Journal of Statistics & Management Systems
ISSN 0972-0510 (Print), ISSN 2169-0014 (Online)
Vol. 27 (2024), No. 5, pp. 877–910
DOI : 10.47974/JSMS-1032

Robust M-estimation for linear regression models : Adaptive aspect

Sanaa M. El-Gayar
Amani A. Ahmed *
Mahmoud A. Mahmoud
Department of Statistics
Faculty of Economics and Political Science
Cairo University
Egypt

Abstract

In this study, adaptive one-step generalized M-estimators for the regression parameters of the linear regression model are derived. The proposed estimators are developed assuming the exponential power family and the t-distribution family. A class of one-step M-estimators with likelihood score function is obtained then the estimators are adapted to the distribution. A Monte Carlo simulation study is conducted to investigate the performance of the suggested estimators. The results show that the proposed adaptive robust estimators are highly efficient than least squares estimators as well as robust counterparts. The results show also that the suggested method for estimating the shape parameter through its relation with the kurtosis coefficient gives efficient estimators.

Subject Classification: *(2010) 62F35.*

Keywords: *Exponential power distribution, t-Distribution, One-step M-estimation, Kurtosis coefficient, Robust weights.*

1. Introduction

From a long time, statisticians have recognized that optimal procedures provided by the classical statistical analysis under exact parametric models lose their optimality for even small deviations from the assumed model. This stimulated statisticians to develop stable methods against large families of distributions obtained by adding new shape parameters to the parametric models. This involves various types ranging from moderate to heavy tailed distributions. That is, developing robust statistical procedures that are valid for

* *E-mail:* a.abubakr@feps.edu.eg

© Taru Publications

large family of distributions and are insensitive to deviations from specific ones [16].

In the 1960's, two theoretical approaches of robust procedures were introduced and robustness measures were basically developed from them. First, Huber introduced the minimax approach and developed M-estimators by minimizing the maximum asymptotic variance over a set of contaminated family of distributions. Then Hampel introduced the influence function (IF) approach and developed M-estimators that minimize the asymptotic variance at a central model subject to a bound on the gross errors sensitivity [30].

Measures of robustness were developed based on these two approaches such as having bounded IF and high breakdown point. For M-estimators, the IF is proportional to the score function ψ. They are robust if ψ is bounded and continuous. All other criteria can be measured relative to the kind and behavior of ψ [18]. Extensions of robust estimation for linear models were then introduced [2, 25, 29, 7, 32, 35, 19, 26 and 10].

In general, robust estimators are not efficient. However, the possibility to find asymptotically efficient estimators under a large class of sufficiently smooth symmetric distributions was first discussed in a general set up by [31]. Estimators of this kind are known in the literature as adaptive estimators. Enhancing efficiency of robust estimators can be achieved by implementing different sorts of adaptive techniques [15, 24, 5, 33, 24, 14 and 21].

The aim of this work is to develop highly efficient robust estimators for the linear regression model. Adaptive M- and generalized M-estimators for the regression parameters are suggested assuming the extended power distribution (EPD) family. This distribution family contains both exponential power and the student t-families of distributions [1 and 22]. To formulate the problem, consider the ordinary linear regression model given by:

$$Y_i = \boldsymbol{x}_i^t \boldsymbol{\theta} + E_i \, , i = 1,2, \dots , n, \tag{1.1}$$

where, Y_i's are n observations of response variables, $\boldsymbol{x}_i$ is a q-vector of $(q-1)$ explanatory variables with $x_{i1} = 1, \forall i$, $\boldsymbol{\theta}$ is a q-vector of unknown regression parameters. $E_i's$ are independent and identically distributed random variables having a common unknown distribution function F with mean zero and variance σ^2[23 and 13]. An M-estimator of the regression parameter $\boldsymbol{\theta}$ is the solution of the minimization problem: $min_\theta \sum_{i=1}^n \rho(Y_i - \boldsymbol{x}_i^t \boldsymbol{\theta})$ with $\rho(.)$ a convex function. Assuming ρ is a derivable function with derivative ψ, an M-estimator of $\boldsymbol{\theta}$ becomes equivalently the solution in $\boldsymbol{\theta}$ of the equation $\sum_{i=1}^n \boldsymbol{x}_i \psi(Y_i - \boldsymbol{x}_i^t \boldsymbol{\theta}) = \mathbf{0}$. For example the least squares (LS) estimator corresponds to $\rho(t) = t^2$, while the least absolute deviations (LAD) estimator corresponds to $\rho(t) = |t|$ [9].

The IF of an M-estimator is bounded in Y whenever ψ is a bounded function, but necessarily unbounded in $\boldsymbol{x}$. Outliers in the **X**-space (leverage

points problem) may render the value of the IF very large and extremely reduce all robustness properties. Generalized M-estimators (GM) were then developed by Mallows in 1975 [16] to overcome such a problem. For the GM estimation, weight functions are incorporated in the estimating equation to limit the bad effects of leverages. Mallows suggested a GM-estimate as the solution of the minimization problem: $min_\theta \sum_{i=1}^{n} w_i \, \rho\left(\frac{Y_i - x_i^t \boldsymbol{\theta}}{\sigma}\right)$. The weight functions w_i are related to the Mahalanobis distance in the **X**-space [3], defined as:

$$w_i = min\left[1, \left(\frac{b}{(x_i - \boldsymbol{m_x})^t \boldsymbol{C_x^{-1}} (x_i - \boldsymbol{m_x})}\right)^{\frac{\alpha'}{2}}\right], \tag{1.2}$$

where, b is an upper percentile of a chi-squared distribution with $(q-1)$ degrees of freedom, $\boldsymbol{m_x}$ and $\boldsymbol{C_x}$ are the classical multivariate location and scatter estimates for $\boldsymbol{x_i}$ given by: $\boldsymbol{m_x} = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{x_i}$, $\boldsymbol{C_x} = \frac{1}{n-1}\sum_{i=1}^{n} (\boldsymbol{x_i} - \boldsymbol{m_x})(\boldsymbol{x_i} - \boldsymbol{m_x})^t$, and $\alpha' \geq 0$. On the other hand, Schweppe (1975) [21] suggested another formula of GM-estimator to improve the efficiency and the robustness level of the resulted estimator. This is obtained as the solution of: $min_\theta \sum_{i=1}^{n} w_i^2 \, \rho\left(\frac{Y_i - x_i^t \boldsymbol{\theta}}{w_i \sigma}\right)$, such that high leverage points are down weighted only if the residuals are large.

It has been shown that obtaining exact solutions to the above mentioned estimating equations is not an easy problem. [4] suggested a good approximation to Huber M-estimate by performing a one-step Newton-Raphson iteration, using consistent initial estimate $\widehat{\boldsymbol{\theta}}_{\boldsymbol{0}}$of $\boldsymbol{\theta}$**.** One-step GM-estimators were then rapidly developed using robust initials and robust weights. Different robust weights can be found based on high breakdown multivariate location and scatter estimates of the predictors such as the minimum volume ellipsoid (MVE) and the minimum covariance determinants (MCD) estimators [34]. Moreover, different types of high breakdown point initial estimates can be also distinguished: the least median of squares (LMS), the least trimmed of squares (LTS) and LAD estimator [26]. Although these initials have high breakdown point, they are poor in terms of efficiency measures under the normal distribution compared with LS estimator [10].

In general, one-step GM-estimator through Newton-Raphson iteration using robust initial estimator $\widehat{\boldsymbol{\theta}}_{\boldsymbol{0}}$ and robust weight functions w_i, v_i [32] can be formulated as:

$$\widehat{\boldsymbol{\theta}} = \widehat{\boldsymbol{\theta}}_{\boldsymbol{0}} + \widehat{\boldsymbol{H}}_{\boldsymbol{0}}^{-\boldsymbol{1}} \widehat{\boldsymbol{g}}_{\boldsymbol{0}} \,, \tag{1.3}$$

where, $\widehat{\boldsymbol{g}}_{\boldsymbol{0}} = \sum_{i=1}^{n} \boldsymbol{x_i} w_i \psi\left(\frac{v_i r_i}{\hat{\sigma}}\right)$, $\widehat{\boldsymbol{H}}_{\boldsymbol{0}} = \frac{1}{\hat{\sigma}}\sum_{i=1}^{n} \boldsymbol{x_i x_i^t} w_i v_i \psi'\left(\frac{v_i r_i}{\hat{\sigma}}\right)$, $r_i = y_i - \boldsymbol{x_i^t \widehat{\theta}_0}$**,** and $\hat{\sigma}$ is a robust estimate of σ.

In this study, assuming the extended power family, a sort of adaptive one-step M- and GM- estimators of $\boldsymbol{\theta}$ are suggested and investigated. First, a class of

one-step M- and GM-estimators with likelihood score function is derived. The kurtosis measure of this distribution family is examined and shown to depend only on the shape parameter in a decreasing manner. The idea is thus to adapt the IF to the underlying distribution by estimating the shape parameter using empirical kurtosis counterpart calculated using the residuals. A Monte Carlo simulation study is conducted to assess the performance of the estimators. The study considers different combinations of consistent robust initials, robust weights of the Mallows and Schweppe types assuming the existence of leverage points or not. The results show that the suggested estimators are highly more efficient than their initials as well as LS estimator especially for heavy tailed distributions. Moreover, the suggested method of estimating the shape parameter is highly efficient.

Accordingly, this paper is organized as follows: Section 2 presents the extended power family of distributions. Section 3 is devoted to formulating the one-step GM-estimators, and deriving the adaptive versions. In Section 4, Monte Carlo simulation study is described and its main results are given. Two illustrative examples with real data sets are given in Section 5. Finally, the conclusions are given in Section 6.

2. Extended power family of distributions

The probability density function (pdf) of the EPD [1] is given by:

$$f(y|\mu,\varphi,\lambda,c) = k(c,\lambda)\varphi^{\frac{1}{2}} \exp\left[-\frac{1}{2} c\, Q_\lambda\left(1+\frac{\varphi(y-\mu)^2}{c-1}\right)\right], \quad (2.1)$$
$$-\infty < y,\mu < \infty, \varphi > 0, \lambda \geq 0, c > 1,$$

Where,

$$Q_\lambda(\upsilon) = \begin{cases} \dfrac{\upsilon^\lambda - 1}{\lambda}, \lambda > 0, \\ \lim\limits_{\lambda\to 0} \dfrac{\upsilon^\lambda - 1}{\lambda} = \log \upsilon, \lambda = 0. \end{cases}$$

and $k(c,\lambda)$is the normalizing constant,

The distribution is symmetric about a location parameter μ, having a scale parameter φ and two shape parameters λ andc. According to the λvalues, the distributions are distinguished in tail length. Sharper tails correspond to $\lambda > 1$ while heavier tails correspond to $0 \leq \lambda < 1$. In particular at $\lambda = 1$, the distribution is exactly the normal distribution with mean μ and variance $\frac{c-1}{c\varphi}$. When $\lambda = 0$, the distribution is the student's t-distribution with $(c-1)$ degrees of freedom. Clearly, for $\lambda > 0$, the density function is approximately proportional to $exp(-D|y-\mu|^{2\lambda})$. This means that the distribution is close to the well-known exponential power (EP) distribution which is very useful in robustness studies.

2.1. *Moments and kurtosis for the EP distribution*

Consider the EP distribution reduced from the family in (2.1). Its pdf is given by:

$$f(y|\mu,\varphi,\lambda,c)=\left(2^{\frac{1}{2\lambda}+1}\Gamma\left(\frac{1}{2\lambda}+1\right)\right)^{-1}\sqrt{\frac{c\varphi}{\lambda(c-1)}}\exp-\frac{1}{2}\left|\sqrt{\frac{c\varphi}{\lambda(c-1)}}(y-\mu)\right|^{2\lambda},$$
$$-\infty<y<\infty,-\infty<\mu<\infty,\varphi,\lambda>0,c>1. \quad (2.2)$$

Clearly, the distribution is symmetric about the location parameter μ, having scale parameters c,φ and a shape parameter λ. For $\lambda=1$, the distribution is the normal distribution. For $\lambda=\frac{1}{2}$, the distribution is the double exponential (Laplace) distribution with mean μ and variance $\frac{4(c-1)}{c\varphi}$. The distribution converges to the uniform distribution as λ→∞.

Due to symmetry of the EP distribution about μ, all central moments of odd orders are exactly zero. As for those of even orders, it can be easily shown that the $2r^{th}$ central moment, $r=1,2,3$ ..., is given by:

$$E(Y-\mu)^{2r}=\left[2^{\frac{1}{\lambda}}\frac{\lambda(c-1)}{c\varphi}\right]^{r}\frac{\Gamma\left(\frac{2r+1}{2\lambda}\right)}{\Gamma\left(\frac{1}{2\lambda}\right)},\varphi,\lambda>0,c>1. \quad (2.3)$$

If $r=1$, the variance σ^2 becomes:

$$\sigma^2=2^{\frac{1}{\lambda}}\lambda\left(\frac{c-1}{c\varphi}\right)\frac{\Gamma\left(\frac{3}{2\lambda}\right)}{\Gamma\left(\frac{1}{2\lambda}\right)},\varphi,\lambda>0,c>1. \quad (2.4)$$

If $r=2$, the fourth central moment becomes:

$$E(Y-\mu)^4=2^{\frac{2}{\lambda}}\lambda^2\left(\frac{c-1}{c\varphi}\right)^2\frac{\Gamma\left(\frac{5}{2\lambda}\right)}{\Gamma\left(\frac{1}{2\lambda}\right)},\varphi,\lambda>0,c>1. \quad (2.5)$$

The kurtosis is thus:

$$K(\lambda)=\frac{\Gamma\left(\frac{1}{2\lambda}\right)\Gamma\left(\frac{5}{2\lambda}\right)}{\left(\Gamma\left(\frac{3}{2\lambda}\right)\right)^2},\lambda>0. \quad (2.6)$$

Evidently, the kurtosis is only a function of λ regardless of all other parameters of the distribution.

Lemma 2.1: *Consider the kurtosis function* $K(\lambda)=\frac{\Gamma\left(\frac{1}{2\lambda}\right)\Gamma\left(\frac{5}{2\lambda}\right)}{\left(\Gamma\left(\frac{3}{2\lambda}\right)\right)^2},\lambda>0.$ *$K(\lambda)$ is a decreasing function of the shape parameter λ.*

Proof: [11] proved that, for any real numbers$(e, f \geq 1)$, the function $g(x) = \frac{\Gamma(ex)\Gamma(fx)}{\left(\Gamma\left(\frac{e+f}{2}x\right)\right)^2}$ is an increasing function in x over the interval $(0, \infty)$. So, by taking $e = 1, f = 5$ and $x = \frac{1}{2\lambda}$, the lemma follows immediately.

Clearly, for $0.5 \leq \lambda \leq 1$, the kurtosis function decreases from 6 at $\lambda = 0.5$ (Laplace distribution) to 3 at $\lambda = 1$ (normal distribution) [Figure (2.1)].

The density (2.2) can be written in a re-parameterized form as:

$$f(y|\mu, \sigma, \lambda) = \left(2^{\frac{1}{2\lambda}+1}\Gamma\left(\frac{1}{2\lambda}+1\right)\right)^{-1} \frac{\sqrt{B}}{\sigma} \exp -\frac{1}{2}\left|\sqrt{B}\left(\frac{y-\mu}{\sigma}\right)\right|^{2\lambda},$$
$$-\infty < y < \infty, -\infty < \mu < \infty, \sigma, \lambda > 0, \tag{2.7}$$

Where, $B = 2^{\frac{1}{\lambda}} \frac{\Gamma\left(\frac{3}{2\lambda}\right)}{\Gamma\left(\frac{1}{2\lambda}\right)}$.

2.2 *Moments and kurtosis for the t-distribution*

As $\lambda \to 0$, the pdf in (2.1) is reduced to the t- distribution given by:

$$f(y|\mu, \varphi, c) = \frac{\varphi^{\frac{1}{2}}\Gamma\left(\frac{c}{2}\right)}{\sqrt{\pi(c-1)}\Gamma\left(\frac{c-1}{2}\right)}\left[1 + \frac{\varphi}{c-1}(y-\mu)^2\right]^{-\frac{c}{2}},$$
$$-\infty < y < \infty, -\infty < \mu < \infty, \varphi > 0, c > 1. \tag{2.8}$$

The distribution is symmetric about the location parameter μ and has a scale parameter $\varphi^{-\frac{1}{2}}$ with $(c-1)$ degrees of freedom. It is well known that, the smaller the value of c, the heavier the tail of the distribution. Moreover, the distribution converges to the normal distribution as $c \to \infty$. For $c = 2$, the t-distribution becomes the Cauchy distribution.

It can be easily shown that the $2r^{th}$ central moment, $r = 1,2,3, \ldots$, of this distribution is given by:

$$E(Y-\mu)^{2r} = \left(\frac{c-1}{\varphi}\right)^r \frac{\beta\left(\frac{2r+1}{2}, \frac{c-2r-1}{2}\right)}{\beta\left(\frac{1}{2}, \frac{c-1}{2}\right)}, \text{ whenever } c > 2r+1. \tag{2.9}$$

If $r = 1$, the variance σ^2 becomes:

$$\sigma^2 = \frac{c-1}{\varphi(c-3)} = \frac{1}{\varphi}\left[1 + \frac{2}{c-3}\right], \text{ whenever } c > 3. \tag{2.10}$$

For any value of c and φ, the kurtosis is thus:

$$K(c) = 3 + \frac{6}{c-5}, \text{ whenever} > 5. \tag{2.11}$$

Clearly, $K(c)$ depends only on the shape parameter c in a decreasing manner. It decreases from the real number 9 as $c = 6$ to 3 as $c \to \infty$, which naturally corresponds to the limiting normal distribution [Figure (2.2)].

Moreover, the density function (2.8) can be re-parameterized in terms of $\sigma^2$and c as:

$$f(y|\mu,\sigma,c)=\frac{\Gamma\left(\frac{c}{2}\right)}{\sigma\sqrt{\pi(c-3)}\Gamma\left(\frac{c-1}{2}\right)}\left[1+\left(\frac{y-\mu}{\sigma\sqrt{c-3}}\right)^2\right]^{-\frac{c}{2}},$$

$$-\infty<y<\infty,-\infty<\mu<\infty,\sigma>0,c>3. \quad (2.12)$$

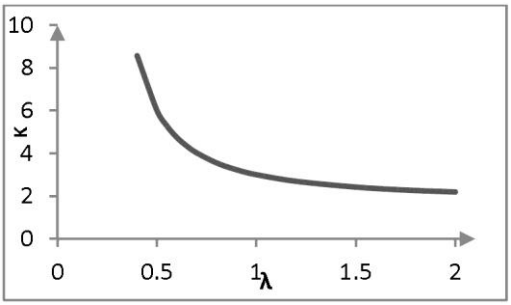

Figure 2.1
$K(\lambda)$ for EP family.

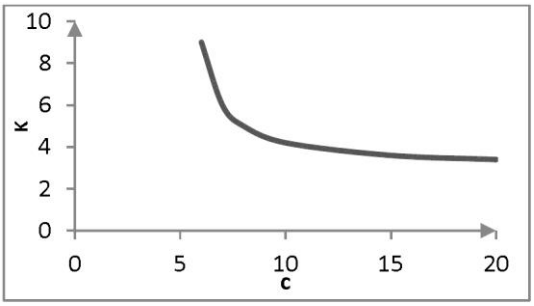

Figure 2.2
$K(c)$ for t-family.

Thus, the kurtosis function for the EP and the student's t- family is a one to one function of the shape parameter λ or c. That is, any distribution of them can be characterized by its kurtosis value. Actually, the kurtosis is a measure of tail heaviness of distributions rather than peakedness [28]. Its behavior in characterizing the distribution shape of either of the two families permits to possibly estimate the distribution shape parameter through kurtosis estimation using the sample.

3. Adaptive one-step robust M-estimators

In this study, adaptive one-step M- and GM-estimators of the vector $\boldsymbol{\theta}$ in the regression model given by (1.1) are proposed. A first stage to get such estimators is to derive their non-adaptive versions. In view of the general representation (1.3), we derive here a class of estimators that conciliating robust

consistent initials, good robust estimator $\hat{\sigma}^2$ of σ^2 and robust weights with the efficiency resulting by applying the likelihood scoring for ψ.

3.1. *The Estimators under the EP Family*

Assuming the pdf of the EP distribution in Equation (2.7) with $\mu = 0$, $0.5 < \lambda \leq 1$, and consider the function $\rho(e) = -\ln f(e|\sigma, \lambda)$.Then the first and second derivatives ψ and $\psi'$of ρ are given by:

$$\psi(e) = \lambda B^{\lambda} \left|\frac{e}{\sigma}\right|^{2\lambda-1} sign(e). \tag{3.1}$$

$$\psi'(e) = \lambda(2\lambda - 1)B^{\lambda} \left|\frac{e}{\sigma}\right|^{2\lambda-2}. \tag{3.2}$$

Clearly, ψ is well defined as a relevant score and it is continuous and its unboundedness for large values of e is compensated by being a likelihood score [Figures (3.1) and (3.2)]. But ψ' diverges to ∞ as e tends to zero. An adjustment of the residuals is needed here to avoid small values by considering $|r^*_i| = \max(0.01, |r_i|)$.

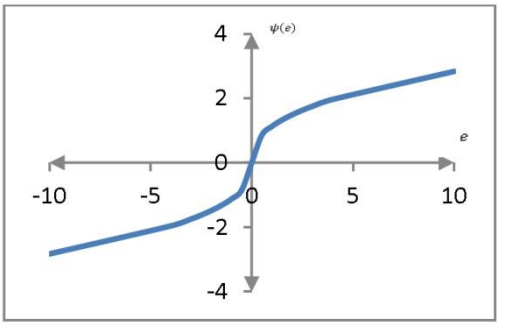

Figure 3.1
ψ function for $\lambda = 0.7, \sigma = 1$

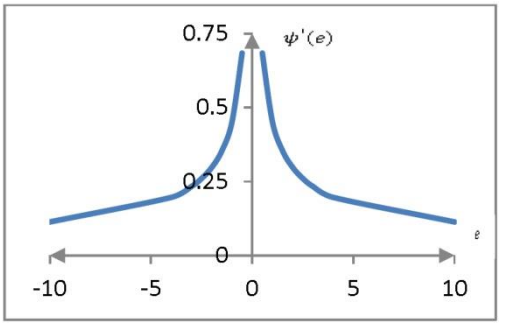

Figure 3.2
ψ' function for $\lambda = 0.7, \sigma = 1$

Returning to the functions $\widehat{\boldsymbol{g}}_{\boldsymbol{0}}$ and $\widehat{\boldsymbol{H}}_{\boldsymbol{0}}$in Equation (1.3) and applying such likelihood functions ψ and ψ', the suggested one-step M-estimators of $\boldsymbol{\theta}$ is thus:

$$\widehat{\boldsymbol{\theta}}_{\boldsymbol{M}} = \widehat{\boldsymbol{\theta}}_{\boldsymbol{0}} + \frac{\sum_{i=1}^{n} \boldsymbol{x_i} |r_i^*|^{2\lambda-1} sign(r_i^*)}{(2\lambda-1)\sum_{i=1}^{n} \boldsymbol{x_i x_i^t} |r_i^*|^{2\lambda-2}}. \tag{3.3}$$

Introducing weights to treat leverage points, the suggested Mallows-type one-step GM-estimators is:

$$\widehat{\boldsymbol{\theta}}_{\boldsymbol{WM}} = \widehat{\boldsymbol{\theta}}_{\boldsymbol{0}} + \frac{\sum_{i=1}^{n} \boldsymbol{x_i} w_i |r_i^*|^{2\lambda-1} sign(r_i^*)}{(2\lambda-1)\sum_{i=1}^{n} \boldsymbol{x_i x_i^t} w_i |r_i^*|^{2\lambda-2}}, \tag{3.4}$$

and the Scheweppe-type one-step GM-estimators, using $v_i = \frac{1}{w_i}$, is thus:

$$\widehat{\boldsymbol{\theta}}_{\boldsymbol{VM}} = \widehat{\boldsymbol{\theta}}_{\boldsymbol{0}} + \frac{\sum_{i=1}^{n} \boldsymbol{x_i} w_i \left|\frac{r_i^*}{w_i}\right|^{2\lambda-1} sign(r_i^*)}{(2\lambda-1)\sum_{i=1}^{n} \boldsymbol{x_i x_i^t} \left|\frac{r_i^*}{w_i}\right|^{2\lambda-2}}. \tag{3.5}$$

The adaptive estimators suggested in this study are obtained by replacing λ in Equations (3.3) to (3.5) by their estimators $\hat{\lambda}$. The idea is to derive $\hat{\lambda}$ through equating $K(\lambda)$ in Equation (2.6) to the sample kurtosis $(\hat{\gamma} + 3)$ computed using the corresponding residuals r_i.

3.2 *The estimators under the t-family*

Consider the pdf given in Equation (2.12) with $\mu = 0$. The functions ψ and ψ' have the forms:

$$\psi(e) = \frac{c}{\sqrt{c-3}} \frac{\frac{e}{\sigma\sqrt{c-3}}}{\left(1+\left(\frac{e}{\sigma\sqrt{c-3}}\right)^2\right)}, \tag{3.6}$$

$$\psi'(e) = \frac{c}{c-3} \frac{1-\left(\frac{e}{\sigma\sqrt{c-3}}\right)^2}{\left(1+\left(\frac{e}{\sigma\sqrt{c-3}}\right)^2\right)^2}, c > 3. \tag{3.7}$$

Clearly for some value of $c > 3, \psi$ is continuous and bounded (it is redescendent) and ψ' is a well behaving function [Figures (3.3) and (3.4)].

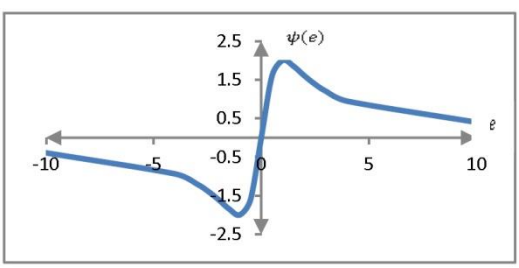

Figure 3.3
ψ function for $c = 4, \sigma = 1$

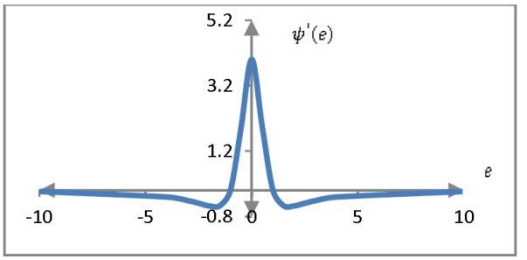

Figure 3.4
ψ' function for $c = 4, \sigma = 1$

Accordingly, putting $z_i = \frac{r_i}{\hat{\sigma}\sqrt{c-3}}$, the suggested one-step M-estimators is thus given by:

$$\widehat{\boldsymbol{\theta}}_{\boldsymbol{M}} = \widehat{\boldsymbol{\theta}}_{\boldsymbol{0}} + \frac{\hat{\sigma}\sqrt{c-3}\sum_{i=1}^{n} \boldsymbol{x_i}\left(\frac{z_i}{1+{z_i}^2}\right)}{\sum_{i=1}^{n} \boldsymbol{x_i x_i^t}\left(\frac{1-{z_i}^2}{(1+{z_i}^2)^2}\right)}. \tag{3.8}$$

The suggested of Mallows-type one-step GM-estimators is:

$$\widehat{\boldsymbol{\theta}}_{\boldsymbol{WM}} = \widehat{\boldsymbol{\theta}}_{\boldsymbol{0}} + \frac{\hat{\sigma}\sqrt{c-3}\sum_{i=1}^{n} \boldsymbol{x_i} w_i\left(\frac{z_i}{1+{z_i}^2}\right)}{\sum_{i=1}^{n} \boldsymbol{x_i x_i^t} w_i\left(\frac{1-{z_i}^2}{(1+{z_i}^2)^2}\right)}, \tag{3.9}$$

and the Scheweppe-type one-step GM-estimators, using $v_i = \frac{1}{w_i}$, is thus:

$$\widehat{\boldsymbol{\theta}}_{\boldsymbol{VM}} = \widehat{\boldsymbol{\theta}}_{\boldsymbol{0}} + \frac{\hat{\sigma}\sqrt{c-3}\sum_{i=1}^{n} \boldsymbol{x_i}\left(\frac{z_i}{1+(z_i/w_i)^2}\right)}{\sum_{i=1}^{n} \boldsymbol{x_i x_i^t}\left(\frac{1-(z_i/w_i)^2}{(1+(z_i/w_i)^2)^2}\right)}. \tag{3.10}$$

The adaptive estimators suggested in this study are obtained by replacing c in Equations (3.8) to (3.10) by its estimator $\hat{c}$. Actually, $\hat{c}$ is estimated by equating $K(c)$ in Equation (2.11) to the sample kurtosis $(\hat{\gamma} + 3)$ computed using the corresponding residuals r_i.

4. Simulation Study

4.1 *Data Generation*

For the data generation process, the linear model (1.1) is considered assuming $q = 3$ with two explanatory variables (X_2, X_3). We followed the approach of [8] in generating the data in two cases of explanatory variables. For the non-leverage points case, data are generated from the standard normal distribution. For leverage points case, we considered two different sources of leverage points. First, heavy tailed explanatory variables where explanatory variables are generated from the t-distribution with three degrees of freedom. Second, gross errors explanatory variables where explanatory variables are generated from a contaminated normal distribution with approximately 5% of observations in each of the two tails come from a normal distribution with mean

zero and variance 100. The BACON algorithm was applied to confirm the existence of outliers in the X-space [6]. Assuming $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3) = (1,2,3)$, and after generating the error term values of the EPD then the response variable is calculated directly from Equation (1.1).

4.2 *Simulation Design*

The performance of the estimators is investigated for different combinations of initials and weights. Four initial estimators, $\widehat{\boldsymbol{\theta}}_{\mathbf{0}}$, are considered: LMS (BP=0.5), LTS (trimming proportion 0.25, BP=0.25), LAD (BP=0.5) and LS (BP=0) estimators. Mallows weights based on Mahalanobis distance are used with $\alpha' = 1$ and $b = 9.23$ for a significance level of 0.01.$\boldsymbol{m_x}$ and $\boldsymbol{C_x}$are estimated using three multivariate location and scatter estimators: MVE (BP=0.5), MCD (with trimming proportion 0.25, BP=0.25), the classical sample mean and variance covariance matrix (BP=0). We considered only the following combinations of weights and initials: the classical sample mean and variance-covariance matrix with LS as an initial estimator, MVE with LMS and MCD with both LTS and LAD. The standard deviation σ is estimated using a robust scale estimate $\hat{\sigma}$ (BP=0.5) given by: $\hat{\sigma} = d\ med_i(med_j(r_i - r_j))$, d =1.1926 [17 and 27].

For the exponential power family, $c\ and\ \varphi$ are kept fixed at $c = 2, \varphi = 0.01$ and seven different values of λ are selected to reflect different levels of tail heaviness of the error's distribution ($\lambda = 0.55, 0.6, 0.65, 0.7, 0.8, 0.9, 1$). For the student's t family, φ was also fixed at 0.01 and seven different values of c are selected ($c = 4, 5, 6, 7, 10, 15, 20$). To investigate the effect of sample size on the behavior of the suggested estimators, we considered four sample sizes ($n = 25, 50, 100$ and 200)

After generating data, steps for each combination are performed in the following order: first $\widehat{\boldsymbol{\theta}}_{\mathbf{0}}$ is calculated, then the residuals r_i are obtained and used to compute $\hat{\sigma}$ and $\hat{\gamma}$. It has been found that, under the distribution assumption of this study, it is better to use the adjusted estimate version of the commonly used empirical kurtosis coefficient, $\hat{\gamma} = \frac{m_4}{m_2^2}\left(\frac{n-1}{n}\right)^2 - 3\frac{(n-1)^3}{n^2(n+1)}$, with m_2, m_4 the 2nd and 4th central sample moments, respectively. The value of $\hat{\lambda}$ or $\hat{c}$ is then obtained using their one to one correspondence with $\hat{\gamma}$. The weights w_i are computed using the X-data, then the components $\hat{\psi}$ and $\hat{\psi}'$are obtained to get finally the value of the proposed estimator. This process is repeated 10,000 times and the bias and mean squared errors are calculated.

4.3 *Simulation Results*

It has to be noted that the bias results of all estimators were calculated and results show that all estimators are asymptotically unbiased. Tables of the bias results together with some other tables concerning the efficiency of the estimators relative to the initials in the leverage points case are omitted to save the space. For the non leverage case, the proposed estimator $\widehat{\boldsymbol{\theta}}_{\boldsymbol{M}(\widehat{\boldsymbol{\theta}}_{\mathbf{0}})}$ is compared

to the LS estimator $\widehat{\boldsymbol{\theta}}_{(\boldsymbol{LS})}$ as well as to its initial estimator $\widehat{\boldsymbol{\theta}}_{\mathbf{0}}$. In the leverage points case, the Scheweppe-type estimator $\widehat{\boldsymbol{\theta}}_{\boldsymbol{VM}(\widehat{\boldsymbol{\theta}}_{\mathbf{0}})}$ is compared to $\widehat{\boldsymbol{\theta}}_{(\boldsymbol{LS})}$, to its initial estimator $\widehat{\boldsymbol{\theta}}_{\mathbf{0}}$, and to the Mallows-type estimator $\widehat{\boldsymbol{\theta}}_{\boldsymbol{WM}(\widehat{\boldsymbol{\theta}}_{\mathbf{0}})}$.

4.3.1. Case (a): Non-leverage points

In general, for the two distribution families [Table (a.1) and (a.2) in Appendix], the efficiency $\boldsymbol{RE_1}$ increases with n and tends to have the same value (100) as the distribution tends to the normal distribution ($\lambda \to 1/\, c \to \infty$). This means that all the estimators tend to be asymptotically equivalent under the normal distribution. The values of $\boldsymbol{RE_2}$ show that $\widehat{\boldsymbol{\theta}}_{\boldsymbol{M}(\widehat{\boldsymbol{\theta}}_{\mathbf{0}})}$outperforms its initial $\widehat{\boldsymbol{\theta}}_{\mathbf{0}}$for most of the cases and such performance increases with n and λ or c. The estimator $\widehat{\boldsymbol{\theta}}_{\boldsymbol{M}(\boldsymbol{LMS})}$ has the worst performance.The superiority of the other estimators to the LS estimator is evident.

Under the EP family, the best estimator is $\widehat{\boldsymbol{\theta}}_{\boldsymbol{M}(\boldsymbol{LAD})}$for all values of n [Figures (4.1)]. It is also almost as efficient as $\widehat{\boldsymbol{\theta}}_{\boldsymbol{M}(\boldsymbol{LTS})}$ under the t family [figure (4.2)], but with higher efficiency relative to LS even for $n = 25$.

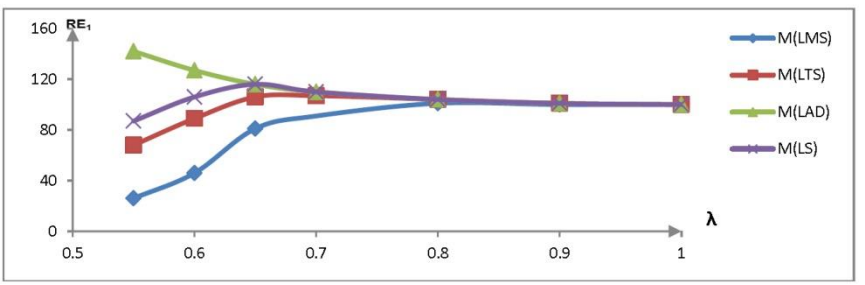

Figure 4.1
***RE₁* of $\widehat{\theta}_{\mathrm{M}(.)}$, EP-family, $n = 200$.**

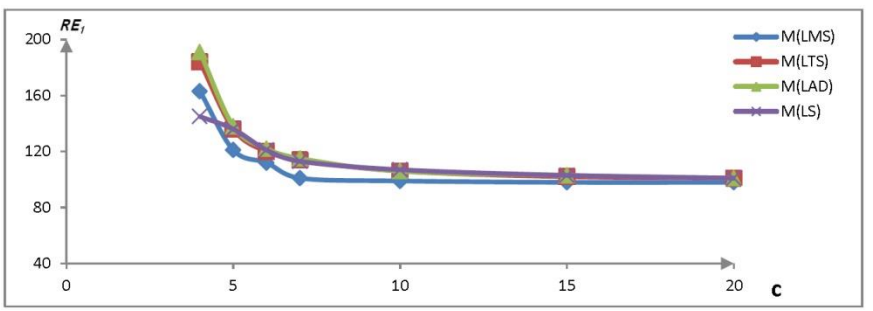

Figure 4.2
***RE₁* of $\widehat{\theta}_{\mathrm{M}(.)}$, t family , $n = 100$.**

4.3.2. Case (b): Leverage points

The results show that under the EP family and assuming leverages generated from heavy tailed distribution, $\widehat{\boldsymbol{\theta}}_{\mathbf{VM(LAD)}}$ is highly superior to LS and all other estimators as well as to its initials for λ approximately equals 0.65 or less [Tables (b-1) and (b-2)]. If leverages are due to gross errors, $\widehat{\boldsymbol{\theta}}_{\mathbf{VM(LTS)}}$ becomes slightly better than the competing estimators for λ up to 0.8 except for $\lambda = 0.55$. The estimators tend to be equivalent to each other and to LS for the normal distribution [Figures (4.3) and (4.4)].

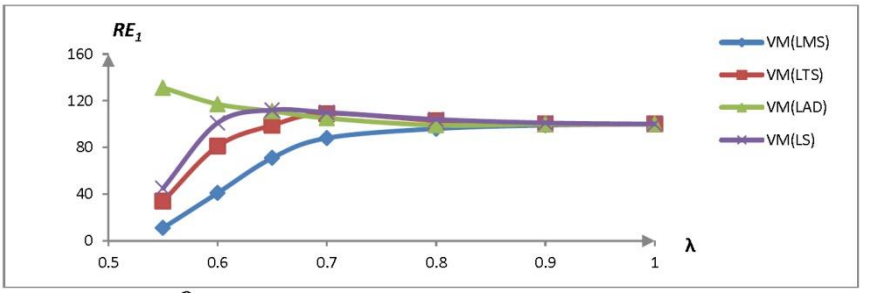

Figure 4.3

***RE*$_1$ of $\widehat{\theta}_{\mathbf{VM(.)}}$, EP family, heavy tailed, *n*=200.**

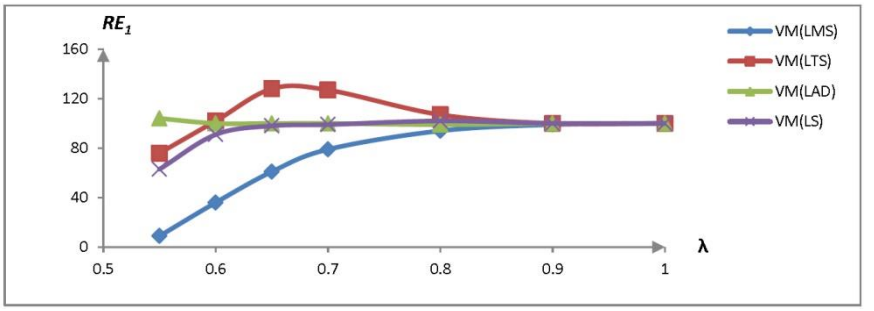

Figure 4.4

***RE*$_1$ of $\widehat{\theta}_{\mathbf{VM(.)}}$, EP family , gross errors, *n* = 200.**

As for the performance of the Scheweppe-type estimator $\widehat{\boldsymbol{\theta}}_{\mathbf{VM(.)}}$ relative to Mallows type $\widehat{\boldsymbol{\theta}}_{\mathbf{WM(.)}}$, results show better performance of Scheweppe-type compared to Mallows one. For example, assuming gross errors leverages, the relative efficiency of $\widehat{\boldsymbol{\theta}}_{\mathbf{VM(LAD)}}$ with respect to $\widehat{\boldsymbol{\theta}}_{\mathbf{WM(LAD)}}$ is (124,123,100) for $n = 200$ and $\lambda = 0.55$. Thus regardless of the leverage source, the Scheweppe-type estimator is recommended using LAD as initial estimator if λ is around 0.55 and otherwise with LTS as initial.

For the t-family, Tables (b-3) and (b-4) show that under heavy tailed explanatory variables, and for $n = 50$, $c \leq 7$, $\widehat{\boldsymbol{\theta}}_{\mathbf{VM(LAD)}}$ is the best [Figure (4.5)]. It is highly efficient than the LS and than its initial. As the sample size increases,

the performance of $\widehat{\boldsymbol{\theta}}_{\mathbf{VM(LTS)}}$ is the best**.** Under gross errors, for large n, $\widehat{\boldsymbol{\theta}}_{\mathbf{VM(LTS)}}$ is the best estimator for $c \leq 7$. It is highly efficient than LS and $\widehat{\boldsymbol{\theta}}_{\mathbf{WM(LTS)}}$. For example, for $c = 5$ and $n = 200$, the relative efficiency of $\widehat{\boldsymbol{\theta}}_{\mathbf{VM(LTS)}}$ with respect to $\widehat{\boldsymbol{\theta}}_{\mathbf{WM(LTS)}}$ is $(374{,}332{,}212)$.

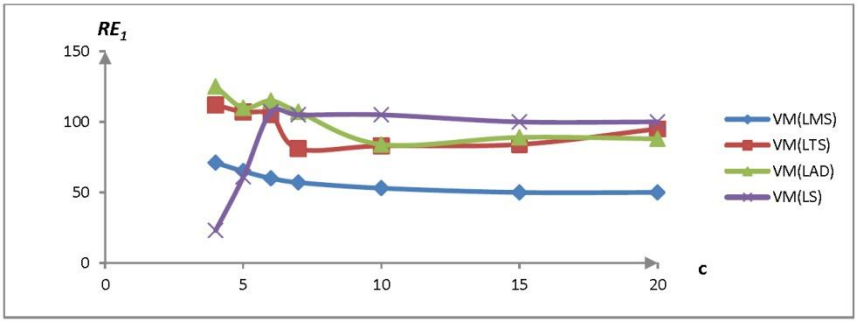

Figure 4.5
***RE*$_1$ of $\widehat{\boldsymbol{\theta}}_{\mathbf{VM(.)}}$, t-family, heavy tailed, $\boldsymbol{n = 50}$.**

This means that under the t-family of distributions, the performance of the suggested estimator is efficient for large n. Scheweppe-type estimator has better performance than Mallows-type where using LTS as initial is recommended whatever the source of leverage is.

Finally, in order to assess how estimating $\hat{\lambda}$ $(\hat{c})$ affects the behavior of the adaptive estimator, the performance of this last estimator is compared to that of the non adaptive estimator assuming λ (c) is known. Tables (4.1) and (4.2) display the bias and standard error (SE) of $\widehat{\boldsymbol{\theta}}_{\boldsymbol{M(LAD)}}$ and the corresponding non adaptive estimator denoted now $\widehat{\boldsymbol{\theta}}^{*}_{\boldsymbol{M(LAD)}}$ (assuming λ or c known).

The results in table (4.1) show the similarity of the performance of $\widehat{\boldsymbol{\theta}}_{\mathbf{M(LAD)}}$ to that of $\widehat{\boldsymbol{\theta}}^{*}_{\mathbf{M(LAD)}}$ at different λ (c). This reflects how $\hat{\lambda}$ $(\hat{c})$ is good and has good effect on the behavior of the adaptive estimator. The results in table (4.2) show also how good is the effect of $\hat{\lambda}$ $(\hat{c})$ on the behavior of $\widehat{\boldsymbol{\theta}}_{\mathbf{VM(LAD)}}$.

Table 4.1
Bias and standard error of $\widehat{\theta}_{M\,(LAD)}$, $\widehat{\theta}^{*}_{M\,(LAD)}$, n=50 , the EP-family / the t-family (non leverage case)

		the EP-family				the t-family			
		Bias		Standard error		Bias		Standard error	
λ	θ	$\widehat{\theta}_{M\,(LAD)}$	$\widehat{\theta}^{*}_{M\,(LAD)}$	$\widehat{\theta}_{M\,(LAD)}$	$\widehat{\theta}^{*}_{M\,(LAD)}$	$\widehat{\theta}_{M\,(LAD)}$	$\widehat{\theta}^{*}_{M\,(LAD)}$	$\widehat{\theta}_{M\,(LAD)}$	$\widehat{\theta}^{*}_{M\,(LAD)}$
0.55	$\widehat{\theta}_2$	0.01	0.02	2.201	2.202	0.01	0.01	1.939	1.939
	$\widehat{\theta}_3$	0.01	0.02	1.494	1.495	0.01	0.01	1.355	1.355
	$\widehat{\theta}_1$	0	0	1.621	1.621	0	0	1.456	1.456
0.6	$\widehat{\theta}_2$	0.02	0.029	2.024	2.031	0.02	0.02	2.096	2.096
	$\widehat{\theta}_3$	0.03	0.01	1.378	1.383	0.03	0.03	1.432	1.432
	$\widehat{\theta}_1$	0	0.01	1.484	1.49	0	0	1.587	1.587
0.65	$\widehat{\theta}_2$	0.01	0.01	1.899	1.9	0.01	0.0101	2.121	2.12
	$\widehat{\theta}_3$	0.03	0	1.273	1.274	0.03	0.03	1.445	1.444
	$\widehat{\theta}_1$	0.01	0	1.374	1.375	0.01	0.0099	1.592	1.591
0.7	$\widehat{\theta}_2$	0.03	0.04	1.76	1.771	0.03	0.0299	2.117	2.117
	$\widehat{\theta}_3$	0.03	0.011	1.181	1.19	0.03	0.0301	1.456	1.457
	$\widehat{\theta}_1$	0.02	0.01	1.322	1.331	0.02	0.0201	1.628	1.628
0.8	$\widehat{\theta}_2$	0.01	0.021	1.585	1.598	0.01	0.0099	2.121	2.12
	$\widehat{\theta}_3$	0.03	0.009	1.072	1.081	0.03	0.03	1.474	1.474
	$\widehat{\theta}_1$	0.01	0.02	1.177	1.185	0.01	0.0101	1.629	1.629
0.9	$\widehat{\theta}_2$	0.02	0.021	1.471	1.485	0.02	0.02	2.068	2.068
	$\widehat{\theta}_3$	0	0.005	0.989	1.001	0	0	1.413	1.414
	$\widehat{\theta}_1$	0	0.007	1.097	1.097	0	0	1.595	1.595
1	$\widehat{\theta}_2$	0.03	0.007	1.395	1.38	0.03	0.03	2.027	2.027
	$\widehat{\theta}_3$	0.02	0.003	0.941	0.927	0.02	0.02	1.386	1.386
	$\widehat{\theta}_1$	0	0.022	1.057	1.05	0	0	1.559	1.559

Table 4.2
Bias and standard error of $\widehat{\theta}_{VM\ (LAD)}$, $\widehat{\theta}^{*}_{VM\ (LAD)}$, $\widehat{\theta}_{(LAD)}$ and $\widehat{\theta}_{(LS)}$, n=50 , the EP-family / t-faimly (leverage case (heavy-tailed))

		the EP-family				The / t-faimly			
		Bias		Standard error		Bias		Standard error	
λ	θ	$\widehat{\theta}_{VM\ (LAD)}$	$\widehat{\theta}^{*}_{VM\ (LAD)}$	$\widehat{\theta}_{VM\ (LAD)}$	$\widehat{\theta}^{*}_{VM\ (LAD)}$	$\widehat{\theta}_{VM\ (LAD)}$	$\widehat{\theta}^{*}_{VM\ (LAD)}$	$\widehat{\theta}_{VM\ (LAD)}$	$\widehat{\theta}^{*}_{VM\ (LAD)}$
0.55	$\widehat{\theta}_2$	0	0	0.151	0.151	0	0	0.27	0.27
	$\widehat{\theta}_3$	0	0	0.104	0.104	0	0	0.178	0.178
	$\widehat{\theta}_1$	0.01	0.01	1.554	1.555	0.01	0.01	2.168	2.168
0.6	$\widehat{\theta}_2$	0	0	0.138	0.138	0.01	0.01	0.207	0.207
	$\widehat{\theta}_3$	0	0	0.097	0.098	0.01	0.01	0.141	0.141
	$\widehat{\theta}_1$	0.01	0.01	1.403	1.41	0.02	0.02	1.845	1.845
0.65	$\widehat{\theta}_2$	0	0.0001	0.128	0.129	0	0	0.192	0.192
	$\widehat{\theta}_3$	0	0	0.09	0.091	0	0	0.132	0.132
	$\widehat{\theta}_1$	0	0.0002	1.345	1.352	0.01	0.01	1.723	1.723
0.7	$\widehat{\theta}_2$	0	0	0.12	0.121	0	0	0.18	0.18
	$\widehat{\theta}_3$	0	0	0.082	0.082	0	0	0.125	0.125
	$\widehat{\theta}_1$	0.01	0.0094	1.272	1.281	0.01	0.01	1.687	1.687
0.8	$\widehat{\theta}_2$	0	0	0.109	0.11	0	0	0.153	0.153
	$\widehat{\theta}_3$	0	0	0.075	0.076	0.01	0.01	0.108	0.108
	$\widehat{\theta}_1$	0.01	0.0103	1.146	1.154	0	0	1.514	1.514
0.9	$\widehat{\theta}_2$	0	0	0.098	0.099	0	0	0.151	0.151
	$\widehat{\theta}_3$	0	0.0001	0.067	0.068	0	0	0.104	0.104
	$\widehat{\theta}_1$	0.01	0.0103	1.065	1.067	0.01	0.01	1.546	1.546
1	$\widehat{\theta}_2$	0	0	0.093	0.093	0	0	0.141	0.141
	$\widehat{\theta}_3$	0	0	0.064	0.064	0	0	0.096	0.096
	$\widehat{\theta}_1$	0.01	0.0095	1.012	1.012	0.02	0	1.513	1.513

5. Illustrative examples

Two real data sets that have been widely used in robust regression studies are considered. One is under the assumption of the EP family and the other under the t- family. The suggested estimators are calculated and Bootstrap bias as well as standard errors are recorded for comparison [12].

5.1 *Prostate cancer data*

The aim of analyzing this data is to investigate the factors affecting the incidence of prostate cancer for a sample of 97 observations. This data set was used by [14] to fit partially adaptive robust estimator for regression models under the assumption of EP family for the regression error. The variables are:

Log of prostate specific antigen (lpsa) (the dependent variable); Log of cancer volume (lcavol); Log of weight (lweight); Age; Log of benign prostatic hyperplasica amount (lbph); Seminal vesicle invasion (svi); Log of capsular penetration; Gleason score (gleason); Percent of Gleason scores 4 or 5.

According to [14], (lcavol), (lweight) and (lbph) are shown among the significant explanatory variables. Accordingly, we consider a linear model given by:

$$\mathrm{lpsa_i} = \theta_1 + \theta_2 \mathrm{lcavol_i} + \theta_3 \mathrm{lweight_i} + \theta_4 \mathrm{lbph_i} + \mathrm{E_i}.$$

Using MCD estimator, all weights w_i as given in (1.2) are found equal one indicating a non leverage setting. Using $\hat{\theta}_{(LAD)}$, say, as an initial estimate, the residuals are obtained, from which we find $\hat{\gamma} = 0.716$. $\hat{\lambda}$ is then calculated and found equal to 0.76. According to formula (4.1), $\hat{\theta}_{M(LAD)}$ is calculated and the estimated regression coefficients are as given in Table (5.1). All other estimates, $\hat{\theta}_{(.)}$ and $\hat{\theta}_{M(.)}$, are calculated in a similar way.

Table 5.1
Estimated regression coefficients for prostate cancer data

Estimates Coefficients	$\widehat{\boldsymbol{\theta}}_{\boldsymbol{M(LMS)}}$	$\widehat{\boldsymbol{\theta}}_{\boldsymbol{(LMS)}}$	$\widehat{\boldsymbol{\theta}}_{\boldsymbol{M(LTS)}}$	$\widehat{\boldsymbol{\theta}}_{\boldsymbol{(LTS)}}$	$\widehat{\boldsymbol{\theta}}_{\boldsymbol{M(LAD)}}$	$\widehat{\boldsymbol{\theta}}_{\boldsymbol{(LAD)}}$	$\widehat{\boldsymbol{\theta}}_{\boldsymbol{M(LS)}}$	$\widehat{\boldsymbol{\theta}}_{\boldsymbol{(LS)}}$
θ_2	**0.661**	**0.5**	**0.633**	**0.645**	**0.633**	**0.602**	**0.631**	**0.658**
Bias	0.028	0.098	0.012	0.046	0.006	0.013	0.008	0.033
(Variance)	(2.99)	(9.144)	(2.636)	(5.272)	(2.561)	(2.919)	(2.489)	(2.689)
θ_3	**0.679**	**0.074**	**0.621**	**0.259**	**0.585**	**0.599**	**0.591**	**0.581**
Bias	0.011	0.022	0.049	0.063	0.008	0.015	0.017	0.027
(Variance)	(4.414)	(13.809)	(3.942)	(8.278)	(3.905)	(4.413)	(3.795)	(4.061)
θ_4	**0.018**	**0.119**	**0.05**	**0.045**	**0.043**	**0.017**	**0.038**	**0.053**
Bias	0.019	0.028	0.006	0.01	0.009	0.044	0.016	0.022
(Variance)	(3.287)	(10.454)	(2.976)	(5.892)	(2.963)	(3.348)	(2.893)	(3.096)
θ_1	**-0.905**	**1.624**	**-0.65**	**0.815**	**-0.511**	**-0.511**	**-0.524**	**-0.523**
Bias	0.147	0.161	0.131	0.321	0.025	0.084	0.048	0.05
(Variance)	(5.5)	(17.875)	(5.043)	(10.489)	(5.116)	(5.935)	(4.95)	(5.346)

Clearly, the values are coherent with the behavior of the estimators for λ about 0.8 and sample size about 100. For example, concerning θ_2, the estimates under LTS or LAD are nearly the same, close to the LS estimates.

5.2 *Stack-loss data*

We consider here a data set given in [2]. It concerns 21 days (observations) of operation of a plant for the oxidation of ammonia to nitric acid. The nitric oxides produced are absorbed in a countercurrent absorption tower. [20] used this data set in fitting linear model under the assumption of t-distribution family. The variables are: Stack.Loss (ten times the percentage of the ingoing ammonia to the plant that escapes from the absorption column unabsorbed; that is, an (inverse) measure of the overall efficiency of the plant). This is the dependent variable. The other variables are: Air.Flow (the rate of operation of the plant); Water.Temp (the temperature of cooling water circulated through coils in the absorption tower) and Acid.Conc (the concentration of the acid circulating, minus 50, times 10: that is, 89 corresponds to 58.9 per cent acid).

Similar to prostate cancer example, we started with investigating explanatory variables for leverage points and the results show non leverage case. Then, the initial estimate $\hat{\theta}_{(\mathrm{LAD})}$is calculated and the corresponding residuals are obtained. The estimated kurtosis value $\hat{\gamma}$ is 15.88. Accordingly, the estimated value of the shape parameter is $\hat{c} = 5.377$. The robust scale estimate is 1.5168. Finally, $\hat{\theta}_{\mathrm{M(LAD)}}$ is calculated and given in Table (5.2) as well as $\hat{\theta}_{(\mathrm{LAD})}$ and $\hat{\theta}_{(\mathrm{LS})}$.

Table 5.2
Estimated regression coefficients for Stack-loss data

Estimates Coefficients	$\widehat{\boldsymbol{\theta}}_{\boldsymbol{M(LMS)}}$	$\widehat{\boldsymbol{\theta}}_{\boldsymbol{(LMS)}}$	$\widehat{\boldsymbol{\theta}}_{\boldsymbol{M(LTS)}}$	$\widehat{\boldsymbol{\theta}}_{\boldsymbol{(LTS)}}$	$\widehat{\boldsymbol{\theta}}_{\boldsymbol{M(LAD)}}$	$\widehat{\boldsymbol{\theta}}_{\boldsymbol{(LAD)}}$	$\widehat{\boldsymbol{\theta}}_{\boldsymbol{M(LS)}}$	$\widehat{\boldsymbol{\theta}}_{\boldsymbol{(LS)}}$
θ_2	**1.37**	**0.75**	**0.92**	**0.82**	**0.85**	**0.83**	**0.87**	**0.72**
Bias	0.21	0.25	0.002	0.09	0.041	0.045	0.011	0.012
(Variance)	(8.9)	(3.72)	(3.86)	(4.05)	(3.72)	(3.85)	(3.14)	(3.84)
θ_3	**0.44**	**0.5**	**0.59**	**0.46**	**0.59**	**0.57**	**0.75**	**1.29**
Bias	0.39	0.32	0.011	0.04	0.021	0.327	0.072	0.079
(Variance)	(17.09)	(3.4)	(3.92)	(4.29)	(3.64)	(3.72)	(3.12)	(3.75)
θ_4	**-0.33**	**0.02**	**-0.12**	**-0.08**	**-0.09**	**-0.06**	**-0.12**	**-0.15**
Bias	0.22	0.14	0.011	0.03	0.026	0.063	0.087	0.01
(Variance)	(12.24)	(3.98)	(3.64)	(4.03)	(3.52)	(3.57)	(3.62)	(3.61)
θ_1	**-43.77**	**-39.25**	**-39.44**	**-35.33**	**-38.37**	**-39.69**	**-40.49**	**-39.92**
Bias	2.35	3.6	0.012	0.171	0.009	0.131	0.013	0.015
(Variance)	(56.2)	(19.22)	(5.4)	(17.2)	(3.17)	(3.96)	(4.26)	(3.64)

The values in the table are consistent with the behavior of the regression coefficients estimators for c between 5 and 6 and n around 25. For example, the

estimates of θ_2 using LAD or LTS are nearly the same, far from the estimates using LMS. Same results are evident for θ_3.

6. Conclusions

For the linear regression model and under two location-scale large families of distributions, namely the exponential power and t-families, a sort of robust one-step M- and GM-estimators are suggested. The suggested estimators have adapted likelihood and their efficiency relative to LS estimator as well as to their initial estimators is investigated. This is easily made due to the existing one-one correspondence between the parameter and the kurtosis measure proved to be a decreasing function in the shape parameter. Relying on a Monte Carlo simulation study, it is shown that the adaptive one-step M- estimator is highly more efficient than LS as well as its initial. The estimator using LAD estimate as initial is shown to outperform all the competing estimators and that with LMS initial is the worst.

In assessing the estimators' sensitivity against two different sources of leverage points, it is found that the Scheweppe type one-step GM-estimator is more efficient than Mallows type. Using LAD as initial and MCD in calculating weights is recommended under the EP family for λ around 0.55, and with LTS and same weight for moderate values of λ. Under the t-family, using LTS as initial estimator and MCD in calculating weights has the best performance.

Appendix

Table (a.1)

RE_1, RE_2 of different estimators $\hat{\theta}_{M(.)}$ relative to $\hat{\theta}_{(LS)}$ and the initial one $\hat{\theta}_{(.)}$, respectively, for the EP-family (non leverage case).

		$\hat{\theta}_{M\,(LMS)}$								$\hat{\theta}_{M\,(LTS)}$								$\hat{\theta}_{M\,(LAD)}$								$\hat{\theta}_{M\,(LS)}$			
		RE_1				RE_2				RE_1				RE_2				RE_1				RE_2				RE_1			
	n	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200
0.55	$\hat{\theta}_2$	17	18	22	26	44	46	55	70	42	43	54	68	62	65	73	90	110	115	131	142	103	104	104	105	45	54	69	87
	$\hat{\theta}_3$	18	18	23	26	43	47	57	72	42	45	56	69	62	67	74	87	108	116	129	141	103	104	105	105	45	52	72	88
	$\hat{\theta}_1$	21	24	27	30	47	54	69	86	45	53	65	79	62	71	83	96	112	124	139	144	103	104	104	105	50	66	80	101
0.6	$\hat{\theta}_2$	31	32	39	46	86	94	112	142	63	67	76	89	105	112	121	136	101	105	116	127	104	105	106	107	70	78	91	106
	$\hat{\theta}_3$	29	32	41	48	83	95	114	152	64	65	77	87	107	115	119	130	102	106	116	126	104	105	107	107	71	80	95	110
	$\hat{\theta}_1$	35	41	45	52	89	105	135	180	67	75	87	97	100	115	127	133	105	114	123	129	104	105	106	108	77	92	106	112
0.65	$\hat{\theta}_2$	59	60	71	81	179	197	240	300	89	91	100	106	162	180	181	183	92	100	109	116	107	108	110	111	97	103	110	116
	$\hat{\theta}_3$	60	63	72	80	179	211	235	299	87	90	101	107	156	172	176	184	91	100	110	116	107	109	109	111	97	104	111	117
	$\hat{\theta}_1$	64	72	82	89	177	225	263	345	91	97	105	111	154	165	175	185	99	108	118	120	107	110	110	111	99	111	115	120
0.7	$\hat{\theta}_2$	74	76	83	91	245	263	304	383	94	96	102	107	187	203	209	212	92	97	105	110	110	112	114	117	101	103	108	110
	$\hat{\theta}_3$	76	76	84	92	242	269	313	378	92	95	102	108	181	205	206	209	93	97	107	109	109	111	114	117	101	103	109	113
	$\hat{\theta}_1$	80	86	92	97	238	275	345	444	97	101	106	109	180	194	199	206	94	101	107	113	110	114	116	117	103	110	111	113

Contd...

0.8	$\hat{\theta}_2$	94	95	99	101	348	407	450	513	100	100	102	104	222	250	251	252	94	98	100	104	125	128	131	131	102	104	104	104
	$\hat{\theta}_3$	94	97	99	102	343	405	450	512	100	100	102	105	221	244	249	254	93	97	102	104	127	128	130	129	102	103	105	104
	$\hat{\theta}_1$	96	99	101	102	346	388	467	575	101	102	104	105	216	234	238	241	97	101	104	105	126	131	129	129	103	105	104	105
0.9	$\hat{\theta}_2$	98	98	99	100	392	485	521	618	100	100	100	101	251	282	283	298	97	98	100	101	140	141	142	145	101	101	101	101
	$\hat{\theta}_3$	98	99	100	100	402	474	542	607	100	100	101	101	252	296	296	297	96	99	100	101	142	142	142	143	100	101	101	101
	$\hat{\theta}_1$	99	100	100	101	403	469	547	689	100	101	101	101	249	276	281	288	99	100	101	101	141	142	145	146	101	101	101	101
1	$\hat{\theta}_2$	100	100	100	100	431	549	613	728	100	100	100	100	267	330	334	348	100	100	100	100	154	154	155	156	100	100	100	100
	$\hat{\theta}_3$	100	100	100	100	428	519	612	704	100	100	100	100	267	328	341	353	100	100	100	100	154	155	155	159	100	100	100	100
	$\hat{\theta}_1$	100	100	100	100	441	533	657	774	100	100	100	100	268	321	327	357	100	100	100	100	154	156	158	159	100	100	100	100

Remark: Relative efficiency of estimator (a) with respect to estimator (b) is given by: $\text{RE} = \frac{MSE(b)}{MSE(a)} \times 100$

$\boldsymbol{RE_1}$: The relative efficiency (%) with respect to $\hat{\theta}_{(\text{LS})}$,

$\boldsymbol{RE_2}$: The relative efficiency (%) with respect to $\hat{\theta}_{(\cdot)}$.

Table (a.2)

RE_1, RE_2 of different estimators $\hat{\theta}_{M(.)}$ relative to $\hat{\theta}_{(LS)}$ and the initial one $\hat{\theta}_{(.)}$, respectively, for the t-family (non leverage case).

		$\hat{\theta}_{M\,(LMS)}$								$\hat{\theta}_{M\,(LTS)}$								$\hat{\theta}_{M\,(LAD)}$								$\hat{\theta}_{M\,(LS)}$			
		RE_1				RE_2				RE_1				RE_2				RE_1				RE_2				RE_1			
	n	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200
4	$\hat{\theta}_2$	2	95	163	182	3	172	315	394	159	166	184	189	162	178	182	183	154	173	191	205	103	118	118	121	113	120	145	178
	$\hat{\theta}_3$	1	90	167	188	1	159	314	399	160	169	185	186	161	175	180	185	154	162	184	186	108	116	120	121	109	126	154	165
	$\hat{\theta}_1$	2	155	191	193	3	261	365	428	164	183	188	198	156	171	172	172	164	175	193	196	110	118	118	121	128	137	147	180
5	$\hat{\theta}_2$	1	67	121	139	3	187	357	460	126	132	136	137	181	202	203	205	128	131	138	141	123	124	125	127	115	116	136	138
	$\hat{\theta}_3$	1	87	128	135	2	240	367	440	123	127	137	139	186	200	208	210	128	129	135	138	123	125	126	127	110	121	136	140
	$\hat{\theta}_1$	2	113	137	138	5	281	401	480	128	134	140	142	177	193	195	196	132	135	138	141	124	125	126	126	118	133	137	139
6	$\hat{\theta}_2$	2	66	112	119	5	218	398	457	114	114	120	124	199	217	219	230	115	122	122	123	127	128	130	130	111	118	121	123
	$\hat{\theta}_3$	1	79	112	122	2	262	397	488	112	116	122	123	196	219	225	228	116	118	121	124	128	130	131	131	110	117	123	124
	$\hat{\theta}_1$	3	102	121	123	7	317	430	518	118	120	125	125	191	211	213	214	116	122	122	124	128	129	131	132	110	121	123	124

Contd...

7	$\hat{\theta}_2$	3	52	101	113	11	195	396	489	106	106	114	116	206	232	235	241	110	112	115	116	129	131	131	132	110	111	113	114	
	$\hat{\theta}_3$	1	69	108	112	3	261	419	501	106	108	115	116	203	234	237	242	108	111	112	113	130	131	132	132	104	112	114	115	
	$\hat{\theta}_1$	3	90	111	115	8	312	460	547	108	113	115	116	199	224	229	229	113	112	116	116	130	131	132	132	108	114	115	116	
10	$\hat{\theta}_2$	23	67	99	105	82	292	454	552	101	102	106	107	227	259	262	265	103	106	106	108	137	138	139	139	104	105	107	107	
	$\hat{\theta}_3$	12	73	100	105	43	312	451	552	101	103	105	106	222	259	264	264	104	105	106	107	135	136	138	138	105	104	106	106	
	$\hat{\theta}_1$	28	96	104	105	98	376	497	577	103	106	107	107	215	241	252	254	105	107	107	107	135	136	138	144	105	105	107	107	
15	$\hat{\theta}_2$	28	76	98	101	69	356	512	593	98	99	102	103	238	276	285	287	101	102	103	103	142	143	143	144	102	102	103	102	
	$\hat{\theta}_3$	17	84	98	101	28	388	489	590	99	99	102	103	236	281	283	286	101	102	102	103	143	143	147	147	102	102	103	103	
	$\hat{\theta}_1$	30	96	102	102	74	442	552	656	101	101	103	103	234	267	270	285	101	102	103	103	142	141	143	145	102	102	103	103	
20	$\hat{\theta}_2$	30	80	98	101	79	389	537	625	98	98	101	101	243	289	296	305	100	101	101	101	145	146	147	148	101	101	101	101	
	$\hat{\theta}_3$	19	87	99	101	35	406	523	621	99	99	101	101	237	293	301	305	101	101	102	102	145	146	145	148	101	101	102	102	
	$\hat{\theta}_1$	32	96	101	101	76	453	572	684	100	100	101	102	242	271	279	286	101	101	102	102	144	145	147	149	100	100	101	101	

Table (b.1)

RE_1, RE_3 of different estimators $\hat{\theta}_{VM(.)}$ relative to $\hat{\theta}_{(LS)}$ and the Mallows one $\hat{\theta}_{WM(.)}$, respectively, EP-family (leverage case(heavy-tailed)).

		$\hat{\theta}_{VM\ (LMS)}$								$\hat{\theta}_{VM\ (LTS)}$								$\hat{\theta}_{VM(LAD)}$								$\hat{\theta}_{VM\ (LS)}$			
		RE_1				RE_3				RE_1				RE_3				RE_1				RE_3				RE_1			
	n	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200
0.55	$\hat{\theta}_2$	3	8	7	11	98	99	99	100	16	23	25	34	16	23	25	34	94	112	114	131	100	100	100	101	24	28	31	45
	$\hat{\theta}_3$	6	8	9	10	100	99	100	100	22	24	26	32	22	24	26	32	101	111	116	120	100	100	100	101	22	29	33	41
	$\hat{\theta}_1$	10	11	13	14	100	100	100	100	27	31	39	51	27	31	39	51	114	125	138	141	100	100	100	100	30	36	52	60
0.6	$\hat{\theta}_2$	10	28	30	41	95	99	100	100	51	62	68	81	51	62	68	81	86	101	104	117	100	100	100	100	68	79	80	101
	$\hat{\theta}_3$	22	30	33	37	99	99	100	100	61	65	69	79	61	65	69	79	96	101	106	115	100	100	100	101	71	80	87	96
	$\hat{\theta}_1$	36	43	47	53	99	100	100	100	71	78	88	99	71	78	88	99	107	117	123	129	100	100	100	100	81	94	106	116
0.65	$\hat{\theta}_2$	22	49	55	71	95	99	101	100	70	82	88	99	70	82	88	99	85	97	98	111	100	100	100	100	88	100	103	112
	$\hat{\theta}_3$	43	54	58	63	99	100	101	102	84	86	89	95	84	86	89	95	92	95	103	106	100	100	100	101	95	99	105	104
	$\hat{\theta}_1$	62	69	79	85	99	100	100	100	94	96	107	112	94	96	107	112	103	106	117	119	100	100	100	100	101	108	114	117

Contd...

0.7	$\hat{\theta}_2$	34	65	73	88	94	100	104	101	81	91	96	109	81	91	96	85	94	94	105	100	101	95	103	105	110
	$\hat{\theta}_3$	61	71	74	81	99	101	102	103	91	95	97	105	91	95	97	89	94	99	101	99	101	102	103	103	107
	$\hat{\theta}_1$	81	86	92	99	99	100	100	100	98	101	107	110	98	101	107	97	102	108	112	100	100	107	110	111	113
0.8	$\hat{\theta}_2$	59	86	92	96	101	101	103	112	92	97	99	103	92	97	99	84	91	93	99	100	101	99	103	102	104
	$\hat{\theta}_3$	84	89	90	93	99	104	106	107	96	98	98	100	96	98	98	91	94	96	98	99	101	102	102	103	102
	$\hat{\theta}_1$	93	98	99	101	99	101	101	100	100	102	103	104	100	102	103	96	99	103	104	100	101	100	104	105	106
0.9	$\hat{\theta}_2$	86	96	98	99	103	105	111	120	98	100	99	100	98	100	99	91	99	97	100	104	102	105	114	101	101
	$\hat{\theta}_3$	97	97	98	99	100	107	110	111	100	99	99	100	100	99	99	96	98	98	98	101	103	109	109	101	101
	$\hat{\theta}_1$	99	100	100	100	100	101	101	101	101	100	101	101	101	100	101	99	100	101	100	101	101	101	101	101	101
1	$\hat{\theta}_2$	100	100	100	100	103	107	120	125	100	100	100	100	100	100	100	100	100	100	100	104	106	126	131	100	100
	$\hat{\theta}_3$	100	100	100	100	100	108	112	115	100	100	100	100	100	100	100	100	100	100	100	103	108	115	113	100	100
	$\hat{\theta}_1$	100	100	100	100	100	102	102	101	100	100	100	100	100	100	100	100	100	100	100	102	102	102	101	100	100

RE_3 : the relative efficiency of $\hat{\theta}_{\mathrm{VM}(\theta_0)}$ with respect to $\hat{\theta}_{\mathrm{WM}(\theta_0)}$.

Table (b.2)

RE_1, RE_3 of different estimators $\hat{\theta}_{VM(.)}$ relative to $\hat{\theta}_{(LS)}$ and the Mallows one $\hat{\theta}_{WM(.)}$, respectively, for the EP-family (leverage case(gross errors)).

		$\hat{\theta}_{VM(LMS)}$								$\hat{\theta}_{VM(LTS)}$								$\hat{\theta}_{VM(LAD)}$								$\hat{\theta}_{VM(LS)}$			
		RE_1				RE_3				RE_1				RE_3				RE_1				RE_3				RE_1			
	n	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200
0.55	$\hat{\theta}_2$	1	3	6	9	83	93	103	108	11	15	56	76	80	101	135	158	64	76	101	104	99	104	113	124	21	24	44	63
	$\hat{\theta}_3$	2	3	4	6	96	102	109	114	9	9	51	71	87	93	133	165	64	78	100	105	99	103	113	123	21	23	47	55
	$\hat{\theta}_1$	12	12	13	14	101	102	102	102	25	27	89	94	96	100	101	107	129	134	135	152	100	100	103	100	36	50	58	77
0.6	$\hat{\theta}_2$	2	7	24	36	69	78	113	117	39	49	102	102	61	97	194	233	63	81	99	100	99	104	111	113	61	71	84	91
	$\hat{\theta}_3$	6	9	18	23	90	101	128	136	33	33	105	107	77	80	188	257	63	83	99	100	99	103	112	115	62	69	86	87
	$\hat{\theta}_1$	43	43	46	53	101	103	104	104	67	67	101	109	91	99	101	110	117	121	123	135	100	100	101	100	94	106	113	124
0.65	$\hat{\theta}_2$	3	11	38	61	60	70	107	125	56	73	125	128	68	89	274	293	63	80	91	100	99	104	112	115	86	96	96	98
	$\hat{\theta}_3$	13	17	34	43	87	101	142	160	55	57	127	121	67	72	262	336	63	81	90	99	98	103	113	119	88	94	95	96
	$\hat{\theta}_1$	68	72	75	92	101	105	106	106	88	90	110	117	90	98	100	109	108	115	117	122	100	100	102	100	109	114	116	123

Contd...

0.7	$\hat{\theta}_1$	82	84	101	88	108	101	87	109	108	94	120	88	97	109	348	63	108	87	114	99	100	112	114	99
	$\hat{\theta}_3$	19	26	61	52	102	87	57	126	129	65	126	35	83	302	348	63	81	87	99	98	108	123	130	96	98	98	100	100
	$\hat{\theta}_2$	4	18	79	63	74	57	126	129	65	86	127	35	83	302	348	63	81	87	100	99	108	122	124	95	99	99	99	99
0.8	$\hat{\theta}_1$	95	93	109	95	114	105	91	114	114	97	108	91	99	100	110	101	101	103	104	100	100	100	104	104	104	105	105	102
	$\hat{\theta}_3$	34	47	84	73	127	99	63	180	219	83	103	52	59	411	464	64	82	83	99	103	113	173	178	100	100	100	101	100
	$\hat{\theta}_2$	12	39	94	75	89	63	124	132	78	91	107	33	74	342	484	65	82	84	99	101	120	163	170	99	100	100	102	102
	$\hat{\theta}_1$	99	100	100	100	139	158	135	148	121	98	99	100	370	435	823	99	99	100	100	380	421	316	433	100	100	100	100	100
0.9	$\hat{\theta}_3$	99	100	100	100	236	218	227	251	100	100	100	353	438	529	760	99	99	100	100	361	402	433	502	100	100	100	100	100
	$\hat{\theta}_2$	37	73	99	94	115	91	141	139	100	99	100	68	162	355	641	74	89	94	100	135	191	271	339	100	100	100	100	100
	$\hat{\theta}_3$	68	78	96	92	182	153	213	238	94	96	99	116	119	418	603	75	90	93	100	160	165	312	335	100	100	100	100	100
	$\hat{\theta}_1$	98	98	101	98	120	107	118	121	98	98	99	100	100	101	112	99	100	101	101	100	100	101	109	101	101	101	101	101
1	$\hat{\theta}_2$	99	100	100	100	139	158	135	148	100	100	100	370	381	435	823	99	99	100	100	380	421	316	433	100	100	100	100	100
	$\hat{\theta}_3$	99	100	100	100	236	218	227	251	100	100	100	353	438	529	760	99	99	100	100	361	402	433	502	100	100	100	100	100
	$\hat{\theta}_1$	100	100	100	100	122	109	123	124	100	100	100	100	101	105	115	100	100	100	100	100	101	105	115	100	100	100	100	100

Table (b.3)

RE_1, RE_3 of different estimators $\hat{\theta}_{VM(.)}$ relative to $\hat{\theta}_{(LS)}$ and the Mallows one $\hat{\theta}_{WM(.)}$, respectively, t-family (leverage case (heavy-tailed)).

		$\hat{\theta}_{VM\ (LMS)}$								$\hat{\theta}_{VM\ (LTS)}$								$\hat{\theta}_{VM(LAD)}$								$\hat{\theta}_{VM\ (LS)}$			
		RE_1				RE_3				RE_1				RE_3				RE_1				RE_3				RE_1			
	n	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200
4	$\hat{\theta}_2$	0	71	108	165	60	66	90	95	1	112	159	183	1	74	96	98	0	125	160	175	15	87	94	98	11	23	131	168
	$\hat{\theta}_3$	0	109	111	141	66	73	80	88	7	144	147	159	4	86	93	94	1	149	159	172	18	92	93	96	15	18	142	160
	$\hat{\theta}_1$	0	169	173	192	94	96	97	99	53	178	190	195	34	98	99	100	0	181	190	198	22	99	99	100	44	55	172	184
5	$\hat{\theta}_2$	0	65	88	121	74	86	93	95	3	107	122	133	3	82	97	99	2	110	112	131	22	90	98	99	12	61	116	134
	$\hat{\theta}_3$	0	77	82	114	80	85	90	90	22	108	126	129	28	90	95	96	11	125	129	122	28	95	95	98	18	40	120	125
	$\hat{\theta}_1$	0	122	131	141	82	97	98	99	50	135	137	141	38	99	99	100	46	134	136	140	35	99	99	100	63	88	130	141
6	$\hat{\theta}_2$	0	60	69	115	80	88	91	96	14	105	109	120	20	86	97	100	7	115	114	117	28	92	99	100	19	107	114	119
	$\hat{\theta}_3$	0	62	78	109	88	88	92	98	98	106	111	107	91	94	97	98	30	108	110	109	36	96	98	99	51	106	115	116
	$\hat{\theta}_1$	2	109	115	124	84	97	98	100	90	121	121	124	77	99	99	100	91	119	123	123	76	100	100	100	108	116	123	124

Contd...

7	$\hat{\theta}_2$	0	57	64	102	81	89	90	97	56	81	105	109	90	89	99	100	24	107	110	111	38	94	100	100	61	105	109	114
	$\hat{\theta}_3$	0	63	75	106	89	88	89	96	96	98	98	103	96	96	98	99	75	105	107	109	70	98	98	99	111	110	110	110
	$\hat{\theta}_1$	2	102	107	114	82	97	98	100	109	111	113	113	99	99	100	100	108	113	114	115	95	100	100	102	110	114	114	116
10	$\hat{\theta}_2$	0	53	55	95	86	87	93	97	48	83	99	103	91	97	97	102	72	84	84	102	85	98	98	102	101	105	102	106
	$\hat{\theta}_3$	0	54	69	82	93	86	88	97	82	93	94	100	90	97	99	102	74	94	96	99	99	100	102	104	105	104	104	102
	$\hat{\theta}_1$	1	94	100	105	87	96	98	99	100	104	106	106	101	100	100	100	104	105	105	106	100	100	100	100	107	106	106	107
15	$\hat{\theta}_2$	0	50	55	91	88	90	98	98	47	84	90	100	91	98	102	103	85	89	85	100	98	102	104	103	101	100	102	102
	$\hat{\theta}_3$	0	52	68	81	95	88	89	99	84	90	93	95	95	96	103	104	100	94	98	95	101	104	104	106	102	100	101	101
	$\hat{\theta}_1$	4	91	95	101	88	95	97	100	97	100	102	103	101	100	101	101	101	102	102	102	101	101	101	101	102	103	104	103
20	$\hat{\theta}_2$	0	50	64	92	90	92	97	97	36	95	85	99	92	99	105	104	85	88	100	100	103	104	106	108	101	100	102	101
	$\hat{\theta}_3$	0	51	66	81	99	89	90	99	79	89	92	94	88	96	104	105	95	96	99	99	101	105	106	107	101	100	101	100
	$\hat{\theta}_1$	0	92	95	100	88	95	97	100	92	99	101	101	100	99	101	101	100	100	101	102	101	101	101	101	101	101	101	102

Table (b.4)

RE_1, RE_3 of different estimators $_{VM(.)}$relative to $_{(LS)}$and the Mallows one$_{WM(.)}$, respectively, t-family (leverage (gross errors)).

		$\hat{\theta}_{VM\ (LMS)}$								$\hat{\theta}_{VM\ (LTS)}$								$\hat{\theta}_{VM(LAD)}$								$\hat{\theta}_{VM\ (LS)}$			
		RE_1				RE_3				RE_1				RE_3				RE_1				RE_3				RE_1			
	n	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200
4	$\hat{\theta}_2$	0	6	101	115	11	62	121	145	5	14	120	149	127	225	120	149	33	43	50	74	71	788	874	106	6	30	44	74
	$\hat{\theta}_3$	0	8	108	113	11	33	106	129	6	14	122	142	113	192	102	105	35	40	43	60	57	469	948	117	6	27	34	64
	$\hat{\theta}_1$	15	63	113	118	12	28	104	127	136	167	190	198	104	155	121	323	182	195	200	207	99	102	155	183	7	17	80	198
5	$\hat{\theta}_2$	0	6	93	107	13	47	199	250	3	10	114	138	179	303	158	212	20	22	38	56	55	624	622	124	11	15	49	102
	$\hat{\theta}_3$	0	7	96	106	14	38	139	178	4	10	116	132	128	248	138	332	23	27	32	43	42	371	712	128	12	17	25	102
	$\hat{\theta}_1$	8	56	112	116	19	25	154	218	101	122	142	143	117	168	130	374	124	137	140	142	99	114	162	195	104	134	134	137
6	$\hat{\theta}_2$	0	5	86	94	21	89	241	380	3	9	112	132	294	306	179	353	15	18	34	48	48	503	582	262	13	57	102	108
	$\hat{\theta}_3$	0	5	87	93	22	45	219	288	3	9	114	127	177	325	163	427	19	23	28	36	35	275	540	203	15	69	109	113
	$\hat{\theta}_1$	7	21	104	110	31	84	183	223	90	109	120	123	137	195	159	479	112	119	123	125	98	196	181	204	121	123	124	126

Contd...

7	$\hat{\theta}_2$	0	4	62	73	38	78	316	433	2	6	113	130	326	338	192	404	17	26	31	45	45	473	422	246	14	86	104	110	
	$\hat{\theta}_3$	0	3	73	78	38	76	245	415	3	7	113	125	214	346	479	544	18	26	31	34	33	254	493	224	16	89	110	109	
	$\hat{\theta}_1$	5	17	92	94	39	63	227	398	72	100	115	113	169	221	222	566	103	114	116	116	98	190	201	224	115	116	116	118	
10	$\hat{\theta}_2$	0	7	57	67	30	78	564	648	1	7	91	118	416	506	484	517	15	24	29	38	38	346	360	373	77	91	101	104	
	$\hat{\theta}_3$	0	5	69	71	41	34	309	488	1	6	92	113	216	420	530	604	16	24	27	29	27	208	352	355	81	92	99	102	
	$\hat{\theta}_1$	2	12	86	92	41	87	321	405	70	93	106	106	198	228	322	568	94	103	107	107	97	202	210	320	105	106	107	107	
15	$\hat{\theta}_2$	0	5	55	61	40	48	587	694	2	5	80	104	447	606	592	622	14	23	28	32	32	266	270	403	88	97	99	100	
	$\hat{\theta}_3$	0	4	63	69	44	71	559	660	3	3	81	1011	426	589	658	840	15	23	24	28	23	190	292	425	90	95	99	100	
	$\hat{\theta}_1$	1	13	82	91	51	100	327	499	70	88	101	102	231	333	434	650	91	99	101	103	97	221	224	462	101	102	102	103	
20	$\hat{\theta}_2$	0	4	54	60	63	138	633	756	2	3	71	100	541	938	723	875	13	23	28	31	31	251	256	469	89	98	100	100	
	$\hat{\theta}_3$	0	3	58	61	50	150	668	782	2	2	68	100	493	778	765	976	14	22	23	28	22	169	180	455	89	95	100	100	
	$\hat{\theta}_1$	0	9	71	85	43	171	420	586	69	70	94	100	461	593	691	942	89	98	101	101	97	250	267	476	98	100	100	100	

References

[1] Albert, J., Delampady, M., & Polasek, W., A Class of Distributions for Robustness Studies. *Journal of Statistical Planning and Inference*, 28, 291-304 (1991).

[2] Andrews, D. F., A Robust Method for Multiple Linear regression. *Technometrics*, 16, 523-531 (1974).

[3] Bedrick, E.J., Lapidus, J. & Powell, J.F., Estimating the Mahalanobis Distance from Mixed Continuous and Discrete Data. *Biometrics*, 56,394-401 (2000).

[4] Bickel, P. J., One-Step Huber Estimation in the Linear Model. *Journal of the American Statistical Association*, 70, 428-434 (1975).

[5] Bickel, P. J., The 1980 Wald Memorial Lectures on Adaptive Estimation. *Annals of Statistics*, 3, 647-671 (1982).

[6] Billor, N., Hadi, A. S., and Velleman, P. F., BACON: Blocked Adaptive Computationally Efficient Outlier Nominators . *Computational Statistics & Data Analysis*, 34, 279-298 (2000).

[7] Coakley, C. W., & Hettmansperger, T. P., A Bounded Influence, High Breakdown, Efficient Regression Estimator. *Journal of the American Statistical Association*, 88, 872-880 (1993).

[8] De Jongh, P. J., De Wet, T., & Welsh, A. H., Mallows-Type Bounded Influence Regression Trimmed Means. *Journal of The American Statistical Association*, 83, 805-810 (1988).

[9] De Menezes, D.F, Prata, D.M., Secchi, A.R. & Pinto, J. C., A review on robust M-estimators for regression analysis. *Computers and Chemical Engineering*,147, 107254 (2021).

[10] Desgagné, A., Efficient and robust estimation of regression and scale parameters, with outlier detection. *Computational Statistics and Data Analysis*, 155, 107-114 (2021).

[11] El Gayar, S. M., On the Geometric and Analytical Characteristics of Complete Even Power Exponential Distribution. *Bulletin of The Faculty of Science*, 25, (2-c), 1-5, Assiut University, Egypt (1996).

[12] Efron, B. & Tibshirani, R., Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy. *Statistical Science*, 1: 54-57 (1986).

[13] Gupta, A., Singh, V. Mathur, P. and Travieso-Gonzalez, C. M., Prediction of COVID-19 pandemic measuring criteria using support

vector machine, prophet and linear regression models in Indian scenario, *Journal of Interdisciplinary Mathematics*, 24:1, 89-108 (2021), DOI: 10.1080/09720502.2020.1833458

[14] Hansen, J. V., McDonald, J. B., & Turley, R. S., Partially Adaptive Robust Estimation of Regression Models and Applications. *European Journal of Operational Research*, 170, 132-143 (2006).

[15] Hogg, R. V. Adaptive Robust Procedures: A Partial Review and Some Suggestions for Future Applications and Theory. *Journal of the American Statistical Association*, 69, 909-923 (1974).

[16] Huber, P. J., & Ronchetti, E. M., *Robust Statistics.* 2nd edition, John Wiley & Sons, Inc., Canada (2009).

[17] Jurecková, J. , Picek, J. & Schindler, M., *Robust Statistical Methods with R*. Taylor & Francis group, LLC, New York (2019).

[18] Kafadar, K., John Tukey and Robustness. *Institute of Mathematical Statistics*, 18, 319-331 (2003).

[19] Kong , D., Bondell, H.D. & Wu, Y., Fully efficient robust estimation, outlier detection and variable selection via penalized regression. *Statistica Sinica*, 28, 1031-1052 (2018).

[20] Lange, K.L., Little, J. A., & Taylor, J. M., Robust Statistical Modeling Using the t Distribution. *Journal of the American Statistical Association*, 84, 881-896 (1989).

[21] Maronna, R.A., Martin, D., Yohai , V.J. & Salibián, B., *Robust Statistics: Theory and Methods (with R)*. Wiley Series in Probability and Statistics, 2nd Edition (2019).

[22] McDonald, J. B., & Newey, W. K., Partially Adaptive Estimation of Regression Models Via the Generalized T- Distribution . *Econometric Theory*, 4, 428-457 (1988).

[23] Miao, R., High technology investment risk prediction using partial linear regression model under inequality constraints, *Journal of Interdisciplinary Mathematics*, 21:4, 869-881 (2018), DOI: 10.1080/09720502.2018.1475067.

[24] Moberg, T. F., Ramberg, J. S., & Randles, R. H., An Adaptive Multiple Regression Procedure Based on M-Estimators. *Technometrics*, 33, 213-224 (1980).

[25] Ramsay, T. O., A Comparative Study of Several Robust Estimates of Slope, Intercept and Scale in Linear Regression. *Journal of the American Statistical Association*, 72, 608-615 (1977).

[26] Riazoshams, H., Midi ,H. & Ghilagaber, G., *Robust Nonlinear Regression: with Applications using R*. JohnWiley & Sons Ltd. (2019).

[27] Rousseeuw, P. J., & Croux, C., Alternatives to the Median Absolute Deviation. *Journal of the American Statistical Association,* 88, 1273-1283 (1993).

[28] Ruppert, D., What Is Kurtosis?: An Influence Function Approach. *The American Statistician*, 41, 1-5 (1987).

[29] Simpson, D. G., Ruppert, D., & Carroll, R. J., On One-Step GM-Estimates and Stability of Inferences in Linear Regression. *Journal of the American Statistical Association,* 87, 439-450 (1992).

[30] Stefanski, L.A. & Boos, D.D., The Calculus of M-Estimation. *The American Statistician*, 56, 29-38 (2002).

[31] Stein, C., Efficient Nonparametric Testing and Estimation. *Third Berkeley Symposium on Mathematical Statistics and Probability,* 1, 187-196 (1956), University of California Press.

[32] Welsch, A. H., & Ronchetti, E., A Journey in Single Steps: Robust One-Step M-Estimation in Linear Regression. *Journal of Statistical Planning and Inference,* 103, 287-310 (2002).

[33] Yuh, L., & Hogg, R. V., On Adaptive Regression. *Biometrics*, 44, 433-445 (1988).

[34] Wilcox, R., *Introduction to Robust Estimation and Hypothesis Testing*. Elsevier Inc. (2012).

[35] Wu, W.B., M-Estimation of Linear Models with Dependent Errors. *Annals of Statistics*, 35, 495-521 (2007).

Received August, 2021