Improved preliminary test and Stein-rule Liu estimators for the ill-conditioned elliptical linear regression model

M. Arashi\textsuperscript{a}, B.M. Golam Kibria\textsuperscript{b,∗}, M. Norouzirad\textsuperscript{a}, S. Nadarajah\textsuperscript{c}

\textsuperscript{a} Department of Statistics, School of Mathematical Sciences, Shahrood University of Technology, Shahrood, Iran
\textsuperscript{b} Department of Mathematics and Statistics, Florida International University, Miami, FL 33199, USA
\textsuperscript{c} Department of Mathematics, University of Manchester, Manchester, United Kingdom

\begin{abstract}
Recently, Liu (1993) estimator draws an important attention to estimate the regression parameters for an ill-conditioned linear regression model when the vector of errors is distributed according to the law belonging to the class of elliptically contoured distributions (ECDs). This paper proposed some improved Liu type estimators, namely, the unrestricted Liu estimator (ULE), restricted Liu estimator (RLE), preliminary test Liu estimator (PTLE), shrinkage Liu estimator (SLE) and positive rule Liu estimator (PRLE) for estimating the regression parameters $\beta$. The performance of the proposed estimators is compared based on the quadratic bias and risk functions under both null and alternative hypotheses, which specify certain restrictions on the regression parameters. The conditions of superiority of the proposed estimators for parameter $d$ and non-centrality parameter $\Delta$ are given.

© 2014 Elsevier Inc. All rights reserved.
\end{abstract}

\section{Introduction}
\subsection{Literature review}

It is a general practice for the researchers to estimate the regression parameters by using observed data alone. Nevertheless, the inclusion of the prior information in the estimation process may improve the quality of the estimators in the sense of smaller quadratic risk \cite{45,48}. It is well known that the estimators with the prior information (called the restricted estimator, RE) perform better than the estimators with no prior information (called the unrestricted estimator, UE). However, when the prior information is doubtful (or not sure), one may combine the restricted and unrestricted estimators to obtain a better performance of the estimator, which leads to the preliminary test’s least squares estimator. The preliminary test approach estimation under the Gaussian assumption has been pioneered by Bancroft \cite{12}, followed by Bancroft \cite{13}, Han and Bancroft \cite{26}, Judge and Bock \cite{30}, Giles \cite{21}, Benda \cite{14}, Saleh \cite{43}, and very recently Kibria and Saleh \cite{35} among others. Note that, the preliminary test estimator (PTE) 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 UE and RE? To overcome this shortcoming, one may consider the

\begin{itemize}
\item [∗] Corresponding author.
\item E-mail addresses: m_arashi_stat@yahoo.com (M. Arashi), kibriag@fiu.edu (B.M.G. Kibria), mina.norouzirad@gmail.com (M. Norouzirad), Saralees.Nadarajah@manchester.ac.uk (S. Nadarajah).
\end{itemize}
alternative choices to PTE, namely the Stein-type shrinkage estimator which incorporates the uncertain prior information and combines the restricted and unrestricted estimators in a superior manner. The properties of Stein-type estimators for the linear regression model have been discussed under normal assumption by various researchers. To mention a few, James and Stein [29], Judge and Bock [30], Saleh and Sen [46,45], Ohtani [42] and Saleh [43] among others.

In practice, most of the researchers assumed that the error variables of the regression model are normally and independently distributed. However, such assumptions may or may not be valid in many practical situations (see [22,60]). It happens particularly if the error distribution has heavier tails. For instance, some economic data may be generated by processes whose distribution has more kurtosis than the normal distribution. The multivariate Student t distribution can overcome both the problems of outliers and dependent but uncorrelated data. The shrinkage estimation under the multivariate t is necessary and has been considered by different researchers: Singh [50,51], Tabatabaey et al. [54], to mention a few. Since, the elliptically contoured distribution contains a lot of distributions, shrinkage estimators for the elliptically contoured error distribution would be a valuable asset for the researchers of this topic. The shrinkage estimation for the linear regression model with elliptically contoured error distribution is considered by Arashi et al. [9], Arashi and Tabatabaey [10,11] and the others. The shrinkage estimation for the ill conditioned model is limited and has received less attention. The objective of this paper is to propose some shrinkage type [37] estimators for the ill conditioned linear regression model.

1.2. Model and some preliminaries

We consider the following linear model
\[ y = X \beta + e, \]  
where \( y = (y_1, y_2, \ldots, y_n) \) is an \( n \times 1 \) vector of observations on the dependent variable, \( X \) is an \( n \times p \) matrix of full rank \( p \), \( \beta = (\beta_1, \beta_2, \ldots, \beta_p) \) is a \( p \times 1 \) vector of unknown parameters and \( e = (e_1, e_2, \ldots, e_n) \) is an \( n \times 1 \) vector of errors, which is distributed according to the law belonging to the class of elliptically contoured distributions (ECDs), \( E_n(0, \sigma^2 V, \psi) \) for an un-structured known matrix \( V \) with the following characteristic function
\[ \phi_n(t) = \psi(\sigma^2 t' V t) \]
for some functions \( \psi : [0, \infty) \rightarrow \mathbb{R}^+ \) say characteristic generator [19]. The details on the ECD class distributions are provided in Appendix A.

For the full model, the unrestricted estimator (UE) of \( \beta \) is given by
\[ \hat{\beta}^{\text{UE}} = C^{-1} X' V^{-1} y, \]  
where \( C = X' X \) is the information matrix. The corresponding unbiased estimator of \( \sigma^2 \) is given by
\[ \hat{\sigma}^2_{\text{UE}} = \frac{(y - X \hat{\beta}^{\text{UE}})' V^{-1} (y - X \hat{\beta}^{\text{UE}})}{m}, \quad m = n - p. \]

Our primary interest is to estimate the regression parameters \( \beta \) when it is a priori suspected but not certain that \( \beta \) may be restricted to the subspace
\[ H_0 : H \beta = h, \]  
where \( H \) is a \( q \times p \) known matrix of full rank \( q \) (\(< p \) and \( h \) is a \( q \times 1 \) vector of known constants.

The restricted estimator (RE) of \( \beta \) is given by
\[ \hat{\beta}^{\text{RE}} = \hat{\beta}^{\text{UE}} - C^{-1} H' (HC^{-1} H')^{-1} (H \hat{\beta}^{\text{UE}} - h) \]  
and the corresponding estimator of \( \sigma^2 \) is given by
\[ \hat{\sigma}^2_{\text{RE}} = \frac{(y - X \hat{\beta}^{\text{RE}})' V^{-1} (y - X \hat{\beta}^{\text{RE}})}{m + q}; \quad m = n - p \]
which is unbiased under the null hypothesis \( H_0 \). Note that the restricted least squares estimator satisfies the condition \( H \hat{\beta}^{\text{RE}} = h \).

The estimator of \( \beta \) in (1.2) is usually used in the case when there is no hypothesis information available on the vector of parameter of interest \( \beta \). On the other hand, the estimator of \( \beta \) in (1.4) is useful in the presence of hypothesis (1.3). As a result, one may combine the UE and RE to obtain the preliminary test estimator (PT) of \( \beta \) as
\[ \hat{\beta}^{\text{PT}} = \hat{\beta}^{\text{RE}} I(\mathcal{L}_n \leq \mathcal{L}_{n,a}) + \hat{\beta}^{\text{UE}} I(\mathcal{L}_n > \mathcal{L}_{n,a}) \]
\[ = \hat{\beta}^{\text{UE}} - (\hat{\beta}^{\text{UE}} - \hat{\beta}^{\text{RE}}) I(\mathcal{L}_n \leq \mathcal{L}_{n,a}). \]  
where using Theorem 3.1 of [27]
\[ \mathcal{L}_n = \frac{(H \hat{\beta}^{\text{UE}} - h)' (HC^{-1} H')^{-1} (H \hat{\beta}^{\text{UE}} - h)}{q \hat{\sigma}^2_{\text{UE}}}. \]
is the general test-statistic for testing the null-hypothesis in (1.3), and \( L_{n,\alpha} \) is the upper \( \alpha \)-level critical value of \( L_n \) and \( I(A) \) is the indicator function of the set \( A \). Under the null hypothesis and the normality assumption, \( L_n \) follows a central \( F \)-distribution with \((q, m)\) degrees of freedom while under the alternative, it follows the generalized non-central \( F \)-distribution with pdf

\[
g_{q,m}(L_n) = \sum_{r \geq 0} \frac{(q)_{r} (m/2)^{r} \Gamma(q+2r) \Gamma(m/2)}{r! B(q+2r, m/2) (1 + q m L_n)^{q+2r+1/2}},
\]

and \((q, m)\) degrees of freedom and non-centrality parameter \( \Delta \), say departure parameter where

\[
\Delta = \frac{(H \beta - h)'(H^{-1}H')^{-1}(H \beta - h)}{\sigma^2_e}; \quad \sigma^2_e = \frac{v}{v - 2} \sigma^2
\]

and the mixing distribution,

\[
K^{(0)}(\Delta) = \left(-2\psi'(0)\right)^{t} \left(\frac{\Delta}{2}\right)^{t} \int_{0}^{\infty} t^{t - 1} e^{-\frac{-1\Delta - 2\psi'(0)}{2}} W(t) dt.
\]

Notice that \( \hat{\beta}^{PT} \) is bounded and performs better than \( \hat{\beta}^{UE} \) and \( \hat{\beta}^{RE} \) in some part of the parameter space.

The preliminary test estimator (PT) has two extreme choices, namely, the unrestricted estimator and the restricted estimator. A compromise approach can be suggested by using the Stein-type [29] shrinkage estimator (SE) of \( \beta \) as

\[
\hat{\beta}^{SE} = \hat{\beta}^{UE} - (\hat{\beta}^{UE} - \hat{\beta}^{RE}) c L_n^{-1},
\]

where

\[
c = \frac{(q - 2)m}{q(m + 2)}, \quad q \geq 3.
\]

The SE in (1.7) will provide uniform improvement over \( \hat{\beta}^{UE} \), however it is not a convex combination of \( \hat{\beta}^{UE} \) and \( \hat{\beta}^{RE} \). Both (1.5) and (1.7) involve the statistic \( L_n \) which adjusts the estimator for departure from \( H_0 \). For a large value of \( L_n \), both (1.5) and (1.7) yield \( \hat{\beta}^{RRE} \), while for a small value of \( L_n \), their performance is different. The SE has the disadvantage that the shrinkage factor \((1 - c L_n^{-1})\) becomes negative for \( L_n < c \). This encourages one to find a better estimator, namely, the positive-rule shrinkage estimator (PR) of \( \beta \) as follows:

\[
\hat{\beta}^{PR} = \hat{\beta}^{SE} - (1 - c L_n^{-1}) I(L_n \leq c) (\hat{\beta}^{UE} - \hat{\beta}^{RE}).
\]

The PR estimator in (1.8) will provide uniform improvement over \( \hat{\beta}^{UE} \) and \( \hat{\beta}^{SE} \), and it is a convex combination of \( \hat{\beta}^{RRE} \) and \( \hat{\beta}^{RE} \). The properties of Stein-type estimators have been discussed and studied extensively under normal and non-normal assumptions by various researchers since the seminal works by Stein [53]. To mention a few, James and Stein [29], Judge and Bock [30], Shalabh [49], Chatuvredy et al. [16] and Saleh [43]. The positive part shrinkage estimator has been considered under the normal assumption by Ohtani [42], Adkins and Hill [1], Ahmed et al. [3,4,2], Arashi and Tabatabaey [10] and Withers and Nadarajah [56,57] among others.

It is observed from (1.2) that the usual least squares estimator (LSE) of \( \beta \) depends heavily on the characteristics of the matrix \( C = X'X^{-1}X \). If the \( C \) matrix is ill-conditioned, then the least squares estimator (LSE) produces unduly large sampling variances. To resolve this problem, Hoerl and Kennard [28] suggested to use \( C(k) = X'X^{-1}X + kl_p \), \((k \geq 0)\) rather than \( C \) in the estimation of \( \beta \). The resulting estimator of \( \beta \) is known as the ridge regression estimator (RRE). Recent applications of ridge regressions are given in [39] and the references therein. The ridge regression method has been considered by various researchers. Among them Hoerl and Kennard [28], Gibbons [20], Sarker [47], Saleh and Kibria [44], Gruber [23,24], Wenczeko [55], Kibria [33], and Kibria and Saleh [35] to mention a few.

The drawback of the ridge regression method is that it is a complicated function of \( k \). To overcome this problem, Liu [37] proposed the following estimator, which combines the benefit of Hoerl and Kennard [28] and Stein [53], defined as follows

\[
\hat{\beta}^{RRE}(d) = (C + I_p)^{-1} (X'Y + dF_{d}) = \hat{F}_d \hat{\beta}^{RRE},
\]

where \( F_d = (C + I_p)^{-1} (C + dI_p) \) and \( 0 < d < 1 \) is a biasing parameter. The Liu estimator has been considered by several researchers in several times for different perspectives. To mention a few, Kaciranlar et al. [31], Yuksel and Akdeniz [59], Liu [38], Alheety et al. [5], Yang and Xu [58] and very recently Kibria [34].

The literature of the Liu estimator under the assumption of the elliptically contoured distribution is limited. The main objective of this paper is to define the UE, RE, PT, SE and PR estimators in the line of the Liu estimator and provided them in the following section. The plan of the paper is as follows: in Section 2, we provide some Liu estimators and the corresponding expressions of biases and risks. The relative performance of the estimators based on the quadratic risk is presented in Section 3. To illustrate the findings of the paper, a numerical computation is given in Section 4. Finally, summary and conclusions have been included in Section 5.
2. Proposed estimators, bias and risk functions

2.1. Liu estimators (LEs)

In this section we present the Liu version of the five estimators of \( \beta \) given in Section 1. Accordingly, the unrestricted Liu estimator (ULE) of \( \beta \) is defined by
\[
\hat{\beta}_{\text{UE}}^L = F_d \hat{\beta}_{\text{UE}}.
\]
This is the well-known [37] estimator of \( \beta \).

Now, under the null hypothesis \( H_0 : H \hat{\beta} = h \), the restricted Liu estimator (RLE) of \( \beta \) is defined by
\[
\hat{\beta}_{\text{RE}}^L = F_d \hat{\beta}_{\text{UE}} - F_d C^{-1} H'(H C^{-1} H')^{-1} (H \hat{\beta}_{\text{UE}} - h).
\]

This estimator has been proposed by Kaciranlar et al. [31] based on the work of Sarker [47]. Yuksel and Akdeniz [59] proposed the following preliminary test Liu estimator (PTLE) of \( \beta \) which is defined as follows
\[
\hat{\beta}_{\text{PT}}^L = F_d \hat{\beta}_{\text{PT}} = F_d \hat{\beta}_{\text{UE}} - F_d (\hat{\beta}_{\text{UE}} - \hat{\beta}_{\text{RE}}) I(\mathcal{L}_n \leq \mathcal{L}_n, a).
\]

Parallel to PTLE we define the James–Stein type shrinkage Liu estimator (SLE) of \( \beta \) as
\[
\hat{\beta}_{\text{SE}}^L = F_d \hat{\beta}_{\text{SE}} = F_d \hat{\beta}_{\text{UE}} - F_d (\hat{\beta}_{\text{UE}} - \hat{\beta}_{\text{RE}}) c \mathcal{L}_n^{-1}.
\]

Finally we consider the positive-rule Liu estimator (PRLE) of \( \beta \) defined by
\[
\hat{\beta}_{\text{PR}}^L = F_d \hat{\beta}_{\text{PR}} = F_d \hat{\beta}_{\text{SE}} - (1 - c \mathcal{L}_n^{-1}) I(\mathcal{L}_n < c) F_d (\hat{\beta}_{\text{UE}} - \hat{\beta}_{\text{RE}}).
\]

For \( d = 1 \), the above proposed estimators reduce to UR, RE, PTE, SE and PRE respectively. In the following section we will discuss about the bias of the estimators.

2.2. Bias of the estimators

The biases of the proposed estimators are routinely derived from Saleh [43] and Hassanazadeh Bahtian et al. [27]. Therefore, we omit all derivations, instead, we present the expressions for the bias of the estimators in the following theorem.

**Theorem 2.1.** Biases of the ULE, RLE, PTLE, SLE and PRLE are given, respectively, by

\[
\begin{align*}
B(\hat{\beta}_{\text{UE}}^L) &= E(\hat{\beta}_{\text{UE}}^L - \beta) = -(1 - d)(C + I_p)^{-1} \beta; \\
B(\hat{\beta}_{\text{RE}}^L) &= -[(1 - d)(C + I_p)^{-1} \beta + F_d \eta]; \\
B(\hat{\beta}_{\text{PT}}^L) &= -[(1 - d)(C + I_p)^{-1} \beta + F_d \eta G_{q+2,m}(l_x; \Delta)]; \\
B(\hat{\beta}_{\text{SE}}^L) &= -[(1 - d)(C + I_p)^{-1} \beta + qC_d \eta E^2(\chi_{q+2}(\Delta))] \quad \text{and} \\
B(\hat{\beta}_{\text{PR}}^L) &= - \left[(1 - d)(C + I_p)^{-1} \beta + F_d \eta \left\{ d_1 E^2[F_{q+2,m}(\Delta)I(F_{q+2,m}(\Delta) < d_1)] - d_1 E^2[F_{q+2,m}(\Delta)I(F_{q+2,m}(\Delta) < d_1)] - C_{q+2,m}(l_x; \Delta) \right\} \right],
\end{align*}
\]

where \( \eta = C^{-1} H'(H C^{-1} H')^{-1} (H \beta - h) \), and

\[
G_{q+2,m}(l_x, \Delta) = \sum_{r=0}^{\infty} K_r^{(h)}(\Delta) I_x \left[ \frac{q + 2i}{2} + r, \frac{m}{2} \right],
\]

\[
l_x = \frac{qF_{q,m}(\alpha)}{m + \Psi_{q,m}(\alpha)}, \quad I_x[a, b] = \int_0^x u^{a-1}(1 - u)^{b-1} du / B(a, b) \quad \text{is the Pearson's regularized incomplete beta function. Further,}
\]

\[
E^{(2-h)}[\chi_{q+2}(\Delta)] = \sum_{r=0}^{\infty} K_r^{(h)}(\Delta)(q + s + 2 + 2r)^{-1},
\]

where

\[
K_r^{(h)}(\Delta) = [-2\psi'(0)]^r \left( \frac{\Delta}{2} \right)^r \int_0^\infty \frac{(t-1)^{-r+h}}{t^r} e^{-\frac{-1+\Delta\cdot t'}{2}} W(t)dt.
\]

Finally

\[
E^{(2-h)}[F_{q+2,n-p}(\Delta)I(F_{q+2,n-p}(\Delta) < d_1)] = \sum_{r=0}^{\infty} K_r^{(h)}(\Delta) \left( \frac{q + s}{n - p} \right) \frac{B(\frac{q+s+2r-2}{2}, \frac{m+2j}{2})}{B(\frac{q+s+2r}{2}, \frac{m}{2})} \left[ \frac{q + s + 2r - 2j}{2}, \frac{m + 2j}{2} \right],
\]

where \( d_1 = \frac{qc}{q+2}, x' = \frac{qc}{m+qc} \). In all the above \( h = 0, 1 \).
2.3. Risk of the estimators

In this subsection we will present the quadratic risk function. Suppose \( \hat{\beta}^* \) denotes an estimator of \( \beta \), then for a given non-singular matrix \( W \), the loss function is defined as

\[
L(\hat{\beta}^*; \beta) = (\hat{\beta}^* - \beta)'W(\hat{\beta}^* - \beta)
\]

and the corresponding risk function of the estimator \( \hat{\beta}^* \) is defined as

\[
R(\hat{\beta}^*; \beta) = E(\hat{\beta}^* - \beta)'W(\hat{\beta}^* - \beta) = tr(M); \quad \text{for } W = I_p,
\]

where \( M \) is the mean-square error matrix of the estimator \( \hat{\beta}^* \). The quadratic risk functions of the proposed estimators are routinely derived from Judge and Bock [30, Chapter 10], Saleh [43] and lead to the following theorem.

**Theorem 2.2.** Quadratic risk functions of ULE, RLE, PTLE, SLE and PRLE are given, respectively, by

\[
R(\hat{\beta}^{RE}(d); \beta) = \sigma^2_\epsilon^2 tr(F_0^2C^{-1}F_0) + (1 - d)^2\beta'(C + I_p)^{-2}\beta,
\]

\[
R(\hat{\beta}^{RE}(d); \beta) = \sigma^2_\epsilon^2 tr(F_0^2C^{-1}F_0) - \sigma^2_\epsilon^2 tr(F_0^2A F_0) + \eta^2F_0^2\eta + 2(1 - d)\eta F_0^2(C + I_p)^{-1}\beta + (1 - d)^2\beta'(C + I_p)^{-2}\beta,
\]

\[
R(\hat{\beta}^{PT}(d); \beta) = \sigma^2_\epsilon^2 tr(F_0^2C^{-1}F_0) - \sigma^2_\epsilon^2 tr(F_0^2A F_0)C_{q+2,m}(l_\alpha, \Delta) + \eta^2F_0^2\eta Z(\alpha, \Delta)
\]

\[
+ 2(1 - d)\eta F_0^2(C + I_p)^{-1}\beta \chi_{q+2,2}(\Delta) + (1 - d)^2\beta'(C + I_p)^{-2}\beta,
\]

\[
R(\hat{\beta}^{SE}(d); \beta) = \sigma^2_\epsilon^2 tr(F_0^2C^{-1}F_0) - \sigma^2_\epsilon^2 cqr(F_0^2AF_0) \times X(\Delta) + cqr^2 F_0^2\eta Y(\Delta)
\]

\[
+ 2cqr(1 - d)\eta F_0^2(C + I_p)^{-1}\beta E^{(2)}(\chi_{q+2,2}(\Delta)) + (1 - d)^2\beta'[C + I_p]^{-2}\beta,
\]

and

\[
R(\hat{\beta}^{PT}(d); \beta) = R(\hat{\beta}^{SE}(d); \beta) - \sigma^2_\epsilon^2 \left\{ tr(F_0^2A F_0) \right\} E^{(1)} \left\{ (1 - d_1 F_{q+2,m}^{-1}(\Delta))^2 I(F_{q+2,m}(\Delta) < d_1) \right\}
\]

\[
+ \sigma^2_\epsilon^2 \left( \eta F_0^2\eta \right) E^{(2)} \left\{ (1 - d_2 F_{q+4,m}(\Delta)) I(F_{q+4,m}(\Delta) < d_2) \right\}
\]


\[
- 2(\eta F_0^2\eta) E^{(2)} \left\{ (d_2 F_{q+2,m}(\Delta) - 1) I(F_{q+2,m}(\Delta) < d_1) \right\}
\]

\[
- 2(1 - d)\eta F_0^2(C + I_p)^{-1}\beta E^{(2)} \left\{ (d_2 F_{q+2,m}(\Delta) - 1) I(F_{q+2,m}(\Delta) < d_1) \right\}
\]

\[
\text{where}
\]

\[
E^{(2-h)}(\chi_{q+4}^{-4}(\Delta)) = \sum_{r \geq 0} k_r^{(h)}(\Delta)(q + s - 2 + 2r)^{-1}(q + s - 4 + 2r)^{-1},
\]

and

\[
X(\Delta) = 2E^{(1)}(\chi_{q+2}^{-2}(\Delta)) - (q - 2)E^{(1)}(\chi_{q+2}^{-2}(\Delta))
\]

\[
Y(\Delta) = 2E^{(2)}(\chi_{q+2}^{-2}(\Delta)) - 2E^{(2)}(\chi_{q+4}^{-2}(\Delta)) + (q - 2)E^{(2)}(\chi_{q+4}^{-4}(\Delta))
\]

\[
Z(\alpha, \Delta) = 2c^{(2)}_{q+2,m}(l_\alpha; \Delta) - c^{(2)}_{q+4,m}(l_\alpha; \Delta).
\]

Also, \( d_2 = \frac{\sigma^2}{\eta^2} \) and \( A = C^{-1}H'(HC^{-1}H')^{-1}HC^{-1} \).

Based on the above information we consider the performance of the estimators in the following section. For the brevity of the paper, we have omitted the analysis based on the quadratic biases. However, they are available in [8] upon request.

3. Risk analysis for the proposed estimators

In this section, we will compare the performance of the proposed estimators in the light of the quadratic risk function. Since \( C \) is a psd matrix, there exists an orthogonal matrix \( \Gamma' \) such that

\[
\Gamma' C \Gamma = \Lambda = diag(\lambda_1, \lambda_2, \ldots, \lambda_p),
\]
where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p > 0$ are the eigenvalues of $C$. It is easy to see that the eigenvalues of $F_d = (C + I_p)^{-1}(C + dl_p)$ and $(C + I_p)$ are $(\frac{\lambda_1 + d}{\lambda_1 + 1}, \frac{\lambda_2 + d}{\lambda_2 + 1}, \ldots, \frac{\lambda_p + d}{\lambda_p + 1})$ and $(\lambda_1 + 1, \lambda_2 + 1, \ldots, \lambda_p + 1)$ respectively. Then, we obtain the following identities

\[
\text{tr}(F_d^{-1}C^{-1}F_d) = \sum_{i=1}^{p} \frac{(\lambda_i + d)^2}{\lambda_i(\lambda_i + 1)^2}, \tag{3.1}
\]

\[
\beta'(C + I_p)^{-2} \beta = \sum_{i=1}^{p} \frac{\theta_i^2}{(\lambda_i + 1)^2}; \quad \theta = \Gamma' \beta, \tag{3.2}
\]

\[
\text{tr}(F_d\Gamma F_d') = \sum_{i=1}^{p} h^*_i(\lambda_i + d)^2 \frac{1}{(\lambda_i + 1)^2}, \tag{3.3}
\]

where $h^*_i \geq 0$ is the $i$th diagonal element of the matrix $H^* = \Gamma' A \Gamma$.

\[
\eta^* F_d \eta = \sum_{i=1}^{p} \eta_i^2(\lambda_i + d)^2 \frac{1}{(\lambda_i + 1)^2}, \tag{3.4}
\]

where $\eta_i^*$ is the $i$th element of vector $\eta^* = \eta' \Gamma$. Similarly,

\[
\eta^* F_d (C + I_p)^{-1} \beta = \sum_{i=1}^{p} \frac{\theta_i \eta_i^*(\lambda_i + d)}{(\lambda_i + 1)^2}. \tag{3.5}
\]

### 3.1. Comparison of PTLE with PTE, ULE, and RLE

Since the comparison among, UE, ULE, RE, RLE, PTE and PTLE is available in [34], we omit this analysis.

### 3.2. Comparison of SLE with SE

**Case 1:** Under the null hypothesis $H_0 : H \beta = h$.

The risk difference of these two estimators is

\[
R(\hat{\beta}^{SE}; \beta) - R(\hat{\beta}^{SE}(d); \beta) = \sigma^2 \text{cq}[\text{tr}(C^{-1} - F_d^{-1}F_d)] - \sigma^2 \text{cq}[\text{tr}(A - F_d\Gamma F_d')]X(0) - (1 - d)^2 \beta'(C + I_p)^{-2} \beta
\]

where

\[
X(0) = 2E^{(1)}(\chi^{-2}_{q+2}(0)) - (q - 2)E^{(1)}(\chi^{-d}_{q+2}(0)).
\]

Applying (3.1)–(3.5), the above difference can be expressed in terms of the eigenvalues as

\[
R(\hat{\beta}^{SE}; \beta) - R(\hat{\beta}^{SE}(d); \beta) = \sigma^2 \sum_{i=1}^{p} \frac{(1 - d)}{\lambda_i(\lambda_i + 1)^2} \left\{ (2\lambda_i + 1)(1 - \text{cq} h^*_i X(0)) - \lambda_i \theta_i^2 \sigma^2_{e^2} \right\}
\]

\[
- d(\lambda_i \text{cq} h^*_i X(0) + \lambda_i \theta_i^2 \sigma^2_{e^2} - 1). \tag{3.6}
\]

Let

\[
d_1 = \frac{\min \left\{ (2\lambda_i + 1)(1 - \text{cq} h^*_i X(0) - \lambda_i \theta_i^2 \sigma^2_{e^2}) \right\}}{\max \left\{ \text{cq} \lambda_i h^*_i X(0) + \lambda_i \theta_i^2 \sigma^2_{e^2} - 1 \right\}} \tag{3.7}
\]

and suppose $\tilde{d}_1$ denotes the denominator in the definition of $d_1$ in (3.7), then using (3.6) and following [31], we may state the following theorem.

**Theorem 3.1.** Suppose $d_1 > 0$ and the restriction in (1.3) is true. Then

1. If $\tilde{d}_1 > 0$, it follows that for each positive $d$ with $d < d_1$, $\hat{\beta}^{SE}(d)$ is better in the smaller MSE sense than $\hat{\beta}^{SE}$.
2. If $\tilde{d}_1 < 0$, it follows that for each positive $d$ with $d > d_1$, $\hat{\beta}^{SE}$ is better in the smaller MSE sense than $\hat{\beta}^{SE}(d)$. 


Case 2: Under the alternative hypothesis $H_0 : H \beta \neq h$.

The risk difference of these two estimators is

$$R(\hat{\beta}_e; \beta) - R(\hat{\beta}^{SE}(d); \beta) = \sigma_e^2[\text{tr}(C^{-1} - F_d'F_d)] - cq\sigma_e^2[\text{tr}(A - F_d'AF_d)]X(\Delta)$$

$$+ cq\eta'[I_p - F_d'F_d]\eta Y(\Delta) - 2cq(1 - d)\eta'F_d'(C + I_p)^{-1}\beta E^{(2)}(\chi^2_{q+2}(\Delta))$$

$$-(1 - d)^2\beta'(C + I_p)^{-2}\beta.$$ 

The above difference will be greater than or equal to 0, when

$$\eta'[I_p - F_d'F_d]\eta \geq \frac{f_1(\Delta, d)}{Y(\Delta)}, \quad (3.8)$$

where

$$f_1(\Delta, d) = cq\sigma_e^2[\text{tr}(A - F_d'AF_d)]X(\Delta) - \sigma_e^2[\text{tr}(C^{-1} - F_d'F_d)]$$

$$+ 2cq(1 - d)\eta'F_d'(C + I_p)^{-1}\beta E^{(2)}(\chi^2_{q+2}(\Delta)) + (1 - d)^2\beta'(C + I_p)^{-2}\beta.$$ 

Since $\Delta > 0$, we assume that the numerator of (3.8) is positive. Then following [6, Theorem A 2.4, p. 634], the SLE will dominate SE whenever $\Delta \leq \Delta_1(\Delta, d)$, where

$$\Delta_1(\Delta, d) = \frac{f_1(\Delta, d)}{cqCh_{\text{max}}[I_p - F_d'F_d]C^{-1}Y(\Delta)}.$$ 

However, SE dominates SLE when

$$\Delta < \Delta_2(\Delta, d) = \frac{f_1(\Delta, d)}{cqCh_{\text{min}}[I_p - F_d'F_d]C^{-1}Y(\Delta)}.$$ 

We now compare based on the parameter $d$. The risk function of $\hat{\beta}^{SE}(d)$ can be expressed in terms of the eigenvalues as

$$R(\hat{\beta}^{SE}(d); \beta) = \sigma_e^2\sum_{i=1}^{p} \frac{\lambda_i + d}{\lambda_i(\lambda_i + 1)^2} 2\sum_{i=1}^{p} \frac{\eta_i^2(\lambda_i + d)^2}{(\lambda_i + 1)^2} X(\Delta) + \sum_{i=1}^{p} \frac{cq\eta_i^2(\lambda_i + d)^2}{(\lambda_i + 1)^2} Y(\Delta)$$

$$+ 2cq(1 - d)\sum_{i=1}^{p} \frac{\theta_i\eta_i^2(\lambda_i + d)^2}{(\lambda_i + 1)^2} E^{(2)}(\chi^2_{q+2}(\Delta)) + (1 - d)^2\sum_{i=1}^{p} \frac{\theta_i^2}{(\lambda_i + 1)^2}, \quad (3.9)$$

Now differentiating (3.9) with respect to $d$, we get

$$\frac{\partial R(\hat{\beta}^{SE}(d))}{\partial d} = \sigma_e^2\sum_{i=1}^{p} \frac{2\eta_i^2}{(\lambda_i + 1)^2} \left\{ (\lambda_i + d) - \lambda_i h_{\lambda_i}^*(\lambda_i + d) X(\Delta) + \sigma_e^{-2}\lambda_i\eta_i^2(1 - d)Y(\Delta) \right\}$$

$$+ \sigma_e^{-2}\lambda_i(\eta_i^2 - \theta_i\eta_i^2\lambda_i - 2\theta_i\eta_i d)E^{(2)}(\chi^2_{q+2}(\Delta)) - (cq)^{-1}\sigma_e^{-2}\lambda_i(1 - d)\theta_i^2 \right\}$$

$$= \sigma_e^2\sum_{i=1}^{p} \frac{2}{(\lambda_i + 1)^2} \left( f_2(\Delta, \alpha) d - g_1(\Delta, \alpha) \right)$$

where

$$f_2(\Delta) = \min \left\{ 1 - \lambda_i h_{\lambda_i}^*(\lambda_i - 1) + \sigma_e^{-2}\lambda_i\eta_i^2 Y(\Delta) - 2\sigma_e^{-2}\lambda_i\theta_i\eta_i E^{(2)}(\chi^2_{q+2}(\Delta)) + (cq)^{-1}\sigma_e^{-2}\lambda_i\theta_i^2 \right\}$$

and

$$g_1(\Delta) = \max \left\{ \lambda_i \left( \lambda_i h_{\lambda_i}^*(\lambda_i - 1) - \sigma_e^{-2}\lambda_i\theta_i\eta_i^2 Y(\Delta) - \sigma_e^{-2}\theta_i\eta_i^2(1 - \lambda_i) + (cq)^{-1}\sigma_e^{-2}\theta_i^2 \right) \right\}.$$ 

Now we define

$$d_2(\Delta) = \frac{f_2(\Delta)}{g_1(\Delta)}, \quad (3.10)$$

Suppose $d > 0$, then we have the following.

1. If $g_1(\Delta) > 0$ and $f_2(\Delta) > 0$, it follows that for each positive $d$ with $d > d_2(\Delta)$, $\hat{\beta}^{SE}(d)$ has risk value less than that of $\hat{\beta}^{SE}$. 
2. If $g_1(\Delta) < 0$ and $f_2(\Delta) < 0$, it follows that for each positive $d$ with $d < d_2(\Delta)$, $\hat{\beta}^{SE}(d)$ has risk value less than that of $\hat{\beta}^{SE}$.
3.3. Comparison of SLE with PTLE

In this case the risk difference, \( R(\hat{\beta}_{SE}^{\alpha}(d); \beta) - R(\hat{\beta}_{PT}^{\alpha}(d); \beta) \) will be non-positive if

\[
\eta F_{d}^{\prime} F_{d} \eta \geq \frac{f_{3}(\Delta, d, \alpha)}{Z(\alpha, \Delta) -cqY(\Delta)},
\]

where

\[
f_{3}(\Delta, d, \alpha) = tr(F_{d} A F_{d}) \left\{ C_{q+2,m}^{(1)}(l_{\alpha}; \Delta) -cqX(\Delta) \right\} + 2(1-d)\alpha^{-2} \eta F_{d}^{\prime}(C + I_{p})^{-1} \beta \left\{ cqE_{2}(\chi_{q+2}^{-2}(\Delta)) - G_{q+2,m}^{(2)}(l_{\alpha}; \Delta) \right\}.
\]

Since \( \Delta > 0 \), we assume that both the numerator and the denominator of (3.11) are positive or negative, respectively. Then \( \hat{\beta}_{SE}^{\alpha}(d) \) dominates \( \hat{\beta}_{PT}^{\alpha}(d) \) when \( \Delta \leq \Delta_{3}(\Delta, d, \alpha) \), and \( \hat{\beta}_{PT}^{\alpha}(d) \) dominates \( \hat{\beta}_{SE}^{\alpha}(d) \) when \( \Delta > \Delta_{4}(\Delta, d, \alpha) \).

Now, we consider the risk difference \( R(\hat{\beta}_{SE}^{\alpha}(d); \beta) - R(\hat{\beta}_{PT}^{\alpha}(d); \beta) \) as a function of eigenvalues and define

\[
d_{3}(\Delta, \alpha) = \frac{f_{4}(\Delta, \alpha)}{g_{2}(\Delta, \alpha)},
\]

where

\[
f_{4}(\Delta, \alpha) = \max_{i} \left\{ \lambda_{i} \eta_{i}^{\ast} Z(\alpha, \Delta) -cqY(\Delta) \right\} - \sigma_{e}^{2} h_{i} \lambda_{i} \left\{ C_{q+2,m}^{(1)}(l_{\alpha}; \Delta) -cqX(\Delta) \right\} + 2\theta \eta_{i}^{\ast} \left\{ G_{q+2,m}^{(2)}(l_{\alpha}; \Delta) -cqE_{2}(\chi_{q+2}^{-2}(\Delta)) \right\}
\]

and

\[
g_{2}(\Delta, \alpha) = \min_{i} \left\{ \sigma_{e}^{2} h_{i}^{\ast} \left\{ C_{q+2,m}^{(1)}(l_{\alpha}; \Delta) -cqX(\Delta) \right\} - \eta_{i}^{\ast} Z(\alpha, \Delta) -cqY(\Delta) \right\} + 2\theta \eta_{i}^{\ast} \left\{ G_{q+2,m}^{(2)}(l_{\alpha}; \Delta) -cqE_{2}(\chi_{q+2}^{-2}(\Delta)) \right\}.
\]

Suppose \( d > 0 \), then we have the following.

1. If \( g_{2}(\Delta, \alpha) > 0 \) and \( f_{4}(\Delta, \alpha) > 0 \), it follows that for each positive \( d \) with \( d < d_{3}(\Delta, \alpha) \), \( \hat{\beta}_{SE}^{\alpha}(d) \) has risk value less than that of \( \hat{\beta}_{PT}^{\alpha}(d) \).

2. If \( g_{2}(\Delta, \alpha) < 0 \) and \( f_{4}(\Delta, \alpha) < 0 \), it follows that for each positive \( d \) with \( d > d_{3}(\Delta, \alpha) \), \( \hat{\beta}_{SE}^{\alpha}(d) \) has risk value less than that of \( \hat{\beta}_{PT}^{\alpha}(d) \).

Under \( H_{0} \), the risk difference reduces to

\[
\sigma_{e}^{2} \sum_{i=1}^{p} \frac{h_{i}^{\ast}(\lambda_{i} + d)^{2}}{(\lambda_{i} + 1)^{2}} \left\{ C_{q+2,m}^{(1)}(l_{\alpha}; 0) -cqX(0) \right\},
\]

where \( C_{q+2,m}^{(1)}(l_{\alpha}; 0) = F_{q+2,m}^{(q_{f_{a}} q_{f_{a}})}(0) \) \( = 1 - \alpha \). Thus, the risk of \( \hat{\beta}_{SE}^{\alpha}(d) \) is smaller than that of \( \hat{\beta}_{PT}^{\alpha}(d) \) when the critical value \( l_{\alpha} \) satisfies

\[
l_{\alpha} \leq \frac{q + 2}{q} F_{q+2,m}^{-1}(cqX(0)).
\]

Otherwise, the risk of \( \hat{\beta}_{PT}^{\alpha}(d) \) is smaller than the risk of \( \hat{\beta}_{SE}^{\alpha}(d) \).

**Remark 1.** For \( \alpha = 1 \), we will have the comparison between SLE and ULE and for \( \alpha = 0 \), we will have the comparison between SLE and RLE.
3.4. Comparison of PRLE with PR, ULE, RLE, PTLE and SLE

3.4.1. Comparison of PRLE and PR

Case 1. Under the null hypothesis, \( H_0 : H \beta = h \).

Differentiating the risk function \( R(\hat{\beta}^{PR}(d); \beta) \) with respect to \( d \), we find a sufficient condition so that \( \frac{\partial R(\hat{\beta}^{PR}(d); \beta)}{\partial d} \) to be negative is that \( d \in (0, d_4) \), where

\[
d_4 = \frac{\min_i \left\{ \lambda_i [1 - cq h_i^* X(0) - h_i^* \lambda_i E^{(1)} [(1 - d_1 F_{q+2,m}(0))^2 I(F_{q+2,m}(0) < d_1)] - \sigma_e^{-2} \theta_i^2] \right\}}{\max_i \left\{ cq h_i^* X(0) - h_i^* \lambda_i E^{(1)} [(1 + d_1 F_{q+2,m}(0))^2 I(F_{q+2,m}(0) < d_1)] - \sigma_e^{-2} \theta_i^2 - 1 \right\}}. \tag{3.14}
\]

Suppose the numerator of (3.14) is positive, then, \( \hat{\beta}^{PR}(d) \) is superior to \( \hat{\beta}^{PR} \), when \( d > 0 \) belongs to the region \( 0 < d < d_4 \).

Case 2. Non-null case, \( H_0 : H \beta \neq h \).

Differentiating the risk difference with respect to \( d \), define,

\[
d_5(\Delta) = \frac{f_5(\Delta)}{g_5(\Delta)}, \tag{3.15}
\]

where

\[
f_5(\Delta) = \min_i \left\{ 1 - \lambda_i h_i^*[X(\Delta) + a_1] + \sigma_e^{-2} \lambda_i \eta_i^*[Y(\Delta) - a_2 - 2a_3] - \sigma_e^{-2} \lambda_i \theta_i^* \left( E^{(1)}[\chi_{q+2}^{-1}(\Delta)] + 2a_3 \right) + \lambda_i \theta_i \right\}
\]

and

\[
g_5(\Delta) = \max_i \left\{ \lambda_i h_i^*[X(\Delta) + a_1] - 1 - \sigma_e^{-2} \lambda_i \eta_i^*[Y(\Delta) + a_2 + 2a_3] - \sigma_e^{-2} \lambda_i \theta_i^* \right\}.
\]

Also,

\[
a_1 = E^{(1)} \left[ (1 - d_1 F_{q+2,m}(\Delta))^2 I\left(F_{q+2,m}(\Delta) < d_1 \right) \right]
\]
\[
a_2 = E^{(2)} \left[ (1 - d_2 F_{q+4,m}(\Delta))^2 I\left(F_{q+4,m}(\Delta) < d_2 \right) \right]
\]
\[
a_3 = E^{(2)} \left[ (d_1 F_{q+2,m}^{-1}(\Delta) - 1) I\left(F_{q+2,m}(\Delta) < d_1 \right) \right].
\]

Suppose \( d > 0 \), then we can write the following statements.

1. If \( g_5(\Delta) > 0 \) and \( f_5(\Delta) > 0 \), it follows that for each positive \( d \) with \( d > d_5(\Delta) \), \( \hat{\beta}^{PR}(d) \) has risk value less than that of \( \hat{\beta}^{PR} \).
2. If \( g_5(\Delta) < 0 \) and \( f_5(\Delta) < 0 \), it follows that for each positive \( d \) with \( d < d_5(\Delta) \), \( \hat{\beta}^{PR}(d) \) has risk value less than that of \( \hat{\beta}^{PR} \).

To obtain a condition on \( \Delta \), we consider the risk difference between PRLE and PR estimators as,

\[
R(\hat{\beta}^{PR}(d); \beta) - R(\hat{\beta}^{PR}; \beta) = \sigma_e^2 \left[ \text{tr}(F'_d C^{-1} F_d) - \text{tr}(C^{-1}) \right] + \sigma_e^2 \left[ \text{tr}(A) - \text{tr}(F'_d A F_d) \right] \left[ cq X(\Delta) \right] \\
+ E^{(1)} \left[ (1 - d_1 F_{q+2,m}(\Delta))^2 I(F_{q+2,m}(\Delta) < d_1) \right] \\
+ 2(1 - d) \eta I(P - F'_d F_d) \left[ cq E^{(2)}(X^{-2}_{q+2}(\Delta)) \right] \\
- E^{(2)} \left[ (d_1 F_{q+2,m}^{-1}(\Delta) - 1) I(F_{q+2,m}(\Delta) < d_1) \right] - \eta[I_p - F'_d F_d] \eta g_4(\Delta), \tag{3.16}
\]

where

\[
g_4(\Delta) = cq Y(\Delta) - a_2 - 2a_3.
\]

The right-hand side of (3.16) is non-positive when

\[
\eta [I_p - F'_d F_d] \eta \geq \frac{f_6(\Delta, d)}{g_4(\Delta)}, \tag{3.17}
\]

where

\[
f_6(\Delta, d) = \text{tr}(F'_d C^{-1} F_d - C^{-1}) + \text{tr}(A - F'_d A F_d) \left[ cq X(\Delta) + a_1 \right] \\
+ 2 \sigma_e^2 (1 - d) \eta I(P - F'_d F_d)^{-1} \left[ cq E^{(2)}(X^{-2}_{q+2}(\Delta)) - a_2 \right].
\]

Since \( \Delta > 0 \), assume that both the numerator and the denominator of (3.17) are positive or negative, respectively. Then, \( \hat{\beta}^{PR}(d) \) dominates \( \hat{\beta}^{PR} \) when \( \Delta \leq \Delta_5(\Delta, d) = \frac{f_6(\Delta, d)}{c_{\text{max}}[I_p - F'_d F_d] C^{-1}] \times g_4(\Delta) \) and \( \hat{\beta}^{PR}(d) \) dominates \( \hat{\beta}^{PR}(d) \) when \( \Delta > \Delta_5(\Delta, d) = \frac{f_6(\Delta, d)}{c_{\text{min}}[I_p - F'_d F_d] C^{-1}] \times g_4(\Delta) \).
3.4.2. Comparison of PRLE with ULE, RLE, PTLE and SLE

Since the ULE and the RLE are particular cases of the PTLE, therefore, we skip the comparison between ULE and PRLE, and also between RLE and PRLE.

Comparison between PRLE and PTLE
Case 1. Under the null hypothesis, $H_0 : H \beta = h$.

The risk difference is
\[
R(\hat{\beta}^{\text{PR}}(d); \beta) - R(\hat{\beta}^{\text{PT}}(d); \beta)
\]
\[
= \sum_{i=1}^{p} h_i^*(\lambda_i + d)^{\sigma_i^2} \left\{ C^{(1)}_{q+2,m}(l_\alpha; 0) - cqX(0) - E^{(1)} \left[ (1 - d_1F_{q+2,m}^{-1}(0))^{\gamma}I(F_{q+2,m} \leq d_1) \right] \right\} \geq 0
\] (3.18)
for all $\alpha$ satisfying the condition
\[
\{ \alpha : l_\alpha \geq C_{q+2,m}^{-1} (cqX(0) + E^{(1)} \left[ (1 - d_1F_{q+2,n-p}^{-1}(0))^{\gamma}I(F_{q+2,n-p} \leq d_1) \right] ; 0) \}.
\] (3.19)
Thus, the risk of PTLE is smaller than that of PRLE when the critical value $l_\alpha$ satisfies the relation in (3.19). However, the risk of PRLE is smaller than that of PTLE when the critical value $l_\alpha$ satisfies the opposite relation to (3.19).

Case 2. Under alternative hypothesis, $H_a : H \beta \neq h$.

The risk difference $R(\hat{\beta}^{\text{PR}}(d); \beta) - R(\hat{\beta}^{\text{PT}}(d); \beta)$ will be non-positive when
\[
n_{F_d}F_d\eta_g \geq \frac{f_7(\Delta, d, \alpha)}{g_5(\Delta, \alpha)}
\] (3.20)
where
\[
f_7(\Delta, d, \alpha) = \sigma_i^2 \text{tr}(F_dA_d^F) \left\{ C^{(1)}_{q+2,m}(l_\alpha; \Delta) - cqX(\Delta) + a_1 \right\}
\]
\[
- 2(1 - d)\eta_i F_d(C + l_0)^{-1} \beta \left\{ cE^{(2)}(X_{q+2}^{-2}(\Delta)) - c^{(2)}_{q+2,m}(l_\alpha; \Delta) - a_3 \right\}
\]
and
\[
g_5(\Delta, \alpha) = Z(\alpha, \Delta) - cqY(\Delta) + a_2 + 2a_3.
\]
Since $\Delta > 0$, assume that both the numerator and the denominator of (3.20) are positive or negative, respectively. Then $\hat{\beta}^{\text{PR}}(d)$ dominates $\hat{\beta}^{\text{PT}}(d)$ when $\Delta \leq \Delta_7(\Delta, d, \alpha) = \frac{f_7(\Delta, d, \alpha)}{\inf(\{f_d(\Delta, d, \alpha)\}^{c-1}) \times g_5(\Delta, \alpha)}$, and $\hat{\beta}^{\text{PT}}(d)$ dominates $\hat{\beta}^{\text{PR}}(d)$ when $\Delta > \Delta_8(\Delta, d, \alpha) = \frac{f_7(\Delta, d, \alpha)}{\inf(\{f_d(\Delta, d, \alpha)\}^{c-1}) \times g_5(\Delta, \alpha)}$.

Now, we consider the risk difference of PRLE and PTLE as a function of eigenvalues and define
\[
d_6(\Delta, \alpha) = \frac{f_6(\Delta, \alpha)}{g_6(\Delta, \alpha)}
\] (3.21)
where
\[
f_6(\Delta, \alpha) = \max \left\{ \lambda_i h_i^*P_1 + \lambda_i \eta_i^*Q_2 + \theta \eta_i^*Q_3 \right\}
\]
and
\[
g_6(\Delta, \alpha) = \min \left\{ 2 + \theta \lambda_i \eta_i^*P_3 - \lambda_i h_i^*P_1 - \lambda_i \eta_i^*Q_3 \right\}
\]
Also,
\[
P_1 = C^{(1)}_{q+2,m}(l_\alpha; \Delta) - cqX(\Delta) + a_1,
\]
\[
P_2 = cqY(\Delta) - a_2 - 2a_3 - Z(\alpha, \Delta),
\]
\[
P_3 = cE^{(2)}(X_{q+2}^{-2}(\Delta)) - a_3 - G^{(2)}_{q+2,m}(l_\alpha; \Delta) - cqX(\Delta).
\]
Suppose $d > 0$, then we have the following statements.
1. If $g_6(\Delta, \alpha) > 0$ and $f_6(\Delta, \alpha) > 0$, it follows that for each positive $d$ with $d > d_6(\Delta, \alpha)$, $\hat{\beta}^{\text{PR}}(d)$ has risk value less than that of $\hat{\beta}^{\text{PT}}(d)$.
2. If $g_6(\Delta, \alpha) < 0$ and $f_6(\Delta, \alpha) < 0$, it follows that for each positive $d$ with $d < d_6(\Delta, \alpha)$, $\hat{\beta}^{\text{PR}}(d)$ has risk value less than that of $\hat{\beta}^{\text{PT}}(d)$.

Remark 2. For $\alpha = 0$, we obtain the superiority condition of $\hat{\beta}^{\text{PR}}(d)$ over $\hat{\beta}^{\text{RE}}(d)$ and for $\alpha = 1$, we obtain the superiority condition of $\hat{\beta}^{\text{PR}}(d)$ over $\hat{\beta}^{\text{UE}}(d)$. 
Comparison of PRLE and SLE

\[
R(\hat{\beta}^{PR}(d); \beta) - R(\hat{\beta}^{SE}(d); \beta) = -\sigma^2 \left\{ \text{tr}(F_d'AF_d)E^{(1)} \left[ (1 - d_1 F_{q+2,m}(\Delta))^2 I(F_{q+2,m}(\Delta) < d_1) \right] \right. \\
+ \sigma^2 \left. \left[ (1 - d_2 F_{q+4,m}(\Delta))^2 I(F_{q+4,m}(\Delta) < d_2) \right] \right\} \\
- 2(\eta F_d' \eta)E^{(2)} \left[ (d_1 F_{q+2,m}(\Delta) - 1) I(F_{q+2,m}(\Delta) < d_1) \right] \\
- 2(1 - d) \eta F_d'(C + I_p)^{-1} \beta E^{(2)} \left[ (d_1 F_{q+2,m}(\Delta) - 1) I(F_{q+2,m}(\Delta) < d_1) \right].
\] (3.22)

Case 1: Suppose \( \eta F_d'(C + I_p)^{-1} \beta \) is positive, then the right-hand side of (3.22) is negative, since the expectation of a positive random variable is positive. Thus, for all \( \Delta \) and \( d \)

\[
R(\hat{\beta}^{PR}(d); \beta) \leq R(\hat{\beta}^{SE}(d); \beta).
\]

Therefore under this condition, the PRLE not only confirms the inadmissibility of SLE but also provides a simple superior random variable is positive. Thus, for all \( \Delta \) and \( d \)

\[
R(\hat{\beta}^{PR}(d); \beta) \leq R(\hat{\beta}^{SE}(d); \beta).
\]

Case 2: Suppose \( \eta F_d'(C + I_p)^{-1} \beta \) is negative, then the difference in (3.22) will be positive when

\[
\eta F_d' \eta \geq \frac{f_0(\Delta, d)}{g_1(\Delta)},
\] (3.23)

where

\[
f_0(\Delta, d) = 2(1 - d) \eta F_d'(C + I_p)^{-1} \beta a_3 - \sigma^2 \text{tr}(F_d'AF_d) a_1
\]

and

\[
g_1(\Delta) = -(a_2 + 2a_3).
\]

Since \( \Delta > 0 \), assume that both the numerator and the denominator of (3.23) are positive or negative, respectively. Then \( \hat{\beta}^{PR}(d) \) dominates \( \hat{\beta}^{SE}(d) \), when \( \Delta \leq \Delta_0(\Delta, d) = \frac{f_0(\Delta, d)}{\sup(\{F_d'AF_d\} \times g_1(\Delta))} \) and \( \hat{\beta}^{SE}(d) \) dominates \( \hat{\beta}^{PR}(d) \), when \( \Delta > \Delta_0(\Delta, d) = \frac{f_0(\Delta, d)}{\sup(\{F_d'AF_d\} \times g_1(\Delta))} \).

Thus, it is observed that even though the PR least squares estimator uniformly dominates both SE and UE, the PRLE does not uniformly dominate both ULE and SLE.

3.5. Estimation of \( d \)

Using Eqs. (3.1) and (3.2), the risk function of \( \hat{\beta}^{UE}(d) \) (in Eq. (2.1)) can be expressed as

\[
f(d) = R(\hat{\beta}^{UE}(d); \beta) = \sigma^2 \text{tr}(F_d'C^{-1}F_d) + (1 - d)^2 \beta'(C + I_p)^{-2} \beta,
\]

\[
= \sigma^2 \sum_{i=1}^{p} \frac{(\lambda_i + d)^2}{\lambda_i(\lambda_i + 1)^2} + (1 - d)^2 \sum_{i=1}^{p} \frac{\theta_i^2}{(\lambda_i + 1)^2}
\]

\[
= \delta_1(d) + \delta_2(d).
\] (3.24)

For the Liu estimator one wants to find a value of \( d \), so that the decrease in variance (\( \delta_1(d) \)) is greater than the increase in squared bias (\( \delta_2(d) \)). In order to show that such a value less than 1 exists so that \( R(\hat{\beta}^{UE}(d); \beta) < R(\hat{\beta}^{UE}, \beta) \), we will take derivative of Eq. (3.24) with respect to \( d \) as,

\[
f'(d) = 2\sigma^2 \sum_{i=1}^{p} \frac{(\lambda_i + d)}{\lambda_i(\lambda_i + 1)^2} - 2(1 - d) \sum_{i=1}^{p} \frac{\theta_i^2}{(\lambda_i + 1)^2}.
\] (3.25)

For \( d = 1 \) in (3.25), we obtain,

\[
f'(d) = 2 \sum_{i=1}^{p} \frac{1}{\lambda_i(\lambda_i + 1)} > 0, \quad \text{since} \quad \lambda_i \geq 0.
\] (3.26)

Therefore, there exists a value of \( d \) (0 < \( d < 1 \)), such that \( R(\hat{\beta}^{UE}(d); \beta) \leq R(\hat{\beta}^{UE}; \beta) \). Now from (3.25), the risk of \( \hat{\beta}^{UE}(d) \) is minimized at

\[
d = \frac{\sum_{i=1}^{p} \sigma^2 \sum_{i=1}^{p} \frac{(\lambda_i + d)^2}{\theta_i^2}}{\sum_{i=1}^{p} (\lambda_i + 1)^2}
\] (3.27)
where $\theta_i$ is the $i$th component of the vector, $\theta = \Gamma^T \beta$, which is defined in Eq. (3.2). Replacing $\hat{\theta}_i^2$ and $\sigma^2$ by their corresponding unbiased estimates $\hat{\theta}_i^2 - \hat{\sigma}_i^2$ and $\hat{\sigma}_i^2$, we get the optimum value of $d$ as

$$
d_{opt} = 1 - \hat{\sigma}_i^2 \left( \sum_{i=1}^{p} \frac{1}{\hat{\lambda}(\hat{\lambda} + 1)} \right)^{-2} \left( \sum_{i=1}^{p} \frac{\hat{\theta}_i^2}{(\hat{\lambda} + 1)^2} \right)^{-2},
$$

(3.28)

More on the estimation of $d$ we refer our readers to Liu [37] and Kristofer et al. [36], among others.

### 4. Illustration for the multivariate Student’s t-distribution

In this section, we display the graphs of risk functions for the MT model based on both $\Delta$ and $d$. In this regard, the error term in model (1.1) has the MT$_n(0, \sigma^2 \nu, \nu)$ distribution with the location 0, scale matrix $\sigma^2 \nu$ and $\nu$ degree of freedom with the following pdf

$$
f(e) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{(\nu \pi)^{\frac{\nu+1}{2}}} \left(1 + \frac{|e|^2}{\nu \sigma^2}\right)^{-\frac{\nu+1}{2}}, \quad 0 < \nu, \sigma < \infty.
$$

Therefore, we have $\Sigma_e = \frac{\nu^2}{\nu - 2} \nu$. Also note that the term $\sigma^2$ in (1.6) is equal to $\frac{\nu^2}{\nu - 2}$.

This distribution belongs to the class of ECDs with the weight function (A.3). See [11] for more details and discussion.

It can be directly concluded that $\int_0^\infty t^{-1} W(t) dt = \frac{\nu}{\nu - 2}$. In this case, $G^{(2-h)}_{q_1, m}(x', \Delta)$, $E^{(2-h)}[F_{q_1, m}(\Delta) I(F_{q_1, m}(\Delta) < \frac{qc}{q_s})]$ can be simply computed by using the fact that

$$
K^{(h)}_r(\Delta) = \left(\frac{\nu}{2}\right)^h \Gamma\left(r + \frac{\nu}{2} - h\right) \Gamma\left(r + 1\right) \left(1 + \frac{\Delta}{\nu \nu_s}\right)^{\frac{r}{2} - h}. 
$$

Although, the respective expressions are given in [11], however, there are some minor typos, which we have corrected them and presented here.

$$
G^{(2-h)}_{q_1, m}(\nu, \Delta) = \sum_{r=0}^{\infty} \left(\frac{\nu}{2}\right)^h \frac{\Gamma\left(r + \frac{\nu}{2} - h\right)}{\Gamma\left(r + 1\right) \left(1 + \frac{\Delta}{\nu \nu_s}\right)^{\frac{r}{2} - h}} I_r \left[\frac{1}{2}(q + 2i) + r, \frac{m}{2}\right], 
$$

$$
E^{(2-h)}[F_{q_1, m}(\Delta) I(F_{q_1, m}(\Delta) < \frac{qc}{q_s})] = \sum_{r=0}^{\infty} \left(\frac{\nu}{2}\right)^h \frac{\Gamma\left(r + \frac{\nu}{2} - h\right)}{\Gamma\left(r + 1\right) \left(1 + \frac{\Delta}{\nu \nu_s}\right)^{\frac{r}{2} - h}} I_r \left[\frac{1}{2}(q + s + 4 + 2r) \frac{1}{2}(m + 2)\right], 
$$

$$
E^{(2-h)}[F_{q_1, m}(\Delta) I(F_{q_1, m}(\Delta) < \frac{qc}{q_s})] = \sum_{r=0}^{\infty} \left(\frac{\nu}{2}\right)^h \frac{\Gamma\left(r + \frac{\nu}{2} - h\right)}{\Gamma\left(r + 1\right) \left(1 + \frac{\Delta}{\nu \nu_s}\right)^{\frac{r}{2} - h}} I_r \left[\frac{1}{2}(q + s + 4 + 2r) \frac{1}{2}(m + 4)\right]. 
$$

MT plays an important role in robust statistical inference, particularly for heavy tailed distributions involving outliers and extreme values.

To achieve matrix $X$, following [40,20], the explanatory variables were generated using the following device

$$
x_{ij} = (1 - \gamma^2)z_{ij} + \gamma^2 z_{ip}, \quad i = 1, 2, \ldots, n; \quad j = 1, 2, \ldots, p.
$$

where $z_{ij}$s are independent standard normal pseudo-random numbers, and $\gamma$ is specified so that the correlation between any two explanatory variables is given by $\gamma^2$. 

In order to display the graphs, we take $p = 4$, $\gamma = 0.8$, $n = 30$, $q = 3$, $v = 5$ and $V = \text{diag}(1, \ldots, 30)$ that $v_{i,j} = 0; \forall i \neq j; i, j = 1, 2, \ldots, n$. We further suppose that the parameters $\beta$, $H$ matrix and $h$ vector in (1.3) have the following forms respectively,

$$
\beta = [10 -11]' , \quad H = \begin{bmatrix}
1 & -1 & 3 & 1 \\
3 & 2 & 1 & 0 \\
4 & -2 & 0 & 5
\end{bmatrix}, \quad \text{and} \quad h = [0\ 0\ 0]' .
$$

For $n = 10$ and $30$, the risk functions vs the biasing parameter $d$ and non-central parameter $\Delta$ are presented in Figs. 1 and 2 respectively. As it can be seen that when $n$ increases, the risk function gets more precise. For $n = 30$ and different levels of significance $\alpha$, the risk functions vs the biasing parameter $d$ and $\Delta$ are displayed in Figs. 3 and 4 respectively. To conclude our result, in Figs. 5 and 6, risk functions are displayed comparing Liu type estimators with their counterparts based on $\Delta$ and $d$, respectively. Overall, from these figures, we observed that the proposed improved Liu type estimators performed uniformly better than their counterpart for any $d$ and $\Delta$, which also supported the theoretical results of the paper. To be more specific about the figures, we discuss the result of each figure with the emphasis on the Liu type estimators. From Fig. 1, as the sample size increases the risk values decrease (as always one expects), however the risk functions are decreasing w.r.t the biasing parameter $d$, when the null-hypothesis is not true. As it is seen from Fig. 2, when the sample size increases the risk of the unrestricted Liu estimator decreases which causes a systematic change (fall) in the risks of the preliminary test as well as the positive-rule Liu estimator, since the risk of the restricted Liu estimator remains nearly unchanged. From Fig. 3, the increasing status of the risk function is apparent w.r.t the biasing parameter $d$. It can be also understood that, as the level of significance increases the risk of the preliminary test Liu estimator increases as well. The performance of Liu type estimators is not the same w.r.t the biasing parameter $d$. As it is seen from Fig. 4. Further for a small level of significance $\alpha$, the preliminary test Liu estimator performs better than the James–Stein type shrinkage Liu estimator, however this domination loosens as $\alpha$ increases to the favor of the SLE estimator. About the results of Figs. 5 and 6 together, we can say the risk of Liu type estimators is less than their counterparts. The same performance as described in the above can be seen based on the parameters $d$ and $\Delta$.

Furthermore, to illustrate the impact of proposed theorems based on different parameter values, we tabulate upper/lower boundaries for some specific values. The proposed results given in Tables 1–8 confirm the conditions of specified superiorities. Again to be more precise about the tables, we should note that to estimate both parameters $\Delta$ and $d$ by solving non-linear equations is not easy. Thus in order to have a better understanding, we fix one parameter of interest ($\Delta$ or $d$) at one stage and try to get close to the boundary of superiority. To be more specific, from Table 1, we see that $\Delta$ is highly related to $\Delta_i(\Delta, d), i = 1, 2$ for the superiority of SLE over SE. Thus, we fix the parameters $\Delta$ and $d$, and
Fig. 3. Risk performance for $n = 30$, $\Delta = 5$ and different values of $\alpha$ and $d$.

Fig. 4. Risk performance for $n = 30$, $d = 0.5$ and different values of $\alpha$ and $\Delta$. 
evaluate $\Delta_i(\Delta, d), i = 1, 2$. If the superiority conditions satisfy (i.e., $\Delta \leq \Delta_1(\Delta, d)$ for fixed values) then we have the solution (SLE performs better than SE), otherwise we skip the values. This procedure is the same looking at Table 2. The value of biasing parameter $d$ for which SLE performs better than SE is obtained from the performance of $g_1(\Delta)$, however it is related non-linearly to the value of $\Delta$. Thus to solve this relation in some stage, we fix the value of $\Delta$ and seek for relevant $g_1(\Delta)$ and $d_2(\Delta)$ from Table 2, that the superiority condition holds. The above scenario is the same for the other
Comparison of SLE with SE based on $\Delta$.

<table>
<thead>
<tr>
<th>$n\Delta$</th>
<th>$\Delta_1(\Delta, d)$</th>
<th>$\Delta_2(\Delta, d)$</th>
<th>$n\Delta$</th>
<th>$\Delta_1(\Delta, d)$</th>
<th>$\Delta_2(\Delta, d)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-1.265995</td>
<td>-91.6578</td>
<td>0</td>
<td>0.05413175</td>
<td>14.95358</td>
</tr>
<tr>
<td>0.25</td>
<td>4.675618</td>
<td>376.8689</td>
<td>0.25</td>
<td>3.193983</td>
<td>994.7763</td>
</tr>
<tr>
<td>0.5</td>
<td>18.10744</td>
<td>1603.757</td>
<td>0.5</td>
<td>10.60104</td>
<td>3609.282</td>
</tr>
<tr>
<td>0.75</td>
<td>60.63166</td>
<td>5839.301</td>
<td>0.75</td>
<td>34.42246</td>
<td>13088.89</td>
</tr>
<tr>
<td>1</td>
<td>1.311755 \times 10^{-17}</td>
<td>-8.674545 \times 10^{-16}</td>
<td>1</td>
<td>4.496652 \times e^{-16}</td>
<td>-6.690998 \times 10^{-16}</td>
</tr>
<tr>
<td>0</td>
<td>-13.30998</td>
<td>-963.641</td>
<td>0</td>
<td>-6.382443</td>
<td>-1763.113</td>
</tr>
<tr>
<td>0.25</td>
<td>22.31648</td>
<td>1798.776</td>
<td>0.25</td>
<td>15.71458</td>
<td>4894.495</td>
</tr>
<tr>
<td>0.5</td>
<td>101.7867</td>
<td>9015.145</td>
<td>0.5</td>
<td>66.06943</td>
<td>2268.26</td>
</tr>
<tr>
<td>0.75</td>
<td>352.024</td>
<td>33902.65</td>
<td>0.75</td>
<td>225.8755</td>
<td>85888.14</td>
</tr>
<tr>
<td>1</td>
<td>7.70463 \times 10^{-17}</td>
<td>-5.095024 \times 10^{-17}</td>
<td>30</td>
<td>3.003902 \times 10^{-17}</td>
<td>-4.408482 \times 10^{-17}</td>
</tr>
<tr>
<td>0</td>
<td>-32.47769</td>
<td>-2351.38</td>
<td>0</td>
<td>-16.95362</td>
<td>-4638.34</td>
</tr>
<tr>
<td>0.25</td>
<td>49.09118</td>
<td>3956.898</td>
<td>0.25</td>
<td>25.67463</td>
<td>10828.42</td>
</tr>
<tr>
<td>0.5</td>
<td>230.6768</td>
<td>20430.8</td>
<td>0.5</td>
<td>152.1077</td>
<td>52648.23</td>
</tr>
<tr>
<td>0.75</td>
<td>801.9841</td>
<td>77237.31</td>
<td>0.75</td>
<td>253.8557</td>
<td>199193.8</td>
</tr>
<tr>
<td>1</td>
<td>1.758509 \times 10^{-18}</td>
<td>-1.162891 \times 10^{-18}</td>
<td>1</td>
<td>6.981793 \times 10^{-17}</td>
<td>-1.024914 \times 10^{-18}</td>
</tr>
<tr>
<td>0</td>
<td>-89.83543</td>
<td>-6504.072</td>
<td>0</td>
<td>-49.06056</td>
<td>-13552.7</td>
</tr>
<tr>
<td>0.25</td>
<td>127.4137</td>
<td>10269.93</td>
<td>0.25</td>
<td>90.58807</td>
<td>28214.74</td>
</tr>
<tr>
<td>0.5</td>
<td>610.4267</td>
<td>54064.86</td>
<td>0.5</td>
<td>406.5621</td>
<td>140721.2</td>
</tr>
<tr>
<td>0.75</td>
<td>2129.289</td>
<td>205067.1</td>
<td>0.75</td>
<td>1406.525</td>
<td>534824.8</td>
</tr>
<tr>
<td>1</td>
<td>4.674266 \times 10^{-19}</td>
<td>-3.091063 \times 10^{-18}</td>
<td>1</td>
<td>1.877341 \times 10^{-18}</td>
<td>-2.7559 \times 10^{-18}</td>
</tr>
</tbody>
</table>

Comparison of SLE with SE based on $d$.

<table>
<thead>
<tr>
<th>$n\Delta$</th>
<th>$g_1(\Delta)$</th>
<th>$d_2(\Delta)$</th>
<th>$n\Delta$</th>
<th>$g_1(\Delta)$</th>
<th>$d_2(\Delta)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.7284132</td>
<td>1.344312</td>
<td>0</td>
<td>-0.05702634</td>
<td>-0.08688098</td>
</tr>
<tr>
<td>5</td>
<td>0.7045509</td>
<td>0.7826991</td>
<td>5</td>
<td>-0.177097</td>
<td>-0.1676672</td>
</tr>
<tr>
<td>10</td>
<td>0.676467</td>
<td>0.6883977</td>
<td>10</td>
<td>-0.206038</td>
<td>-0.1789326</td>
</tr>
<tr>
<td>20</td>
<td>0.6465677</td>
<td>0.6212514</td>
<td>20</td>
<td>-0.2270366</td>
<td>-0.1861139</td>
</tr>
</tbody>
</table>

Tables. From Tables 3, 6 and 7 it is also seen that there is no need for the values of $\Delta_i(.)$ to be positive all the time. One should also note that the non-centrality parameter is always positive. Thus at the moment that $\Delta_i$ is negative, the relation $\Delta > \Delta_i(\Delta, d)$ means a positive value. Tables 4, 5 and 8 also give the comparisons based on $d$. There is no specific trend in superiority conditions, as one may expect comparing to the presented graphs, since superiority changes if we fix the sign of functions in the tables. We again remind that solving underlying non-linear equations, do not gives solution of the same sign always.

It should be noted that the above results are based on the hypothesis $H_\beta \neq h$.

5. Summary and conclusions

We have combined the idea of the preliminary test and the Stein-rule estimator with the Liu [37] estimator to obtain a better estimator for the regression parameters $\beta$ in a multiple linear regression model. We assumed a general assumption for the vector of errors of the model, which is distributed according to the law belonging to the class of elliptically contoured distributions (ECDs). Accordingly, we considered the following five Liu type estimators, namely, unrestricted Liu estimator (ULE), restricted Liu estimator (RLE), preliminary test Liu estimator (PTLE), shrinkage Liu estimator (SLE) and finally, positive rule Liu estimator (PRLE) for estimating the parameters ($\beta$) when it is suspected that the parameter $\beta$ may belong to a linear subspace defined by $H_\beta = h$. The performance of the estimators is compared based on the quadratic risk functions under both null and alternative hypotheses, which specify certain restrictions on the regression parameters. Under the restriction $H_\beta$, RLE performed the best compared to other estimators, however, it performed the worst when $\Delta$ moves away from its origin. Note that the risk of ULE is constant while the risk of RLE is unbounded as $\Delta$ goes to infinity. Also under $H_\beta$, the risk of PTLE is smaller than the risk of SLE and PRLE for $\alpha$ satisfying (3.19) for $q \geq 3$. Thus, neither PTLE nor PRLE nor SLE dominates each other uniformly. Note that the application of PRLE and SLE is constrained by the requirement $q \geq 3$, while PTLE does not need such a constraint. Nevertheless, the choice of the level of significance of the test has a significant impact on the nature of the risk function for the PTLE estimator. Thus, when $q \geq 3$, one would use PRLE; otherwise one uses PTLE with some optimum size $\alpha$. The findings of the paper are more general in the sense that many distributions, namely, multivariate normal, Kotz Type, Pearson Type II & VII, multivariate Student’s t, multivariate Cauchy etc. belong to the ECD class.

This work can be further extended for the elliptically contoured distribution. The five Liu type estimators can be compared with the corresponding ridge type estimators when the errors of the model follow an elliptically contoured distribution.
Comparison of PTLE with SLE based on $\Delta$.

<table>
<thead>
<tr>
<th>$n$</th>
<th>$\Delta$</th>
<th>$d$</th>
<th>$\alpha = 0.05$</th>
<th>$\alpha = 0.1$</th>
<th>$\alpha = 0.15$</th>
<th>$\alpha = 0.25$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\Delta_1(d, \alpha)$</td>
<td>$\Delta_2(d, \alpha)$</td>
<td>$\Delta_3(d, \alpha)$</td>
<td>$\Delta_1(d, \alpha)$</td>
<td>$\Delta_2(d, \alpha)$</td>
<td>$\Delta_3(d, \alpha)$</td>
</tr>
<tr>
<td>0</td>
<td>$-1.121395$</td>
<td>$-3.029414$</td>
<td>$-1.209226$</td>
<td>$-3.021498$</td>
<td>$-1.232199$</td>
<td>$-3.0789$</td>
</tr>
<tr>
<td>0.25</td>
<td>$0.235474$</td>
<td>$1.137792$</td>
<td>$0.116279$</td>
<td>$0.561845$</td>
<td>$-0.00781026$</td>
<td>$-0.03773831$</td>
</tr>
<tr>
<td>0.5</td>
<td>$0.8430094$</td>
<td>$6.997624$</td>
<td>$0.6790488$</td>
<td>$5.363662$</td>
<td>$0.5196657$</td>
<td>$4.316233$</td>
</tr>
<tr>
<td>0.75</td>
<td>$1.1462931$</td>
<td>$14.2397$</td>
<td>$0.9603494$</td>
<td>$11.9289$</td>
<td>$0.7837334$</td>
<td>$9.735856$</td>
</tr>
<tr>
<td>1</td>
<td>$1.324191$</td>
<td>$22.6064$</td>
<td>$1.125519$</td>
<td>$21.9471$</td>
<td>$0.9388979$</td>
<td>$16.03027$</td>
</tr>
<tr>
<td>5</td>
<td>$0.9530306$</td>
<td>$12.70271$</td>
<td>$1.532856$</td>
<td>$12.72388$</td>
<td>$1.479357$</td>
<td>$12.2798$</td>
</tr>
</tbody>
</table>

Comparison of PTLE with SLE based on $d$.

<table>
<thead>
<tr>
<th>$n$</th>
<th>$\Delta$</th>
<th>$\alpha = 0.05$</th>
<th>$\alpha = 0.1$</th>
<th>$\alpha = 0.15$</th>
<th>$\alpha = 0.25$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$g_1(d, \alpha)$</td>
<td>$d_1(d, \alpha)$</td>
<td>$g_2(d, \alpha)$</td>
<td>$d_2(d, \alpha)$</td>
<td>$g_3(d, \alpha)$</td>
</tr>
<tr>
<td>10</td>
<td>$-0.02341583$</td>
<td>$-187.9093$</td>
<td>$-0.02190736$</td>
<td>$-177.8385$</td>
<td>$-0.02032495$</td>
</tr>
<tr>
<td>5</td>
<td>$0.1255851$</td>
<td>$-223.9318$</td>
<td>$0.00917808$</td>
<td>$-207.952$</td>
<td>$-0.007004092$</td>
</tr>
<tr>
<td>10</td>
<td>$-0.006502562$</td>
<td>$-253.9273$</td>
<td>$0.003795436$</td>
<td>$-231.7729$</td>
<td>$-0.002621459$</td>
</tr>
<tr>
<td>20</td>
<td>$-0.001517084$</td>
<td>$-381.1839$</td>
<td>$0.005953986$</td>
<td>$-347.8945$</td>
<td>$-0.0003835455$</td>
</tr>
<tr>
<td>10</td>
<td>$0.005271143$</td>
<td>$-713.894$</td>
<td>$0.003505635$</td>
<td>$-607.4777$</td>
<td>$-0.01915946$</td>
</tr>
<tr>
<td>5</td>
<td>$-0.00120518$</td>
<td>$140.347$</td>
<td>$0.0009254258$</td>
<td>$-95.2964$</td>
<td>$-0.000292847$</td>
</tr>
<tr>
<td>10</td>
<td>$9.5398619$</td>
<td>$5278.518$</td>
<td>$7.670124$</td>
<td>$278.1428$</td>
<td>$0.0002421222$</td>
</tr>
<tr>
<td>20</td>
<td>$0.0003028185$</td>
<td>$85.0757$</td>
<td>$0.001272575$</td>
<td>$36.52423$</td>
<td>$3.458791$</td>
</tr>
</tbody>
</table>

Acknowledgments

The authors are thankful to the Associate Editor and the referees for their valuable comments and suggestions which certainly improved the presentation and quality of the paper. This paper was partially written while the second author was on sabbatical leave (2010–2011). He is grateful to Florida International University for awarding him the sabbatical
Table 5
Comparison of PRLE with PRE based on \( d \).

<table>
<thead>
<tr>
<th>( n )</th>
<th>( \Delta )</th>
<th>( \delta_1(\Delta) )</th>
<th>( d_2(\Delta) )</th>
<th>( n )</th>
<th>( \Delta )</th>
<th>( \delta_1(\Delta) )</th>
<th>( d_2(\Delta) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0</td>
<td>-0.5396458</td>
<td>-2.140968</td>
<td>0</td>
<td>9.022202</td>
<td>-0.08775683</td>
<td>-2.140968</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>-0.2828831</td>
<td>-2.596593</td>
<td>5</td>
<td>2.366221</td>
<td>-0.08775683</td>
<td>-2.596593</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>-0.245474</td>
<td>-5.56519</td>
<td>10</td>
<td>5.094712</td>
<td>-0.08775683</td>
<td>-5.56519</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>-0.2278089</td>
<td>-6.089953</td>
<td>20</td>
<td>-0.412356</td>
<td>-0.08775683</td>
<td>-6.089953</td>
</tr>
</tbody>
</table>

Table 6
Comparison of PRLE with PRE based on \( \Delta \).

<table>
<thead>
<tr>
<th>( n )</th>
<th>( \Delta )</th>
<th>( d )</th>
<th>( \Delta_1(\Delta, d) )</th>
<th>( \Delta_2(\Delta, d) )</th>
<th>( n )</th>
<th>( \Delta )</th>
<th>( d )</th>
<th>( \Delta_1(\Delta, d) )</th>
<th>( \Delta_2(\Delta, d) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0</td>
<td>0.25</td>
<td>-3.928908</td>
<td>-245.9488</td>
<td>10</td>
<td>0</td>
<td>0.25</td>
<td>-0.02745928</td>
<td>-7.585467</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.75</td>
<td>12.08853</td>
<td>1134.438</td>
<td>0</td>
<td>0.5</td>
<td>0.25</td>
<td>-3.704924</td>
<td>-1282.366</td>
</tr>
<tr>
<td></td>
<td>0.75</td>
<td>1</td>
<td>46.83421</td>
<td>4510.499</td>
<td>0.75</td>
<td>1</td>
<td>0.25</td>
<td>-11.56615</td>
<td>-4397.976</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>1.040252 × 10^{-17}</td>
<td>-6.879125 × 10^{-16}</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1.040252 × 10^{-17}</td>
<td>-6.879125 × 10^{-16}</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0.25</td>
<td>-20.12822</td>
<td>-1457.28</td>
<td>0</td>
<td>0</td>
<td>0.25</td>
<td>1.923968</td>
<td>531.4852</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.75</td>
<td>46.84462</td>
<td>377.5818</td>
<td>0.75</td>
<td>0.75</td>
<td>1</td>
<td>1.780175</td>
<td>-554.457</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0</td>
<td>5.012602 × 10^{-18}</td>
<td>-3.314082 × 10^{-17}</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>5.012602 × 10^{-18}</td>
<td>-3.314082 × 10^{-17}</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0.25</td>
<td>-47.02088</td>
<td>-3404.739</td>
<td>0</td>
<td>0</td>
<td>0.25</td>
<td>6.732456</td>
<td>1859.802</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.75</td>
<td>7.712369</td>
<td>621.6403</td>
<td>0.75</td>
<td>0.75</td>
<td>1</td>
<td>2.604868</td>
<td>-811.3174</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0</td>
<td>124.7178</td>
<td>11046.13</td>
<td>10</td>
<td>0</td>
<td>0.5</td>
<td>22.56909</td>
<td>-781.179</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0.25</td>
<td>-127.8159</td>
<td>-9255.298</td>
<td>0</td>
<td>0</td>
<td>0.25</td>
<td>29.28546</td>
<td>8089.939</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.75</td>
<td>143.3743</td>
<td>1158.612</td>
<td>0.75</td>
<td>1</td>
<td>0.25</td>
<td>3.874245</td>
<td>-1206.68</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0</td>
<td>318.4766</td>
<td>28207.14</td>
<td>20</td>
<td>0</td>
<td>0.5</td>
<td>74.57785</td>
<td>-25813.24</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0.25</td>
<td>-281.511</td>
<td>121262.1</td>
<td>0.25</td>
<td>0</td>
<td>0.5</td>
<td>292.9694</td>
<td>-111377.5</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>1</td>
<td>2.878001 × 10^{-19}</td>
<td>-1.903204 × 10^{-18}</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2.878001 × 10^{-19}</td>
<td>-1.903204 × 10^{-18}</td>
</tr>
</tbody>
</table>

leave which gave him excellent research facilities. For the first and third authors, this research was partially supported by Shahrood University of Technology, Iran.

Appendix

If \( e \) has a density, then it is of the form

\[
f(e) \propto \sigma^2 V^{-\frac{1}{2}} g \left( \frac{1}{\sigma^2} e V^{-\frac{1}{2}} \right)
\]

where \( g(.) \) is a non-negative function over \( \mathbb{R}^+ \) such that \( f(.) \) is a density function w.r.t (with respect to) a \( \sigma \)-finite measure \( \mu \) on \( \mathbb{R}^p \). In this case, the notation \( e \sim E_{\mu}(0, \sigma^2 V, g) \) would probably be used.

It is sometimes difficult to have complete analysis of the regression model with ECD errors of type (A.1). 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.

The general formula for the above mixture of distributions is given by

\[
f_e(x) = \int_0^\infty W(t) \phi_{N_0(0, t^{-1} \sigma^2 V)}(x) dt,
\]

where \( \phi_{N_0(0, t^{-1} \sigma^2 V)}(.) \) is the pdf (probability density function) of \( N_0(0, t^{-1} \sigma^2 V) \).

(a) If

\[
W(t) = 2 (t^\gamma \Gamma(\gamma/2))^{-1} \left( \frac{\gamma \sigma^2}{2} \right)^{\gamma/2} t^{-\gamma+1} e^{-\frac{\gamma^2}{2t^2}}, \quad 0 < \gamma, \sigma^2, t < \infty
\]
Table 7

Comparison of PRLE with PTLE based on $\Delta$.

<table>
<thead>
<tr>
<th>$n$</th>
<th>$\Delta$</th>
<th>$d$</th>
<th>$\alpha = 0.05$</th>
<th>$\alpha = 0.1$</th>
<th>$\alpha = 0.15$</th>
<th>$\alpha = 0.25$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$\Delta g(\Delta, \alpha)$</td>
<td>$d(\Delta, \alpha)$</td>
<td>$\Delta g(\Delta, \alpha)$</td>
<td>$d(\Delta, \alpha)$</td>
</tr>
<tr>
<td>0</td>
<td>1.05</td>
<td>0.01</td>
<td>0.9904512</td>
<td>0.3898041</td>
<td>0.9880952</td>
<td>0.3898041</td>
</tr>
<tr>
<td>0.25</td>
<td>0.55</td>
<td>0.01</td>
<td>0.9880952</td>
<td>0.3898041</td>
<td>0.9880952</td>
<td>0.3898041</td>
</tr>
<tr>
<td>0.75</td>
<td>0.25</td>
<td>0.01</td>
<td>0.9880952</td>
<td>0.3898041</td>
<td>0.9880952</td>
<td>0.3898041</td>
</tr>
<tr>
<td>1</td>
<td>0.01</td>
<td>0.01</td>
<td>0.9880952</td>
<td>0.3898041</td>
<td>0.9880952</td>
<td>0.3898041</td>
</tr>
<tr>
<td>1.5</td>
<td>0.01</td>
<td>0.01</td>
<td>0.9880952</td>
<td>0.3898041</td>
<td>0.9880952</td>
<td>0.3898041</td>
</tr>
<tr>
<td>2</td>
<td>0.01</td>
<td>0.01</td>
<td>0.9880952</td>
<td>0.3898041</td>
<td>0.9880952</td>
<td>0.3898041</td>
</tr>
<tr>
<td>2.5</td>
<td>0.01</td>
<td>0.01</td>
<td>0.9880952</td>
<td>0.3898041</td>
<td>0.9880952</td>
<td>0.3898041</td>
</tr>
<tr>
<td>3</td>
<td>0.01</td>
<td>0.01</td>
<td>0.9880952</td>
<td>0.3898041</td>
<td>0.9880952</td>
<td>0.3898041</td>
</tr>
</tbody>
</table>

Table 8

Comparison of PRLE with PTLE based on $d$.

<table>
<thead>
<tr>
<th>$n$</th>
<th>$\Delta$</th>
<th>$\alpha = 0.05$</th>
<th>$\alpha = 0.1$</th>
<th>$\alpha = 0.15$</th>
<th>$\alpha = 0.25$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$g(\Delta, \alpha)$</td>
<td>$d(\Delta, \alpha)$</td>
<td>$g(\Delta, \alpha)$</td>
<td>$d(\Delta, \alpha)$</td>
</tr>
<tr>
<td>0</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>0.25</td>
<td>0.55</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>0.75</td>
<td>0.25</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>1</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
</tbody>
</table>

then we have

$$f(e) = \frac{\Gamma \left( \frac{n+1}{2} \right) |V|^{-\frac{1}{2}}}{(\pi \gamma)^{n/2} \Gamma(n/2)} e^{-\frac{1}{2} \left( \frac{e^2}{\sigma^2} \right)}$$

where $E(e) = 0$ and $E(e^2) = \frac{{n+1}\sigma^2}{\gamma^2}$, $V = \sigma^2$ for $\gamma > 2$. 

Remark 3. Regarding the above classifications, we should take the following notes:

1. In all the above classes we have
   \[ \Sigma_e = -2\psi'(0)\sigma^2 V = \left( \int_0^\infty t^{-1}W(t)dt \right) \sigma^2 V, \]
   resulting in \[-2\psi'(0) = \int_0^\infty t^{-1}W(t)dt.\]
2. The subclass (a) is neither contained in the subclass (b) nor in the subclass (c). However, the 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 \(-2\psi'(0) = \int_0^\infty t^{-1}W(t)dt\) 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 Type II & VII, multivariate Student’s t, multivariate Cauchy, Logistic, Bessel and generalized slash distributions. Dating back to [32], 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. [15], Muirhead [41], Fang et al. [19] and Gupta and Varga [25] among others. More details on this topics are available in [8].

(b) Chu [17] considered

\[ W(t) = (2\pi)^{\frac{d}{2}} |\sigma^2 V|^{\frac{1}{2}} t^{-\frac{d}{2}} \mathcal{L}^{-1}[f(s)], \]  

(A.5)

\(\mathcal{L}^{-1}[f(s)]\) denotes the inverse Laplace transform of \(f(s)\) with \(s = [x'(\sigma^2 V)^{-1}x]/2\). For some examples of \(f(.)\) and \(W(.)\) see [11].

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 Bhatta [18] for more specific details.

The mean of \(e\) is the zero-vector and the covariance-matrix of \(e\) is

\[
\Sigma_e = \text{Cov}(e) = \int_0^\infty \text{Cov}(e|t)W(t)dt
\]

\[= \int_0^\infty W(t)\text{Cov}\left\{N_p(0, t^{-1}\sigma^2 V)\right\} dt
\]

\[= \left( \int_0^\infty t^{-1}W(t)dt \right) \sigma^2 V, \]  

(A.6)

provided the above integral exists.

Comparing the models (A.1) and (A.2), since \(\Sigma_e = \text{Cov}(e) = -2\psi'(0)\sigma^2 V\), using (A.6) we can conclude that

\[-2\psi'(0) = \int_0^\infty t^{-1}W(t)dt.\]

Now suppose that \(X \sim E_n(\mu, \sigma^2 V, g)\). Then it is important to point out that since \(\int f(x)dx = 1\), using Fubini’s theorem one can show that \(\int_0^\infty W(t)dt = 1\). Thus for nonnegative function \(W(.)\), it is a density. For nonnegative function \(W(.)\), the elliptical models can be interpreted as a scale mixture of normal distributions.

(c) Srivastava and Bilodeau [52] considered the signed measure, \(W(t)\) such that

(i) \(\int_0^\infty t^{-1}W^+(dt) < \infty\),

(ii) \(\int_0^\infty t^{-1}W^-(dt) < \infty\),

(A.7)

where \(W^+ - W^-\) is the Jordan decomposition of \(W\) in positive and negative parts. Note that from (i)–(ii) of (A.7),

\[\int_0^\infty t^{-1}W(dt) < \infty\]  

(A.8)

and thus, \(\text{Cov}(e)\) exists under the sub-class defined above. This subclass contains the subclass defined by (b).
References


