boldsymbol{H}'\boldsymbol{V}_1\boldsymbol{H}\boldsymbol{C}] + \boldsymbol{\delta}'\boldsymbol{W}\boldsymbol{\delta}. \end{aligned} \tag{51}$$

Note that $\boldsymbol{R} = \boldsymbol{C}_1^{-1/2}\boldsymbol{H}'\boldsymbol{V}_1\boldsymbol{H}\boldsymbol{C}_1^{-1/2}$ is a symmetric idempotent matrix of rank $q \le p$. Thus, there exist an orthogonal matrix $\boldsymbol{Q}$ (see [15]) such that

$$\boldsymbol{Q}\boldsymbol{R}\boldsymbol{Q}' = \begin{bmatrix} \boldsymbol{I}_q & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}, \quad \boldsymbol{Q}\boldsymbol{C}_1^{-1/2}\boldsymbol{W}\boldsymbol{C}_1^{-1/2}\boldsymbol{Q}' = \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{bmatrix} = \boldsymbol{A}. \tag{52}$$

The matrices $\boldsymbol{A}_{11}$ and $\boldsymbol{A}_{22}$ are of order q and $p-q$, respectively.

Now we define random variable

$$\boldsymbol{w} = \boldsymbol{Q}\boldsymbol{C}_1^{1/2}\tilde{\boldsymbol{\beta}} - \boldsymbol{Q}\boldsymbol{C}_1^{-1/2}\boldsymbol{H}'\boldsymbol{V}_1\boldsymbol{h}, \tag{53}$$

then, $\boldsymbol{w} \sim \mathcal{E}_p(\boldsymbol{\eta}, \sigma_{\epsilon}^2\boldsymbol{I}_p, g)$, where

$$\boldsymbol{\eta} = \boldsymbol{Q}\boldsymbol{C}_1^{1/2}\boldsymbol{\beta} - \boldsymbol{Q}\boldsymbol{C}_1^{-1/2}\boldsymbol{H}'\boldsymbol{V}_1\boldsymbol{h}. \tag{54}$$

Partitioning the vectors $\boldsymbol{w} = (\boldsymbol{w}_1', \boldsymbol{w}_2')'$ and $\boldsymbol{\eta} = (\boldsymbol{\eta}_1', \boldsymbol{\eta}_2')'$ where $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$ are sub-vectors of order q and $p-q$, respectively, we obtain

$$\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta} = \boldsymbol{C}_1^{-1/2}\boldsymbol{Q}'(\boldsymbol{w}-\boldsymbol{\eta}). \tag{55}$$

Thus, we can rewrite

$$\mathcal{L}_n = \frac{\boldsymbol{w}_1'\boldsymbol{w}_1}{qS^2}, \quad \boldsymbol{\eta}_1'\boldsymbol{\eta}_1 = \theta = (\boldsymbol{H}\boldsymbol{\beta}-\boldsymbol{h})'\boldsymbol{V}_1(\boldsymbol{H}\boldsymbol{\beta}-\boldsymbol{h}). \tag{56}$$

Also, we obtain

$$\begin{aligned} \mathrm{tr}\{\boldsymbol{W}[\boldsymbol{C}\boldsymbol{H}'\boldsymbol{V}_1\boldsymbol{H}\boldsymbol{C}]\} &= \mathrm{tr}\{\boldsymbol{Q}\boldsymbol{C}_1^{-1/2}\boldsymbol{W}\boldsymbol{C}_1^{-1/2}\boldsymbol{Q}'\boldsymbol{Q}\boldsymbol{R}\boldsymbol{Q}'\} \\ &= \mathrm{tr}\left(\begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{bmatrix}\begin{bmatrix} \boldsymbol{I}_q & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}\right) \\ &= \mathrm{tr}(\boldsymbol{A}_{11}). \end{aligned} \tag{57}$$

And

$$\begin{aligned} \boldsymbol{\delta}'\boldsymbol{W}\boldsymbol{\delta} &= (\boldsymbol{H}\boldsymbol{\beta}-\boldsymbol{h})'\boldsymbol{V}_1\boldsymbol{H}\boldsymbol{C}\boldsymbol{W}\boldsymbol{C}\boldsymbol{H}'\boldsymbol{V}_1(\boldsymbol{H}\boldsymbol{\beta}-\boldsymbol{h}) \\ &= \boldsymbol{\eta}_1'\boldsymbol{A}_{11}\boldsymbol{\eta}_1. \end{aligned} \tag{58}$$

Substituting Equations (57)–(58) in Equation (3.9) yields

$$R(\hat{\boldsymbol{\beta}};\boldsymbol{\beta}) = R(\tilde{\boldsymbol{\beta}};\boldsymbol{\beta}) - \sigma_{\epsilon}^2\mathrm{tr}(\boldsymbol{A}_{11}) + \boldsymbol{\eta}_1'\boldsymbol{A}_{11}\boldsymbol{\eta}_1. \tag{59}$$

Similarly, we obtain

$$
\begin{aligned}
R(\hat{\boldsymbol{\beta}}^{\mathrm{PT}};\boldsymbol{\beta}) &= E[(\hat{\boldsymbol{\beta}}^{\mathrm{PT}}-\boldsymbol{\beta})'\boldsymbol{W}(\hat{\boldsymbol{\beta}}^{\mathrm{PT}}-\boldsymbol{\beta})]\\
&= E\{[(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})-I(\mathcal{L}_n\le F_\alpha)(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})]'\boldsymbol{W}[(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})\\
&\quad -I(\mathcal{L}_n\le F_\alpha)(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})]\}\\
&= E[(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})'\boldsymbol{W}(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})]-2E[I(\mathcal{L}_n\le F_\alpha)(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})'\boldsymbol{W}(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})]\\
&\quad +E[I(\mathcal{L}_n\le F_\alpha)(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})'\boldsymbol{W}(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})]\\
&= R(\tilde{\boldsymbol{\beta}};\boldsymbol{\beta})-E[\boldsymbol{w}_1'\boldsymbol{A}_{11}\boldsymbol{w}_1 I(\mathcal{L}_n\le F_\alpha)]\\
&\quad -2E[\boldsymbol{w}_2'\boldsymbol{A}_{21}\boldsymbol{w}_1 I(\mathcal{L}_n\le F_\alpha)]+2\boldsymbol{\eta}_1'\boldsymbol{A}_{11}E[\boldsymbol{w}_1 I(\mathcal{L}_n\le F_\alpha)]\\
&\quad +2\boldsymbol{\eta}_2'\boldsymbol{A}_{21}E[\boldsymbol{w}_1 I(\mathcal{L}_n\le F_\alpha)].
\end{aligned}
$$

By making use of the representation given in Equation (1.4), we conclude that $\boldsymbol{w}_1 \mid t$ and $\boldsymbol{w}_2 \mid t$ are conditionally independent under normal theory, then one may write

$$
\begin{aligned}
E[\boldsymbol{w}_2'\boldsymbol{A}_{21}\boldsymbol{w}_1 I(\mathcal{L}_n\le F_\alpha)] &= \int_0^\infty \mathcal{W}(t)E_N[\boldsymbol{w}_2'\boldsymbol{A}_{21}\boldsymbol{w}_1 I(\mathcal{L}_n\le F_\alpha)]\,\mathrm{d}t\\
&= \boldsymbol{\eta}_2'\boldsymbol{A}_{21}\int_0^\infty \mathcal{W}(t)E_N[\boldsymbol{w}_1 I(\mathcal{L}_n\le F_\alpha)]\,\mathrm{d}t.
\end{aligned}
$$

Therefore, we obtain

$$
\begin{aligned}
R(\hat{\boldsymbol{\beta}}^{\mathrm{PT}};\boldsymbol{\beta}) &= R(\tilde{\boldsymbol{\beta}};\boldsymbol{\beta})-\sigma_{\boldsymbol{\epsilon}}^2\mathrm{tr}(\boldsymbol{A}_{11})G_{q+2,m}^{(1)}(F_\alpha;\Delta_*^2)\\
&\quad +\boldsymbol{\eta}_1'\boldsymbol{A}_{11}\boldsymbol{\eta}_1[2G_{q+2,m}^{(2)}(F_\alpha;\Delta_*^2)-G_{q+4,m}^{(2)}(F_\alpha;\Delta_*^2)]. \qquad (60)
\end{aligned}
$$

Now consider for every $q\ge 3$, one can obtain

$$E^{(2)}[\chi_q^{*^{-2}}(\Delta_*^2)]-E^{(2)}[\chi_{q+2}^{*^{-2}}(\Delta_*^2)] = 2E^{(2)}[\chi_{q+2}^{*^{-4}}(\Delta_*^2)], \qquad (61)$$

$$E^{(1)}[\chi_{q+2}^{*^{-2}}(\Delta_*^2)]-(q-2)E^{(1)}[\chi_{q+2}^{*^{-4}}(\Delta_*^2)] = \Delta_*^2E^{(2)}[\chi_{q+4}^{*^{-4}}(\Delta_*^2)]. \qquad (62)$$

Then, one can find

$$
\begin{aligned}
R(\hat{\boldsymbol{\beta}}^{\mathrm{S}};\boldsymbol{\beta}) &= E[(\hat{\boldsymbol{\beta}}^{\mathrm{S}}-\boldsymbol{\beta})'\boldsymbol{W}(\hat{\boldsymbol{\beta}}^{\mathrm{S}}-\boldsymbol{\beta})]\\
&= E[(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})'\boldsymbol{W}(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})]-2dE[\mathcal{L}_n^{-1}(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta})'\boldsymbol{W}(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})]\\
&\quad +d^2E[\mathcal{L}_n^{-2}(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})'\boldsymbol{W}(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})]\\
&= R(\tilde{\boldsymbol{\beta}};\boldsymbol{\beta})-2dE[\mathcal{L}_n^{-1}(\boldsymbol{w}_1'\boldsymbol{A}_{11}\boldsymbol{w}_1-\boldsymbol{\eta}_1'\boldsymbol{A}_{11}\boldsymbol{w}_1+\boldsymbol{w}_2'\boldsymbol{A}_{21}\boldsymbol{w}_1\\
&\quad -\boldsymbol{\eta}_2'\boldsymbol{A}_{21}\boldsymbol{w}_1)]+d^2E[\mathcal{L}_n^{-2}(\boldsymbol{w}_1'\boldsymbol{A}_{11}\boldsymbol{w}_1)]\\
&= R(\tilde{\boldsymbol{\beta}};\boldsymbol{\beta})-dq\sigma_{\boldsymbol{\epsilon}}^2\mathrm{tr}(\boldsymbol{A}_{11})\Bigg\{(q-2)E^{(1)}[\chi_{q+2}^{*^{-4}}(\Delta_*^2)]\\
&\quad +\left[1-\frac{(q+2)\boldsymbol{\eta}_1'\boldsymbol{A}_{11}\boldsymbol{\eta}_1}{2\sigma_{\boldsymbol{\epsilon}}^2\Delta_*^2\mathrm{tr}(\boldsymbol{A}_{11})}\right](2\Delta_*^2)E^{(2)}[\chi_{q+4}^{*^{-4}}(\Delta_*^2)]\Bigg\}. \qquad (63)
\end{aligned}
$$

Finally, the risk of PRSE is given by

$$R(\hat{\boldsymbol{\beta}}^{\mathrm{S}+};\boldsymbol{\beta}) = E[(\hat{\boldsymbol{\beta}}^{\mathrm{S}+}-\boldsymbol{\beta})'\boldsymbol{W}(\hat{\boldsymbol{\beta}}^{\mathrm{S}+}-\boldsymbol{\beta})]$$
$$= R(\hat{\boldsymbol{\beta}}^{\mathrm{S}};\boldsymbol{\beta}) + E[(1-d\mathcal{L}_n^{-1})^2 I(\mathcal{L}_n \le d)(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})'\boldsymbol{W}(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})]$$
$$- 2E[(1-d\mathcal{L}_n^{-1}) I(\mathcal{L}_n \le d)(\hat{\boldsymbol{\beta}}^{\mathrm{S}}-\boldsymbol{\beta})'\boldsymbol{W}(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})]. \tag{64}$$

But using the fact that

$$\begin{aligned}
&E[\hat{\boldsymbol{\beta}}^{\mathrm{S}}-\boldsymbol{\beta})'\boldsymbol{W}(1-\rho\mathcal{L}_n^{-1})I(\mathcal{L}_n \le d)(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})]\\
&\quad= E[(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}) + (1-d\mathcal{L}_n^{-1})(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})]'\boldsymbol{W}(1-d\mathcal{L}_n^{-1})I(\mathcal{L}_n \le d)(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})]\\
&\quad= E\{(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})'\boldsymbol{W}[(1-d\mathcal{L}_n^{-1})I(\mathcal{L}_n \le d)(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})]\}\\
&\qquad+ E[(1-d\mathcal{L}_n^{-1})^2 I(\mathcal{L}_n \le d)(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})'\boldsymbol{W}(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})],
\end{aligned}$$

we obtain

$$\begin{aligned}
R(\hat{\boldsymbol{\beta}}^{\mathrm{S}+};\boldsymbol{\beta}) &= R(\hat{\boldsymbol{\beta}}^{\mathrm{S}};\boldsymbol{\beta}) - E[(1-d\mathcal{L}_n^{-1})^2 I(\mathcal{L}_n \le d)(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})'\boldsymbol{W}(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})]\\
&\quad - 2E\{(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})'\boldsymbol{W}[(1-d\mathcal{L}_n^{-1})I(\mathcal{L}_n \le d)(\tilde{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}})]\}\\
&= R(\hat{\boldsymbol{\beta}}^{\mathrm{S}};\boldsymbol{\beta}) - E[(1-d\mathcal{L}_n^{-1})^2 I(\mathcal{L}_n \le d)\boldsymbol{w}_1'\boldsymbol{A}_{11}\boldsymbol{w}_1]\\
&\quad - 2E[(1-d\mathcal{L}_n^{-1})I(\mathcal{L}_n \le d)(\boldsymbol{w}_1'\boldsymbol{A}_{11}\boldsymbol{w}_1 - \boldsymbol{\eta}_1'\boldsymbol{A}_{11}\boldsymbol{w}_1 + \boldsymbol{w}_2'\boldsymbol{A}_{21}\boldsymbol{w}_1 - \boldsymbol{\eta}_2'\boldsymbol{A}_{21}\boldsymbol{w}_1)]\\
&= R(\hat{\boldsymbol{\beta}}^{\mathrm{S}};\boldsymbol{\beta}) - \sigma_\epsilon^2\left\{\mathrm{tr}(\boldsymbol{A}_{11})E^{(1)}\left[\left(1-\frac{qd}{q+2}F_{q+2,m}^{-1}(\Delta_*^2)\right)^2 I\left(F_{q+2,m}(\Delta_*^2) \le \frac{qd}{q+2}\right)\right]\right.\\
&\quad\left.+\frac{\boldsymbol{\eta}_1'\boldsymbol{A}_{11}\boldsymbol{\eta}_1}{\sigma_\epsilon^2}E^{(2)}\left[\left(1-\frac{qd}{q+2}F_{q+2,m}^{-1}(\Delta_*^2)\right)^2 I\left(F_{q+2,m}(\Delta_*^2) \le \frac{qd}{q+2}\right)\right]\right\}\\
&\quad - 2\boldsymbol{\eta}_1'\boldsymbol{A}_{11}\boldsymbol{\eta}_1 E^{(2)}\left[\left(1-\frac{qd}{q+2}F_{q+2,m}^{-1}(\Delta_*^2)\right) I\left(F_{q+2,m}(\Delta_*^2) \le \frac{qd}{q+2}\right)\right].
\end{aligned} \tag{65}$$

4. Comparison

Providing risk analysis of the underlying estimators with the weight matrix $\boldsymbol{W}$, from Theorem A.2.4. of Anderson [16], we have

$$\theta ch_1(\boldsymbol{A}_{11}) \le \boldsymbol{\eta}_1'\boldsymbol{A}_{11}\boldsymbol{\eta}_1 \le \theta ch_q(\boldsymbol{A}_{11}),$$

or equivalently

$$\sigma_\epsilon^2\Delta_*^2 ch_1(\boldsymbol{A}_{11}) \le \boldsymbol{\eta}_1'\boldsymbol{A}_{11}\boldsymbol{\eta}_1 \le \sigma_\epsilon^2\Delta_*^2 ch_q(\boldsymbol{A}_{11}),$$

where $ch_1(\boldsymbol{A}_{11})$ and $ch_q(\boldsymbol{A}_{11})$ are the minimum and maximum eigenvalue of $\boldsymbol{A}_{11}$, respectively, and $\Delta_*^2 = \theta/\sigma_\epsilon^2$.

Then by (3.8) and (3.17), one may easily see that

$$R(\hat{\boldsymbol{\beta}};\boldsymbol{\beta}) \geq R(\tilde{\boldsymbol{\beta}};\boldsymbol{\beta}) - \sigma_{\boldsymbol{\epsilon}}^2 \mathrm{tr}(\boldsymbol{A}_{11}) + \sigma_{\boldsymbol{\epsilon}}^2 \Delta_*^2 ch_1(\boldsymbol{A}_{11}),$$
$$R(\hat{\boldsymbol{\beta}};\boldsymbol{\beta}) \leq R(\tilde{\boldsymbol{\beta}};\boldsymbol{\beta}) - \sigma_{\boldsymbol{\epsilon}}^2 \mathrm{tr}(\boldsymbol{A}_{11}) + \sigma_{\boldsymbol{\epsilon}}^2 \Delta_*^2 ch_q(\boldsymbol{A}_{11}).$$

When Δ_*^2 is equal to zero, we have the above in equalities. Thus the restricted estimator ($\hat{\boldsymbol{\beta}}$) dominates the GLS estimator ($\tilde{\boldsymbol{\beta}}$) denoting by $\hat{\boldsymbol{\beta}} \succeq \tilde{\boldsymbol{\beta}}$ whenever

$$\Delta_*^2 \leq \frac{\mathrm{tr}(\boldsymbol{A}_{11})}{ch_q(\boldsymbol{A}_{11})}, \tag{66}$$

while $\hat{\boldsymbol{\beta}} \preceq \tilde{\boldsymbol{\beta}}$ whenever

$$\Delta_*^2 \geq \frac{\mathrm{tr}(\boldsymbol{A}_{11})}{ch_1(A_{11})}. \tag{67}$$

For $\boldsymbol{W} = \boldsymbol{C}$ because $\mathrm{tr}(\boldsymbol{A}_{11}) = q$, we conclude that $\hat{\boldsymbol{\beta}}$ performs better than $\tilde{\boldsymbol{\beta}}$ ($\hat{\boldsymbol{\beta}} \succeq \tilde{\boldsymbol{\beta}}$) in the interval $[0, q\sigma_{\boldsymbol{\epsilon}}^2)$ and worse outside this interval.

Comparing $\hat{\boldsymbol{\beta}}^{\mathrm{PT}}$ versus $\tilde{\boldsymbol{\beta}}$, using risk difference, we have

$$\begin{aligned} R(\tilde{\boldsymbol{\beta}};\boldsymbol{\beta}) - R(\hat{\boldsymbol{\beta}}^{\mathrm{PT}};\boldsymbol{\beta}) &= \sigma_{\boldsymbol{\epsilon}}^2 \mathrm{tr}(\boldsymbol{A}_{11}) G_{q+2,m}^{(1)}(F_\alpha;\Delta_*^2) \\ &\quad - \boldsymbol{\eta}_1' \boldsymbol{A}_{11} \boldsymbol{\eta}_1 [2G_{q+2,m}^{(2)}(F_\alpha;\Delta_*^2) - G_{q+4,m}^{(2)}(F_\alpha;\Delta_*^2)]. \end{aligned} \tag{68}$$

It follows that the right-hand side of Equation (68) is non-negative whenever

$$\Delta_*^2 \leq \frac{\mathrm{tr}(\boldsymbol{A}_{11})}{ch_q(\boldsymbol{A}_{11})} \times \frac{G_{q+2,m}^{(1)}(F_\alpha;\Delta_*^2)}{2G_{q+2,m}^{(2)}(F_\alpha;\Delta_*^2) - G_{q+4,m}^{(2)}(F_\alpha;\Delta_*^2)}, \tag{69}$$

and vice versa. Also Under $H_0 : \boldsymbol{H}\boldsymbol{\beta} = \boldsymbol{h}$, $\hat{\boldsymbol{\beta}}^{\mathrm{PT}} \succeq \tilde{\boldsymbol{\beta}}$.

Now we compare $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\beta}}^{\mathrm{PT}}$ by the risk difference as follows:

$$\begin{aligned} R(\hat{\boldsymbol{\beta}};\boldsymbol{\beta}) - R(\hat{\boldsymbol{\beta}}^{\mathrm{PT}};\boldsymbol{\beta}) &= -\sigma_{\boldsymbol{\epsilon}}^2 \mathrm{tr}(\boldsymbol{A}_{11})[1 - G_{q+2,m}^{(1)}(F_\alpha;\Delta_*^2)] \\ &\quad + \boldsymbol{\eta}_1' \boldsymbol{A}_{11} \boldsymbol{\eta}_1 [1 - 2G_{q+2,m}^{(2)}(F_\alpha;\Delta_*^2) + G_{q+4,m}^{(2)}(F_\alpha;\Delta_*^2)]. \end{aligned} \tag{70}$$

Thus, $\hat{\boldsymbol{\beta}} \succeq \hat{\boldsymbol{\beta}}^{\mathrm{PT}}$ whenever

$$\Delta_*^2 \leq \frac{\mathrm{tr}(\boldsymbol{A}_{11})}{ch_q(\boldsymbol{A}_{11})} \times \frac{1 - G_{q+2,m}^{(1)}(F_\alpha;\Delta_*^2)}{1 - 2G_{q+2,m}^{(2)}(F_\alpha;\Delta_*^2) + G_{q+4,m}^{(2)}(F_\alpha;\Delta_*^2)}, \tag{71}$$

and vice versa. However, under H_0, the dominance order of $\tilde{\boldsymbol{\beta}}$, $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\beta}}^{\mathrm{PT}}$ is as follows:

$$\hat{\boldsymbol{\beta}} \succeq \hat{\boldsymbol{\beta}}^{\mathrm{PT}} \succeq \tilde{\boldsymbol{\beta}}. \tag{72}$$

In order to determine the superiority of $\hat{\boldsymbol{\beta}}^{\mathrm{S}}$ to $\tilde{\boldsymbol{\beta}}$, it is enough to see that the following risk difference

$$\begin{aligned} R(\tilde{\boldsymbol{\beta}};\boldsymbol{\beta}) - R(\hat{\boldsymbol{\beta}}^{\mathrm{S}};\boldsymbol{\beta}) &= \sigma_{\boldsymbol{\epsilon}}^2 dq \mathrm{tr}(\boldsymbol{A}_{11}) \Bigg\{ (q-2) E^{(1)}[\chi_{q+2}^{*^{-4}}(\Delta_*^2)] \\ &\quad + \left[1 - \frac{(q+2)\boldsymbol{\eta}_1' \boldsymbol{A}_{11} \boldsymbol{\eta}_1}{2\sigma_{\boldsymbol{\epsilon}}^2 \Delta_*^2 \, \mathrm{tr}(\boldsymbol{A}_{11})} \right] (2\Delta_*^2) E^{(2)}[\chi_{q+4}^{*^{-4}}(\Delta_*^2)] \Bigg\} \end{aligned} \tag{73}$$

is positive for all $\mathbb{A}$ such that

$$\left\{\mathbb{A} : \frac{\mathrm{tr}(\boldsymbol{A}_{11})}{ch_q(\boldsymbol{A}_{11})} \geq \frac{q+2}{2}\right\}, \tag{74}$$

which asserts $\hat{\boldsymbol{\beta}}^{\mathrm{S}}$ uniformly dominates $\tilde{\boldsymbol{\beta}}$.

Furthermore, we show that the shrinkage factor d of the Stein-type estimator is robust with respect to $\boldsymbol{\beta}$ and the unknown mixing distribution.

THEOREM 4.1 *Consider the model* (1) *where the error-vector belongs to the ECD, $\mathcal{E}_n(\boldsymbol{0}, \sigma^2\boldsymbol{V}, g)$. Then the Stein-type shrinkage estimator, $\hat{\boldsymbol{\beta}}^{\mathrm{S}}$ of $\boldsymbol{\beta}$ given by*

$$\hat{\boldsymbol{\beta}}^{\mathrm{S}} = \tilde{\boldsymbol{\beta}} - d^*\mathcal{L}_n^{-1}(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}),$$

uniformly dominates the unrestricted estimator $\tilde{\boldsymbol{\beta}}$ with respect to the quadratic loss function given by Equation (43) *for $\boldsymbol{W} = \boldsymbol{C}$, and is minimax if and only if $0 < d^* \leq 2m/(m+2)$. The largest reduction of the risk is attained when $d^* = m/(m+2)$.*

For the proof see the appendix.

Remark 4.1 Consider the coefficient d given by (2.31). From $q \geq 3$, we obtain $0 < d = (q-2)m/q(m+2) < 2m/(m+2)$ and thus using Theorem 4.1, $\hat{\boldsymbol{\beta}}^{\mathrm{S}}$ in Equation (2.30) uniformly dominates $\tilde{\boldsymbol{\beta}}$ on the whole parameter space under a quadratic loss function.

To compare $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\beta}}^{\mathrm{S}}$, we may write

$$\begin{aligned} R(\hat{\boldsymbol{\beta}}^{\mathrm{S}};\boldsymbol{\beta}) &= R(\hat{\boldsymbol{\beta}};\boldsymbol{\beta}) + \sigma_{\boldsymbol{\epsilon}}^2\mathrm{tr}(\boldsymbol{A}_{11}) - \boldsymbol{\eta}_1'\boldsymbol{A}_{11}\boldsymbol{\eta}_1 - dq\sigma_{\boldsymbol{\epsilon}}^2\mathrm{tr}(\boldsymbol{A}_{11})\left\{(q-2)E^{(1)}[\chi_{q+2}^{*^{-4}}(\Delta_*^2)]\right. \\ &\quad \left. + \left[1 - \frac{(q+2)\boldsymbol{\eta}_1'\boldsymbol{A}_{11}\boldsymbol{\eta}_1}{2\sigma_{\boldsymbol{\epsilon}}^2\Delta_*^2\mathrm{tr}(\boldsymbol{A}_{11})}\right](2\Delta_*^2)E^{(2)}[\chi_{q+4}^{*^{-4}}(\Delta_*^2)]\right\}. \end{aligned} \tag{75}$$

Under H_0, this becomes

$$\begin{aligned} R(\hat{\boldsymbol{\beta}}^{\mathrm{S}};\boldsymbol{\beta}) &= R(\hat{\boldsymbol{\beta}};\boldsymbol{\beta}) + \sigma_{\boldsymbol{\epsilon}}^2(1-d)\mathrm{tr}(\boldsymbol{A}_{11}) \\ &\geq R(\hat{\boldsymbol{\beta}};\boldsymbol{\beta}), \end{aligned} \tag{76}$$

while

$$\begin{aligned} R(\hat{\boldsymbol{\beta}};\boldsymbol{\beta}) &= R(\tilde{\boldsymbol{\beta}};\boldsymbol{\beta}) - \sigma_{\boldsymbol{\epsilon}}^2\mathrm{tr}(\boldsymbol{A}_{11}) \\ &\leq R(\tilde{\boldsymbol{\beta}};\boldsymbol{\beta}). \end{aligned} \tag{77}$$

Therefore, $\hat{\boldsymbol{\beta}}$ performs better that $\hat{\boldsymbol{\beta}}^{\mathrm{S}}$ under H_0. However, as $\boldsymbol{\eta}_1$ moves away from $\boldsymbol{0}$, $\boldsymbol{\eta}_1'\boldsymbol{A}_{11}\boldsymbol{\eta}_1$ increases and the risk of $\hat{\boldsymbol{\beta}}$ becomes unbounded while the risk of $\tilde{\boldsymbol{\beta}}^{\mathrm{S}}$ remains below the risk of $\tilde{\boldsymbol{\beta}}$; thus, $\tilde{\boldsymbol{\beta}}^{\mathrm{S}}$ dominates $\hat{\boldsymbol{\beta}}$ outside an interval around the origin.

Consider under H_0

$$\begin{aligned} R(\hat{\boldsymbol{\beta}}^{\mathrm{S}};\boldsymbol{\beta}) &= R(\hat{\boldsymbol{\beta}}^{\mathrm{PT}};\boldsymbol{\beta}) + \sigma_{\boldsymbol{\epsilon}}^2\mathrm{tr}(\boldsymbol{A}_{11})[1-\alpha-d] \\ &\geq R(\hat{\boldsymbol{\beta}}^{\mathrm{PT}};\boldsymbol{\beta}), \end{aligned} \tag{78}$$

for all α such that $F_{q+2,m}^{-1}(d,0) \leq qF_\alpha/(q+2)$. This means the estimator $\hat{\boldsymbol{\beta}}^{\mathrm{S}}$ does not always dominates $\hat{\boldsymbol{\beta}}^{\mathrm{PT}}$ under H_0.

Thus, under H_0 with α satisfying $F^{-1}_{q+2,m}(d,0) \leq qF_\alpha/(q+2)$, we can conclude that

$$\hat{\boldsymbol{\beta}} \succeq \hat{\boldsymbol{\beta}}^{\text{PT}} \succeq \hat{\boldsymbol{\beta}}^{\text{S}} \succeq \tilde{\boldsymbol{\beta}}. \tag{79}$$

We compare, the risks of $\hat{\boldsymbol{\beta}}^{\text{S+}}$ and $\hat{\boldsymbol{\beta}}^{\text{S}}$. The risk difference is given by

$$\begin{aligned}
&R(\hat{\boldsymbol{\beta}}^{\text{S+}};\boldsymbol{\beta}) - R(\hat{\boldsymbol{\beta}}^{\text{S}};\boldsymbol{\beta}) \\
&\quad = -\sigma_\epsilon^2 \left\{ \operatorname{tr}(\boldsymbol{A}_{11}) E^{(1)} \left[\left(1 - \frac{qd}{q+2} F^{-1}_{q+2,m}(\Delta_*^2)\right)^2 I\left(F_{q+2,m}(\Delta_*^2) \leq \frac{qd}{q+2}\right) \right] \right. \\
&\qquad \left. + \frac{\boldsymbol{\eta}_1' \boldsymbol{A}_{11} \boldsymbol{\eta}_1}{\sigma_\epsilon^2} E^{(2)} \left[\left(1 - \frac{qd}{q+2} F^{-1}_{q+2,m}(\Delta_*^2)\right)^2 I\left(F_{q+2,m}(\Delta_*^2) \leq \frac{qd}{q+2}\right) \right] \right\} \\
&\qquad - 2\boldsymbol{\eta}_1' \boldsymbol{A}_{11} \boldsymbol{\eta}_1 E^{(2)} \left[\left(1 - \frac{qd}{q+2} F^{-1}_{q+2,m}(\Delta_*^2)\right) I\left(F_{q+2,m}(\Delta_*^2) \leq \frac{qd}{q+2}\right) \right].
\end{aligned}$$

The right-hand side of the above equality is negative since for $F_{q+2,m}(\Delta_*^2) \leq qd/(q+2)$, $((qd/(q+2))F_{q+2,m}(\Delta_*^2) - 1) \geq 0$ and also the expectation of a positive random variable is positive. That for all $\boldsymbol{\beta}$, $R(\hat{\boldsymbol{\beta}}^{\text{S+}};\boldsymbol{\beta}) \leq R(\hat{\boldsymbol{\beta}}^{\text{S}};\boldsymbol{\beta})$. And using Theorem 4.1, $R(\hat{\boldsymbol{\beta}}^{\text{S+}};\boldsymbol{\beta}) \leq R(\tilde{\boldsymbol{\beta}};\boldsymbol{\beta})$.

To compare $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\beta}}^{\text{S+}}$, first consider the case under H_0 i.e., $\boldsymbol{\eta}_1 = \mathbf{0}$. In this case,

$$\begin{aligned}
R(\hat{\boldsymbol{\beta}}^{\text{S+}};\boldsymbol{\beta}) &= R(\hat{\boldsymbol{\beta}};\boldsymbol{\beta}) + \sigma_\epsilon^2 \operatorname{tr}(\boldsymbol{A}_{11}) \left\{ (1-d) \right. \\
&\qquad \left. - E^{(1)} \left[\left(1 - \frac{qd}{q+2} F^{-1}_{q+2,m}(0)\right)^2 I\left(F_{q+2,m}(0) \leq \frac{qd}{q+2}\right) \right] \right\} \\
&\geq R(\hat{\boldsymbol{\beta}};\boldsymbol{\beta}),
\end{aligned} \tag{80}$$

since

$$\begin{aligned}
&E^{(1)} \left[\left(1 - \frac{qd}{q+2} F^{-1}_{q+2,m}(0)\right)^2 I\left(F_{q+2,m}(0) \leq \frac{qd}{q+2}\right) \right] \\
&\quad \leq E^{(1)} \left[\left(1 - \frac{qd}{q+2} F^{-1}_{q+2,m}(0)\right)^2 \right] = 1 - d.
\end{aligned} \tag{81}$$

Thus under H_0, $\hat{\boldsymbol{\beta}} \succeq \hat{\boldsymbol{\beta}}^{\text{S+}}$. However, as $\boldsymbol{\eta}_1$ moves away from $\mathbf{0}$, $\boldsymbol{\eta}_1' \boldsymbol{A}_{11} \boldsymbol{\eta}_1$ increases and the risk of $\hat{\boldsymbol{\beta}}$ becomes unbounded while the risk of $\tilde{\boldsymbol{\beta}}^{\text{S+}}$ remains below the risk of $\tilde{\boldsymbol{\beta}}$; thus $\tilde{\boldsymbol{\beta}}^{\text{S+}}$ dominates $\hat{\boldsymbol{\beta}}$ outside an interval around the origin.

Now, we compare $\hat{\boldsymbol{\beta}}^{\text{S+}}$ and $\hat{\boldsymbol{\beta}}^{\text{PT}}$. When H_0 holds, because $G^*_{q+2,m}(F_\alpha, 0) = 1 - \alpha$,

$$\begin{aligned}
R(\hat{\boldsymbol{\beta}}^{\text{S+}};\boldsymbol{\beta}) &= R(\hat{\boldsymbol{\beta}}^{\text{PT}};\boldsymbol{\beta}) + \sigma_\epsilon^2 \operatorname{tr}(\boldsymbol{A}_{11}) \left\{ 1 - \alpha - d \right. \\
&\qquad \left. - E^{(1)} \left[\left(1 - \frac{qd}{q+2} F^{-1}_{q+2,m}(0)\right)^2 I\left(F_{q+2,m}(0) \leq \frac{qd}{q+2}\right) \right] \right\} \\
&\geq R(\hat{\boldsymbol{\beta}}^{\text{PT}};\boldsymbol{\beta}),
\end{aligned} \tag{82}$$

for all α satisfying

$$E^{(1)}\left[\left(1-\frac{qd}{q+2}F^{-1}_{q+2,m}(0)\right)^2 I\left(F_{q+2,m}(0)\leq\frac{qd}{q+2}\right)\right]\leq 1-\alpha-d. \tag{83}$$

Thus, $\hat{\boldsymbol{\beta}}^{\mathrm{S+}}$ does not always dominates $\hat{\boldsymbol{\beta}}^{\mathrm{PT}}$ when the null-hypothesis H_0 holds.

Since always $\hat{\boldsymbol{\beta}}^{\mathrm{S+}} \succeq \hat{\boldsymbol{\beta}}^{\mathrm{S}} \succeq \tilde{\boldsymbol{\beta}}$, and under H_0, the restricted estimator $\hat{\boldsymbol{\beta}}$ performs better that all others, the dominance order of the five estimators under the null hypothesis H_0, can be determined under the following two categories:

(1) $\hat{\boldsymbol{\beta}} \succeq \hat{\boldsymbol{\beta}}^{\mathrm{PT}} \succeq \hat{\boldsymbol{\beta}}^{\mathrm{S+}} \succeq \hat{\boldsymbol{\beta}}^{\mathrm{S}} \succeq \tilde{\boldsymbol{\beta}}$, satisfying (4.19),
(2) $\hat{\boldsymbol{\beta}} \succeq \hat{\boldsymbol{\beta}}^{\mathrm{S+}} \succeq \hat{\boldsymbol{\beta}}^{\mathrm{S}} \succeq \hat{\boldsymbol{\beta}}^{\mathrm{PT}} \succeq \tilde{\boldsymbol{\beta}}$, $\forall\alpha \ni F^{-1}_{q+2,m}(d,0) > qF_\alpha/(q+2)$.

5. Concluding remarks

In this paper, we proposed five different estimators for the regression parameters of a linear regression model. In this approach, the prior non-sample information $\boldsymbol{H\beta}=\boldsymbol{h}$ is suspected. Based on the two ordering superiority categories of the estimators, under the constraint $\boldsymbol{H\beta}=\boldsymbol{h}$, two non-linear estimators $\hat{\boldsymbol{\beta}}^{\mathrm{S+}}$ and $\hat{\boldsymbol{\beta}}^{\mathrm{S}}$ perform better than unbiased GLSE.

The behavior of Stein-type estimators is restricted by the condition $q \geq 3$. However, The PTE requires the size of testing $H_0 : \boldsymbol{H\beta}=\boldsymbol{h}$. For $\mathcal{W}(t)$ as a dirac delta function, the maximal savings in risk for the shrinkage estimator is $m(q-2)/p(m+2)$, while for $\mathcal{W}(t)$ as inverse-gamma function, it is equal to $(m(q-2)/p(m+2))(\nu-2)/\nu$, where ν is d.f. of multivariate student's t-distribution.

Remarkably, the behavior of all estimators comparing with each others under elliptical symmetry are exactly the same as under normal theory as exhibited in Saleh [9]. This phenomenon shows the dominance order of estimators and regarding substantial conditions under normal theory are significantly robust.

Another noteworthy fact is that under the subclass (c) of elliptical models we could present fundamental Theorem 4.1 under elliptical models rather than scale mixture of normal distributions. It is important to point out that just under the signed measure $\mathcal{W}(t)$ under subclass (c) we are able to prove Theorem 4.1 even for non-positive measures. See, Srivastava and Bilodeau [4] for more discussion in this regard.

Finally, if ν_1 remains constant, the distribution of F_{ν_1,ν_2} tends to that of $\chi^2_{\nu_1}/\nu_1$ as ν_2 tends to infinity (see [17]).

Now consider the class of local alternatives $\{K_{(n)}\}$ defined by

$$K_{(n)} : \boldsymbol{H\beta}=\boldsymbol{h}+n^{-1/2}\boldsymbol{\xi}. \tag{84}$$

Furthermore, following Saleh [9], consider the following regularity conditions hold:

(i) $\max_{1\leq i\leq n}\boldsymbol{x}_i'(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{x}_i \to 0$ as $n\to\infty$
where $\boldsymbol{x}_i'$ is the ith row of $\boldsymbol{X}$;
(ii) $\lim_{n\to\infty}\{n^{-1}(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})\}=\boldsymbol{C}$

for finite $\boldsymbol{C}\in S(p)$.

Then using Theorem 7.8.3 from Saleh [9] in addition to Theorem 2.1 we obtain the following important result for the test statistic

$$\lim_{n\to\infty} P(\mathcal{L}_n \leq x) = \mathcal{H}_q^*(x;\delta^2).$$

Based on the above results, one can easily obtain the asymptotic distributional bias, risk and MSE matrix of each estimator under study using the following definition:

$$G_p(\boldsymbol{x}) = \lim_{n\to\infty} P_{K_{(n)}}\{\sqrt{n}S^{-2}(\boldsymbol{\beta}^* - \boldsymbol{\beta}) \leq \boldsymbol{x}\}.$$

Then

$$\boldsymbol{b}(\boldsymbol{\beta}^*) = \int \boldsymbol{x} dG_p(\boldsymbol{x}), \quad \boldsymbol{M}(\boldsymbol{\beta}^*) = \int \boldsymbol{x}\boldsymbol{x}' dG_p(\boldsymbol{x}), \quad R(\boldsymbol{\beta}^*;\boldsymbol{\beta}) = \mathrm{tr}[\boldsymbol{W}\boldsymbol{M}(\boldsymbol{\beta}^*)],$$

which have similar notation to those are given in this paper.

Acknowledgements

The authors would like to thank an anonymous referee for his/her really helpful suggestions which led us to put many details in the paper and pointing out some mistakes in the earlier version of the manuscript. This research was supported by a grant from Ferdowsi University of Mashhad; (No. MS89184TAB).

References

[1] K.T. Fang, S. Kotz, and K.W. Ng, *Symmetric Multivariate and Related Distributions*, Chapman and Hall, London, 1990.
[2] K.C. Chu, *Estimation and decision for linear systems with elliptically random process*, IEEE Trans. Autom. Cont. 18 (1973), pp. 499–505.
[3] L. Debnath and D. Bhatta, *Integral Transforms and Their Applications*, Chapman and Hall, London, 2007.
[4] M. Srivastava and M. Bilodeau, *Stein estimation under elliptical distribution*, J. Mult. Annal. 28 (1989), pp. 247–259.
[5] D. Kelker, *Ditribution theory of spherical distributions and location-scale parameter generalization*, Sankhya 32 (1970), pp. 419–430.
[6] S. Cambanis, S. Huang, and G. Simons, *On the theory of elliptically contoured distributions*, J. Mult. Anal. 11 (1981), pp. 368–385.
[7] R.J. Muirhead, *Aspect of Multivariate Statistical Theory*, Wiley, New York, 1982.
[8] A.K. Gupta and T. Varga, *Elliptically Contoured Models in Statistics*, Kluwer Academic Press, Dordrecht, 1993.
[9] A.K.Md.E. Saleh, *Theory of Preliminary Test and Stein-type Estimation with Applications*, Wiley, New York, 2006.
[10] S.M.M. Tabatabaey, *Preliminary Test Approach Estimation: Regression Model with Spherically Symmetric Errors*, Unpublished Ph.D. Thesis, Carleton University, Canada, 1995.
[11] M. Arashi and S.M.M. Tabatabaey, *Stein-type improvement under stochastic constraints: Use of multivariate Student-t model in regression*, Statist. Probab. Lett. 78 (2008), pp. 2142–2153.
[12] M. Arashi and S.M.M. Tabatabaey, *A note on classical Stein-type estimators in elliptically contoured models*, J. Statist. Plann. Inference 140 (2010), pp. 1206–1213.
[13] T.W. Anderson, K.T. Fang, and H. Hsu, *Maximum-likelihood estimates and likelihood-ratio criteria for multivariate elliptically contoured distributions*, Canad. J. Statist. 14 (1986), pp. 55–59.
[14] G.G. Judge and M.E. Bock, *The Statistical Implication of Pre-Test and Stein-Rule Estimators in Econometrics*, North-Holland, New York, 1978.
[15] S.R. Searle, *Matrix Algebra Useful for Statistics*, Wiley, New York, 1982.
[16] T.W. Anderson, *An Introduction to Multivariate Statistical Analysis*, 3rd ed., Wiley, New York, 2003.
[17] N.L. Johnson and S. Kotz, *Continuous Univariate Distributions-2*, Wiley, New York, 1970.

Appendix

Proof of Theorem 2.1 The likelihood ratio is given by

$$\begin{aligned}
\lambda &= \frac{\max_{\omega} L(\boldsymbol{y})}{\max_{\Omega} L(\boldsymbol{y})} \\
&= \frac{d_n|\hat{\sigma}^2\boldsymbol{V}|^{-1/2}\max_y g[((\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})'\boldsymbol{V}^{-1}(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}}))/2\hat{\sigma}^2]}{d_n|\tilde{\sigma}^2\boldsymbol{V}|^{-1/2}\max_y g[((\boldsymbol{y}-\boldsymbol{X}\tilde{\boldsymbol{\beta}})'\boldsymbol{V}^{-1}(\boldsymbol{y}-\boldsymbol{X}\tilde{\boldsymbol{\beta}}))/2\tilde{\sigma}^2]} \\
&= \left(\frac{\tilde{\sigma}^2}{\hat{\sigma}^2}\right)^{n/2}\frac{g(y_g)}{g(y_g)} \\
&= \left(\frac{1}{1+((\boldsymbol{H}\tilde{\boldsymbol{\beta}}-\boldsymbol{h})'(\boldsymbol{H}\boldsymbol{C}^{-1}\boldsymbol{H}')^{-1}(\boldsymbol{H}\tilde{\boldsymbol{\beta}}-\boldsymbol{h}))/((\boldsymbol{y}-\boldsymbol{X}\tilde{\boldsymbol{\beta}})'\boldsymbol{V}^{-1}(\boldsymbol{y}-\boldsymbol{X}\tilde{\boldsymbol{\beta}}))}\right)^{n/2}.
\end{aligned}$$

Therefore,

$$\lambda^{2/n} = \left(\frac{1}{1+q/(n-p))\mathcal{L}_n}\right),$$

which is decreasing with respect to $\mathcal{L}_n$. Now consider that

$$\begin{aligned}
(n-p)S^2|t &= (\boldsymbol{y}-\boldsymbol{X}\tilde{\boldsymbol{\beta}})'\boldsymbol{V}^{-1}(\boldsymbol{y}-\boldsymbol{X}\tilde{\boldsymbol{\beta}})|t \\
&= \boldsymbol{y}'[(\boldsymbol{V}^{-1}-\boldsymbol{V}^{-1}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{V}^{-1})]\boldsymbol{y}|t \sim \chi^2_{n-p}
\end{aligned}$$

also $(\boldsymbol{H}\boldsymbol{C}^{-1}\boldsymbol{H}')^{-1/2}(\boldsymbol{H}\tilde{\boldsymbol{\beta}})|t \sim \mathcal{N}_q((\boldsymbol{H}\boldsymbol{C}^{-1}\boldsymbol{H}')^{-1/2}(\boldsymbol{H}\boldsymbol{\beta}-\boldsymbol{h}), t^{-1}\sigma^2\boldsymbol{I}_q)$. Then

$$\tilde{\boldsymbol{\beta}}'\boldsymbol{H}'(\boldsymbol{H}\boldsymbol{C}^{-1}\boldsymbol{H}')^{-1}\boldsymbol{H}\tilde{\boldsymbol{\beta}}|t \sim \chi^2_{q,\Delta_t^2},$$

where $\Delta_t^2 = t\theta/\sigma^2$. Also using the fact that $(\boldsymbol{y}-\boldsymbol{X}\tilde{\boldsymbol{\beta}})'\boldsymbol{V}^{-1}(\boldsymbol{y}-\boldsymbol{X}\tilde{\boldsymbol{\beta}})|t$ and $\tilde{\boldsymbol{\beta}}'\boldsymbol{H}'(\boldsymbol{H}\boldsymbol{C}^{-1}\boldsymbol{H}')^{-1}\boldsymbol{H}\tilde{\boldsymbol{\beta}}|t$ are independent, we get

$$\mathcal{L}_n|t = \frac{\tilde{\boldsymbol{\beta}}'\boldsymbol{H}'(\boldsymbol{H}\boldsymbol{C}^{-1}\boldsymbol{H}')^{-1}\boldsymbol{H}\tilde{\boldsymbol{\beta}}|t}{qS^2|t} \sim F_{q,n-p,\Delta_t^2}.$$

Hence

$$\begin{aligned}
\boldsymbol{g}^*_{q,m}(\mathcal{L}_n) &= \int_0^\infty \mathcal{W}(t)F_{q,m,\Delta_t^2}(\mathcal{L}_n|t)\,\mathrm{d}t \\
&= \int_0^\infty \mathcal{W}(t)\sum_{r=0}^\infty \frac{\mathrm{e}^{-\Delta_t^2/2}(\Delta_t^2/2)^r}{\Gamma(r+1)}\left(\frac{q}{m}\right)^{q/2+r} \\
&\quad \times \frac{\mathcal{L}_n^{q/2+r-1}}{B((q/2)+r,m/2)(1+(q/m)\mathcal{L}_n)^{(q+m)/2+r}}\,dt \\
&= \sum_{r=0}^\infty \left(\frac{q}{m}\right)^{q/2+r}\frac{\mathcal{L}_n^{q/2+r-1}K_r^{(0)}(\Delta_*^2)}{B(q/2+r,m/2)(1+(q/m)\mathcal{L}_n)^{(q+m)/2+r}}.
\end{aligned}$$

■

Proof of Theorem 4.1 By making use of $\dot{\boldsymbol{z}} = \boldsymbol{H}'\boldsymbol{V}_1(\boldsymbol{H}\tilde{\boldsymbol{\beta}}-\boldsymbol{h})$, the SE can be rewritten as

$$\begin{aligned}
\hat{\boldsymbol{\beta}}^{\mathrm{S}} &= \tilde{\boldsymbol{\beta}} - qd^*S^2[(\boldsymbol{H}\tilde{\boldsymbol{\beta}}-\boldsymbol{h})'\boldsymbol{V}_1(\boldsymbol{H}\tilde{\boldsymbol{\beta}}-\boldsymbol{h})]^{-1}\boldsymbol{C}^{-1}\boldsymbol{H}'\boldsymbol{V}_1(\boldsymbol{H}\tilde{\boldsymbol{\beta}}-\boldsymbol{h}) \\
&= \tilde{\boldsymbol{\beta}} - qd^*S^2(\dot{\boldsymbol{z}}'\boldsymbol{C}^{-1}\dot{\boldsymbol{z}})^{-1}\boldsymbol{C}^{-1}\dot{\boldsymbol{z}}.
\end{aligned}$$

Then, the risk difference of the SE and the UE under quadratic loss function, is given by

$$\begin{aligned}
\boldsymbol{D}_4 &= E(\hat{\boldsymbol{\beta}}^{\mathrm{S}} - \boldsymbol{\beta})'\boldsymbol{C}(\hat{\boldsymbol{\beta}}^{\mathrm{S}} - \boldsymbol{\beta}) - E(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})'\boldsymbol{C}(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}) \\
&= (d^*)^2 E[q^2 S^4 (\dot{\boldsymbol{z}}'\boldsymbol{C}^{-1}\dot{\boldsymbol{z}})^{-1}] - 2d^* E[qS^2(\dot{\boldsymbol{z}}'\boldsymbol{C}^{-1}\dot{\boldsymbol{z}})^{-1}(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})'\dot{\boldsymbol{z}}] \\
&= (d^*)^2 E_t\{E_N[q^2 S^4(\dot{\boldsymbol{z}}'\boldsymbol{C}^{-1}\dot{\boldsymbol{z}})^{-1}|t]\} \\
&\quad - 2d^* E_t\{E_N[qS^2(\dot{\boldsymbol{z}}'\boldsymbol{C}^{-1}\dot{\boldsymbol{z}})^{-1}(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})'\boldsymbol{H}'\boldsymbol{V}_1(\boldsymbol{H}\tilde{\boldsymbol{\beta}} - \boldsymbol{h})|t]\} \\
&= \frac{q^2(m+2)}{m}(d^*)^2 E_t\left(\frac{\tau^{-2}}{\dot{\boldsymbol{z}}'\boldsymbol{C}^{-1}\dot{\boldsymbol{z}}}\right) - 2q^2 d^* E_t\left(\frac{t^{-2}}{\dot{\boldsymbol{z}}'\boldsymbol{C}^{-1}\dot{\boldsymbol{z}}}\right),
\end{aligned}$$

since

$$\left(\frac{mS^2}{\sigma^2}\right)|t \sim t^{-1}\chi_m^2 \quad \text{and} \quad \tilde{\boldsymbol{\beta}}'\boldsymbol{H}'\boldsymbol{V}_1\boldsymbol{H}\tilde{\boldsymbol{\beta}} \mid t \sim t^{-2}\sigma^4\chi_q^2(\dot{\boldsymbol{\delta}}),$$

where $\dot{\boldsymbol{\delta}} = \boldsymbol{\beta}'\boldsymbol{H}'\boldsymbol{V}_1\boldsymbol{H}\boldsymbol{\beta}$.

Therefore, $\boldsymbol{D}_4 \leq 0$ if and only if $0 < d^* \leq 2m/(m+2)$ since $\int_0^\infty (t^{-2}/\dot{\boldsymbol{z}}'\boldsymbol{C}^{-1}\dot{\boldsymbol{z}})\, d\mathcal{W}(t) > 0$ (see [4]). ■