A STUDY ON ROBUST REGRESSION METHODS

Submitted to
University of Calicut
in partial fulfillment of the requirements
for the award of the Degree of

DOCTOR OF PHILOSOPHY
in
STATISTICS

by
Lakshmi R

Under the guidance of
Dr. Sajesh T A

VERITAS VOS LIBERABIT
Estd. 1889

Research and Postgraduate Department of Statistics
St. Thomas College (Autonomous)
Thrissur - 680001, Kerala.
June 2025

CERTIFICATE

This is to certify that the adjudicators of the PhD thesis of Ms. Lakshmi R, titled ***"A STUDY ON ROBUST REGRESSION METHODS"*** have not given any directions for corrections or suggestions for change in their reports. The content of the soft copy is the same as in the hard copy.

Place: Thrissur
Date: 23-06-2025

Dr. Sajesh T A
Research Supervisor
Associate Professor
St. Thomas College (Autonomous)
Thrissur, Kerala

Dr. SAJESH. T. A, PhD
HEAD OF THE DEPARTMENT
DEPARTMENT OF STATISTICS
ST. THOMAS COLLEGE (AUTONOMOUS)
THRISSUR, KERALA - 680 001

CERTIFICATE

This to certify that the thesis titled ***"A STUDY ON ROBUST REGRESSION METHODS"*** submitted by **Lakshmi R** to the University of Calicut in partial fulfillment of the requirements for the award of the Degree of Doctor of Philosophy in **Statistics** is a record of original research work carried out by her under my supervision. The content of the thesis, in full or in parts, has not been submitted by any other candidate to any other University for the award of any degree or diploma.

Place: Thrissur
Date: 28-11-2024

Dr. Sajesh T A
Research Supervisor
Assistant Professor
St. Thomas College (Autonomous)
Thrissur, Kerala

Dr. SAJESH. T.A, PhD
HEAD OF THE DEPARTMENT
DEPARTMENT OF STATISTICS
ST. THOMAS COLLEGE (AUTONOMOUS)
THRISSUR, KERALA - 680 001

DECLARATION

I, **Lakshmi R**, hereby declare that the work presented in the thesis entitled, ***"A STUDY ON ROBUST REGRESSION METHODS"*** is based on the original work done by me under the guidance of Dr. Sajesh T A, Assistant Professor, Department of Statistics, St. Thomas College (Autonomous), Thrissur, Kerala and has not been included in any other thesis submitted previously for the award of any degree. The contents of the thesis are undergone plagiarism check using iThenticate software at C.H.M.K. Library, University of Calicut, and the similarity index found within the permissible limit. I also declare that the thesis is free from AI generated contents.

23/6/25

Lakshmi R

Dr. Sajesh T A

Dr. SAJESH. T.A, PhD
HEAD OF THE DEPARTMENT
DEPARTMENT OF STATISTICS
ST. THOMAS COLLEGE (AUTONOMOUS)
THRISSUR, KERALA - 680 001

Place: Thrissur

Date: 28-11-2024

ABSTRACT

A key goal of regression analysis is to estimate the unknown parameters of the model, a process often referred to as fitting the model to the data. For that, the most commonly used regression method is the method of ordinary least squares (OLS). The classical technique to estimate the unknown regression parameters is the OLS estimator, which minimizes the sum of squares of the residuals. OLS is a fundamental method in the field of statistical estimation. These estimators are known as the best linear unbiased estimators (BLUE). Despite of the advantages like simplicity, mathematical properties, major drawback of OLS estimates, they are sensitive to the presence of outliers. However, when the classical assumptions of regression models are satisfied, OLS calculated directly from the data yields efficient results and generally outperforms nearly all other estimation methods. Generally, regression data can have good leverage points, bad leverage points, vertical outliers, and regression outliers. There is no necessity that a single data contains all these four type of extreme values in a regression data. Thus to overcome this draw-back, it is always better to make use of robust regression methods. The thesis titled "A Study on robust regression methods" has been arranged into 5 chapters. Chapter 1 introduce the basic meaning of simple linear regression, multiple linear regression, multivariate regression, outliers, leverage points, robust estimators, their properties, importance of robust regression methods, and an extensive review of literature on robust regression methods in multiple linear regression. Disadvantages of existing methods too discussed. Objectives of the thesis also discussed there, following outline of the thesis.

Chapter 2 discuss how the classical location and scatter matrix estimators performs in the presence of outliers, and discussed the need of using robust location, scatter estimators. Introduced a novel robust shrinkage based S_n covariance estimator and shrinkage based L_1 location estimator. Sensitivity curve of the proposed estimator discussed. To check the efficiency of proposed estimators, a new robust Mahalanobis distance proposed based on robust shrinkage S_n covariance estimator and L_1 location vector. Evaluated the performance and properties of distance mea-

sure, extensively through simulation study. Real - life benchmark datasets used to evaluate the efficacy of the distance measure. Apart from benchmark dataset, we collected Lung cancer data of patients from Malabar Cancer center and evaluated the performance of the measure in identifying deviations in the data with respect to the variables Albumin, Absolute and surival days of the patients.

Chapter 3 introduces the robust shrinkage reweighted regression estimator in multiple regression, utilizing the robust covariance matrix and location median proposed in Chapter 2. The performance of the reweighted estimator is evaluated by comparing it with OLS and several other existing robust techniques. An extensive simulation study was conducted to assess the robustness and efficiency of the proposed method. Additionally, the breakdown point, affine equivariance, and sensitivity curve were empirically examined. The proposed estimator was also applied and evaluated on benchmark datasets.

Chapter 4 presents a robust shrinkage two-step reweighted regression estimator for multivariate regression. The performance of the two-step reweighted estimator is evaluated against the classical maximum likelihood based estimator and other existing robust estimators. A monte carlo simulation study was conducted to assess the robustness and efficiency of the method. The affine equivariance property of the proposed estimator was also evaluated empirically. Thus, we evaluated our proposed method in benchmark life dataset.

Finally, Chapter 5 presents the conclusion of the thesis and discusses the proposed estimators. It also outlines the scope of the research and suggests directions for future work.

I dedicate my thesis to my appa & amma, my sister, my fellow research friends and beloved teachers. I am always thankful to them for reminding me of the saying "When the going gets tough, the tough get going".

ACKNOWLEDGEMENT

I am deeply obligated to a number of persons for the fulfillment of my thesis. Without the blessing of the Almighty, this would have never happened.

I express my sincere gratitude to my Research Supervisor and my mentor, Dr. Sajesh T A, Assistant Professor, Department of Statistics, St.Thomas College (Autonomous), Thrissur. This thesis, in its current form is the result of his guidance. Without his motivation and constant support, this would not have been possible.

I would like to thank Dr. Chacko V M, Professor and Dr. Rasin R S, Assistant Professor, Department of Statistics for their constant moral support and encouragement throughout my research period.

Besides, I would take this opportunity to express my thanks to entire staff of St. Thomas College (Autonomous), Thrissur, especially, our Principal, the office staff, the faculties of Department of Statistics and everyone else for their support during my time here.

I owe my thankfulness to University of Grant Commission for providing me financial support to carry out this work under fellowship.

I extend my heartfelt gratitude to Dr. Manil T Mohan, Assistant Professor at the Indian Institute of Technology, Roorkee, and Dr. Baiju K V, Assistant Professor at Women's College, Trivandrum, for their insightful teaching and support greatly contributed to my success in securing the CSIR- UGC JRF.

I sincerely thank my current institution, Christ University, Bengaluru, for supporting me and allowing me to attend the defence. Also, I would like to express my deep gratitude to my sister, my brother in law and all other family members who stood by my side and supported me.

I would like to acknowledge my friend Dr.Dileepkumar M, Assistant Professor in the Department of Statistics at the University of Calicut, as well as my fellow research friends from the Statistics and Mathematics department at St.Thomas College (Autonomous), Thrissur - Anakha, Deepthi, Abhijith, Jincy, Amrutha, Gouthami, Ann Sania, Akshara, Athira, Greeshma, Vijayalakshmi, Megha, Sini. My fellow companions unwavering mental support and joy they provided, brought me to this moment. I would like to extend a special owe to Dr.Sujesh A S for his sincere suggestions and advice throughout my PhD journey and in the preparation of my thesis.

Last but not least, I am overwhelmed with gratitude for my PARENTS; because of them, I am who I am today.

Lakshmi R

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Statistical tools and analyses, both univariate and multivariate, are widely employed across various fields and play a vital role in research studies. One of the notable case in multivariate space is the regression problems, which is extensively utilized across various fields. Linear regression itself can be classified in to three: simple linear, multiple linear and multivariate linear regression. Regression models are nothing but true functional relationship between response variable, usually denoted as y and regressors, denoted as $\mathbf{x}$. A simple linear regression involves a single response variable and a single explanatory variable. Multiple linear regression consist of multiple regressors and single response variable. In multivariate, both regressors and response will be more than one. Consider a multiple linear regression model:

$$y = X\beta + \epsilon, \tag{1.0.1}$$

where $y = (y_1, \ldots, y_n)^t$ is a $n \times 1$ vector of n observations on response variable,

$$X = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{2p} \\ 1 & x_{31} & x_{32} & \cdots & x_{3p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}$$

is a $n \times (p+1)$ matrix of n observations on each p variables, $\epsilon = (\epsilon_1, \ldots, \epsilon_n)$ is a $n \times 1$ vector of error components which follow *i.i.d.* normal distribution with mean zero and constant variance and β is joined $(p+1) \times 1$ regression coefficient vector. The Ordinary Least Squares (OLS) estimator is a classical method for estimating unknown regression parameters by minimizing the sum of squared residuals (the estimated model errors). OLS is a fundamental method in the field of statistical estimation and these estimators are best linear unbiased estimators (BLUE). However, when the classical assumptions of regression models are satisfied, OLS estimates calculated directly from the data yields efficient results and generally outperforms nearly all other estimation methods. Despite of the advantages of OLS such as, simplicity in calculation, mathematical properties like BLUE, minimum variance, major drawback of OLS estimates, they are sensitive to the presence of outliers,[P. J. Rousseeuw & Leroy (2005)].

Statistical analysis relies on certain assumptions about the data, but these can easily be broken by outliers or extreme values. It is important to check the data carefully to identify and address any unexpected patterns. Outliers are observations that markedly differ from the majority of the dataset and can significantly influence the results of an analysis. Numerous definitions of outliers exist in the literature but there is no single, definitive definition of outliers.

- An observation having substantially large residual is an outlier [Anscombe (1960)].
- Grubbs (1969) defined outlier as one that appears to deviate markedly from other members of the sample in which it occurs.

The comprehensive studies conducted by Barnett & Lewis (1994) and D. Hawkins (1980) provides invaluable insight and extensive survey on the topic of outliers. Outliers can manifest as individual observations or clusters of two or more points. It is crucial to detect and rigorously investigate these observations, as they have a substantial influence on statistical tools. The presence of outliers often reveals inherent flaws in the model, the data, or both.

Barnett & Lewis (1994) classified the causes of outliers into the following categories:

- Inherent Variability: This clearly demonstrates the inherent variability present in the population being studied. Such variation is an essential characteristic of the population and cannot be controlled.

- Measurement error: This includes flaws in the measuring equipment or inaccuracies in recording values. Limitations of the measuring instrument introduce an additional layer of variability beyond the inherent factors.

- Execution error: This includes instances where observations are chosen that fall outside the populations of interest or when biased or incorrect samples are selected. It is crucial to recognize that we may unintentionally select a biased sample or incorporating observations who do not accurately represent the population we intended to study.

Erroneous observations due to measurement or execution error can be identified and removed from the data, but points resulting from inherent variability should remain in the dataset. Potential outliers must be thoroughly examined to confirm whether they are erroneous. The purpose of studying outliers, is to explore advanced methods for addressing these anomalous observations. This process can be tedious. Once outliers are identified in the dataset, these questions must be thoroughly addressed by the researcher to make critical decisions: Should they remove the outliers, or should these points be included in the fitting process? Additionally, which data points will be classified as outliers by the researcher? Are these outlying observations genuinely bad data points in the dataset, or do they arise from an inappropriate mathematical model fitted to the data? Is there a valid experimental rationale for removing the suspected outlier? Analyzing outliers can uncover deficiencies in the dataset and may prompt further experimentation [Beckman & Cook (1983)].

There are Univariate and Multivariate outliers. Univariate outliers can be identified by examining the distribution of values within a single feature space. When scientists carry out experiments in or out of the lab, they usually gather random measurements of a particular characteristic. Analyzing these measurements forms the basis for conclusions and possible predictions. These single-sample datasets might include one or more outliers, so caution is advised during analysis.

Multivariate outliers are more challenging than univariate outliers Gnanadesikan & Kettenring (1972), as visually detecting them is nearly impossible; they do not stand out at the extremes. Identifying outliers in multivariate data is essential in statistics because these observations can distort any statistical analysis. Some of the multivariate standard techniques like principal component analysis, linear discriminant analysis, are inherently prone to atypical observations. Nowadays, there are numerous real-world scenarios in the field of outlier detection where the data consists of a large number of variables. For instance, in neuro imaging, the data often includes rare observations caused by issues such as acquisition problems, pre-processing artifacts, or inter - subject variability. A specific example of this is Functional Magnetic Resonance Imaging (fMRI). In fMRI data analysis, even slight movements of the patient's head, as well as the subject's heartbeat and breathing, can introduce significant artifacts into the signals and noise present in the data. Data having outliers do not need to be extremely low or high values, they can simply be observations don't fit well with overall pattern of the dataset. Due to this, classical estimator based methods fails to detect multivariate outliers, because in multivariate we need to consider both the distance of an observation from the centre of the distribution of the data and the shape of the data. The Mahalanobis distance, introduced by Mahalanobis in 1930, is a key statistical measure in multivariate analysis, especially for detecting multivariate outliers. The classical Mahalanobis distance of a $n \times p$ data matrix is defined as:

$$\mathrm{MD}^2(\mathbf{x}_i) = (\mathbf{x}_i - \hat{\mu})\hat{\Sigma}^{-1}(\mathbf{x}_i - \hat{\mu})^t$$

where $\hat{\mu}$ is the classical multivariate location estimate and $\hat{\Sigma}$ is the classical covariance estimate. It calculates the distance of a data point from the mean while considering variable correlations, helping in identify observations that significantly deviate from the expected pattern. The Mean vector and Dispersion matrix are the primary components used in calculating the Mahalanobis distance (MD). Although, in theory, the Mahalanobis distance can be a useful quantitative measure for identifying potential outliers, there are several situations where it may not be as effective. For instance, there may be cases where certain observations yield large Mahalanobis distance but are not actually outliers, and conversely, observations with small Mahalanobis distance might still be outliers. These issues are typically classified into two main phenomena: Swamping and Masking. Swamping refers to the phenomenon of incorrectly labeling normal events as anomalies. Masking is the phenomenon where true outliers are hidden among a larger group of observations, preventing them from being identified due to the influence of other data points. This behavior is well-documented even in moderately large samples, particularly as the number of dimensions increases [Cerioli et al. (2009) and Riani et al. (2008)]. In a regression modelling, data can have four types of atypical observations like, good leverage points, bad leverage points, vertical outliers and regression outliers. In regression, $\mathbf{x_i} = (x_{i1}, \ldots, x_{ip})$ is a p - dimensional variable and is sometimes called as $factor\ space$. A leverage point is defined as i^{th} point $(x_{k1}, x_{k2}, \ldots, x_{kp}, y_k)$, where $(x_{k1}, x_{k2}, \ldots, x_{kp})$ is an outlier with respect to the $\mathbf{x}$ data P. J. Rousseeuw & Leroy (2005). Apart from leverage points there might occur regression outliers or vertical outliers in data. Vertical outlier is a point $(x_{k1}, x_{k2}, \ldots, x_{kp}, y_k)$ where y_k remains far away from the $\hat{y}_k$ on the regression model. Vertical outliers affect the residuals directly and often have large residual values (either positive or negative). They occur in the y direction, meaning the predictor values are typically within the expected range, but the response is unusual P. J. Rousseeuw et al. (2004). Regression outliers or influential points are those points $(x_{i1}, x_{i2}, \ldots, x_{ip}, y_i)$ which deviates from the linear pattern followed by the maximum of the data points. These points can have high leverage points and large residuals. And it distorts the regression coefficients significantly. There is no necessity that a single dataset contains all these four type of extreme values in a regression data. To address the impact of outliers, robust regression methods are highly recommended. This leads to a critical question: Which robust regression estimator is most appropriate, and how can the performance of these methods be effectively compared? Relative efficiency, breakdown point, influence function, affine equivariance are traditional criteria used to compare existing robust methodologies. A single outlier can completely invalidate the OLS estimator. Exploring an estimators capability of handling a certain percentage of outliers effectively is the concept behind the measure breakdown point. Hodges Jr (1967) was the first to explain the concept of a breakdown point, focusing solely on evaluating location in a single dimension. Hampel (1971) provided broad description of it, but it was highly mathematical in nature and asymptotic. Donoho & Huber (1983) suggested a limited sample version of breakdown point. For a sample Z of n observations,

$$Z = (x_{11}, \ldots, x_{1p}, y_1), \ldots\ldots (x_{n1}, \ldots, x_{np}, y_n).$$

Let T represents a regression estimator. When T is applied to such a sample, the result is a regression coefficient vector as $T(Z) = \hat{\beta}$. Let j of the sample data points swapped by arbitrary values and call them as corrupted sample Z'. The maximal bias generated by such contamination is then calculated as:

$$\text{bias}(j; T, Z) = \text{sup}_{Z'} \|T(Z') - T(Z)\|,$$

where the supremum is over all possible Z' and $\|.\|$ is Euclidean norm. If the bias is infinite, j outliers have a significant impact on the estimator. Thus, breakdown of the estimator T at the sample Z is defined as

$$\epsilon_n^*(T, Z) = \min \left\{ \frac{j}{n}; \text{bias}(j; T, Z) \text{ is infinite} \right\},$$

or the least amount of contamination that an estimator can tolerate is known as the breakdown point. The asymptotic breakdown point is defined as the limit of the finite sample breakdown point as n approaches infinity. Generally, the maximum asymptotic breakdown point is $\frac{1}{2}$, because if more than half of the observations are contaminated, you can't differentiate the original data from the contamination P. J. Rousseeuw & Leroy (2005).The finite sample breakdown point of OLS is $\epsilon_n^*(T, Z) = \frac{1}{n}$. That is, even the presence of single outlier in the dataset can affect least square estimators. Thus, asymptotic breakdown point of OLS is 0.

The influence function is an another important tool for assessing the robustness of an estimator. It was first introduced by Hampel (1968) and further studied in Hampel (1974) to examine the infinitesimal property of robust estimators. PJ & Stahel (1986) described it as the standardized impact of an outlier at a point on the estimator. It characterizes the infinitesimal stability of an estimator. For a robust estimator, the influence function should ideally be bounded. For an estimator θ, the influence function is defined as:

$$\text{IF}_{\hat{\theta}}(x_0, F) = \lim_{\epsilon \downarrow 0} \frac{\hat{\theta}_\infty((1-\epsilon)F + \epsilon\delta_{x_0}) - \hat{\theta}_\infty(F)}{\epsilon},$$

where δ_{x_0} is the point mass contamination at x_0 and $\downarrow$ denotes the limit from right side. $\hat{\theta}_\infty((1-\epsilon)F + \epsilon\delta_{x_0}) - \hat{\theta}_\infty(F)$ is asymptotic estimate value of the estimator, when F is the underlying distribution and x_0, a fraction of outliers. Thus, when fraction of outliers is small, $\hat{\theta}_\infty((1-\epsilon)F+\epsilon\delta_{x_0})-\hat{\theta}_\infty(F)$ can be approximated to $\hat{\theta}_\infty(F)+\epsilon\mathrm{IF}_{\hat{\theta}}(x_0,F)$, and bias is approximated by $\epsilon\mathrm{IF}_{\hat{\theta}}(x_0,F)$. Consequently, influence function of an estimator is considered to be as limit version of sensitivity curve.

Affine equivariance is another key property of a robust estimator R. A. Maronna & Zamar (2002). It ensures that the estimate will respond in a predictable manner when the samples undergo an affine transformation. Let $\hat{\mu}$ and $\hat{\Sigma}$ be the location and covariance estimates, is said to be affine equivariant if and only if, for any vector $a \ \epsilon \ R^p$ and non-singular matrix B of order p, satisfies the given condition below:

$$\hat{\mu}(BX+a)=B\hat{\mu}(X)+a, \quad \hat{\Sigma}(BX+a)=B^t\hat{\Sigma}.$$

More specifically, affine equivariant estimates adjust appropriately in response to data rotations and location-scale transformations. Relaxing this requirement can increase the number of available robust estimators or may allow to use non-affine equivariant estimators that perform better in certain cases.

1.1 Robust Regression Techniques

Robust regression is a crucial statistical tool for analyzing data that contains outliers and influential observations. The primary goal of robust regression is to provide reliable (stable) results, including estimators and tests, that perform well when the assumed model is correct and are less sensitive to small deviations from model assumptions, especially in the presence of outliers. Robust regression can be used for outlier detection, as it reduces the influence of outlying observations to enhance stability. In other words, robust regression is a statistical method that aims to reduce or eliminate the impact of influential observations or outliers, in order to derive more reliable results from the majority of the data. The true purpose is to examine the residuals from the robust fit in order to identify outliers. Robust regression techniques have been employed to tackle mainly three categories of issues related to outliers. First, problems with influential points in y-direction, second, problems with outliers in $\mathbf{x}$-space and third, problems with outliers in both $\mathbf{x}$-space, y direction. Researchers have developed a variety of procedures to address these issues.

1.1.1 Least Absolute Value (LAV) regression estimator

In 1757, Boscovich introduced the method of least absolute deviation (LAD) and Edgeworth (1887) improved the idea of Boscovich, develops the least absolute deviation as it is now. Least absolute value (LAV) regression or LAD regression which is also called as L_1 estimator, where the sum of the absolute values of the residuals is minimized to find this value.

$$\min \sum_{i=1}^{n} |e_i| = \min \sum_{i=1}^{n} |y_i - \sum x_{ij}\beta_j| \tag{1.1.1}$$

The proposed idea is straight forward, but it is not as easy as OLS estimates to calculate. Thus, it was not much recognized in history for a long time. Charnes et al. (1955) simplified the LAV based regression method to a linear programming problem and reduced the computational complexity in the method. Koenker & Portnoy (1990) have provided a comprehensive summary on LAV regression estimator. Large sample properties of the LAV regression estimates are obtained in Bassett Jr & Koenker (1978) and Pollard (1991). The LAV regression estimation technique has grown in popularity as a result of these theoretical and computational advances. LAV regression estimate is very resistant to the presence of outliers. LAV regression estimate is less affected than OLS in the presence of unusual y values, but it is unable to determine leverage values [Mosteller & Tukey (1977)]. Thus, its breakdown point is not better than $\frac{1}{n}$. Also, they have a low efficiency. Combination of less efficiency and least breakdown point makes it less attractive than other estimators.

1.1.2 Robust regression based on M estimator

The next step in this direction was M regression estimator introduced by Huber in 1964. M regression estimator is defined as a solution of the minimization problem:

$$\min \sum_{i=1}^{n} \rho(y_i - \sum x_{ij}\beta_j) = \min \sum_{i=1}^{n} e_i^2, \quad \text{or is a root of} \tag{1.1.2}$$

$$\sum_{i=1}^{n} \psi(y_i - \sum x_{ij}\beta_j)x_{ij} = 0. \tag{1.1.3}$$

where the function ρ is a properly chosen arbitrary function, gives contribution of each residual to the objective function and $\psi \equiv \frac{\partial\rho}{\partial\beta}$ is called as influence function. The following are the characteristics of a reasonable ρ:

- ρ is continuous
- $\rho(e) \geq 0$, ρ must be strictly positive and integrable
- $\rho(0) = 0$
- ρ is symmetric, $\rho(e) = \rho(-e)$ and
- ρ is monotonically increasing function, $\rho(e_i) \geq \rho(e_i^t)$ for $|e_i| \geq |e_i^t|$.

Thus, to define M regression estimator it is necessary to specify the function ρ or ψ. Solving the above set of non - linear equations yields the M regression estimate. However, the problem is that the solutions produced are not scale equivariant. Standardizing residuals by means of estimate of σ helps to overcome this, so that system becomes,

$$\sum_{i=1}^{n} \psi(\frac{(y_i - \sum x_{ij}\beta_j)}{\hat{\sigma}})x_{ij} = 0. \tag{1.1.4}$$

As in the case of M estimates of location, the median absolute deviation of σ is often used as the estimate. i.e., $\hat{\sigma} = \frac{\text{median}|e_i - \text{median}(e_i)|}{0.6745}$. The choice of a good ψ function is based on the choice of how much weight to assign outliers. Large outliers are not as heavily weighted by a monotone ψ function as they are by OLS estimates. Even when the outlying distance rises, a redescending ψ function increases the weight of an assigned outlier up to a certain distance and then drops the weight to zero. An iterative method is necessary to find M regression estimates, because residual cannot be found without the model is fitted. Two methods to solve M regression estimates nonlinear normal equations are Newton Raphson and Iteratively Reweighted Least Squares. Susanti et al. (2014) has discussed about the algorithm steps of M estimator.

- For I = 0, OLS estimator is employed to obtain the initial estimates of regression coefficients, $\hat{\beta}^{(0)}$.
- Based on the obtained initial estimates residuals, $e_i^{(0)}$ are calculated and is used to calculate initial weight estimates.

- To obtain initial weights a weight function is chosen and applied to the initial OLS estimate of residuals, $w(e_i^{(0)})$.
- For I = 1 use weighted OLS estimates to obtain $\hat{\beta}^{(1)}$. i.e., $\hat{\beta}^{(1)} = (\mathbf{x}^t\mathbf{w}\mathbf{x})^{-1}\mathbf{x}^t\mathbf{w}y$.
- The process is continued by calculating new weights using the residuals obtained from the initial weighted OLS estimates. These new weights are used in the next iteration step to obtain the estimate $\hat{\beta}^{(2)}$.
- Step 4 and 5 are repeated until the estimate $\hat{\beta}$ converges.

It is easy to study the asymptotic properties of M regression estimators when $\psi(\mathbf{x}, \beta)$ is monotone in β. Huber provided complete proof on asymptotic distribution of M estimators, i.e., M regression estimators have asymptotic variance given in equation below and are normally distributed asymptotically. (P. J. Rousseeuw et al. (1986) p. 103), Shevlyakov et al. (2008):

$$V(\psi(\epsilon), f(\epsilon)) = \frac{\int_{-\infty}^{\infty} \psi^2(\epsilon) f(\epsilon) d\epsilon}{[\int_{-\infty}^{\infty} \psi'(\epsilon) f(\epsilon) d\epsilon]^2}, \tag{1.1.5}$$

where f is the density of true error. Also, M regression estimators are asymptotically consistent. The breakdown point of M regression estimator does not depend on its probability density. M regression estimators are highly resistant against y outliers with breakdown point of 0.5 as they are robust against non-constant error variance and heavy tailed error distribution. M regression estimator is the simplest estimator and is not robust against the leverage points. The estimator is extensively used in analyzing data. M regression estimates are more efficient than OLS estimates. The major drawback of these iterative estimators are that we can never be sure a root exists, until we find it.

Bickel (1975) suggest an alternative to M regression estimator referred to as one step M estimators. In this estimation method, to estimate the parameter with the presence of nuisance parameter, first choose initial parameter for both estimates. Then nuisance parameter would be considered fixed and equal to its initial estimate. Suppose we have an initial estimator $\hat{\beta}_0$ of β in 1.0.1. Let r denote the residuals from $\hat{\beta}_0$. Let $\hat{\sigma}$ be the standard deviation of r. Then, a one step M estimator of β is written in the form:

$$\hat{\beta} = \hat{\beta}_0 + \hat{H_0}^{-1}\hat{g}_0, \tag{1.1.6}$$

where $\hat{g}_0 = \sum_{i=1}^{n} \mathbf{x}_i w_i \psi(v_i r_i/\hat{\sigma})$ for some odd function ψ, $w_i = w(\mathbf{x}_i)$ and $v_i = v(\mathbf{x}_i)$ are weight functions and an appropriate matrix $\hat{H}_0$, One step Huber M estimator has $w_i = v_i = 1$, One step Mallow's estimator has $v_i = 1$, One step Andrew's estimator has $w_i = 1$, and One step Hill and Ryan estimator has $v_i = w_i$. The common choices for the matrix $\hat{H}_0$ are the Newton Raphson form or iteratively reweighted least squares form. The major advantage of one step M regression estimator is easy to compute and they have asymptotic properties.
Redescending M regression estimators are famous ψ type M estimators, in which ψ functions are non-decreasing at the origin, but decreases to zero in the region far from the origin. According to Holland & Welsch (1977), redescending estimators can be characterized as soft - redescending (pseudo-convex) and hard redescending (quasi-convex), depending whether the corresponding influence functions are nearly null and exactly null for polluted values of high magnitude. Hard redescending regression estimators influence function, i.e., ψ function is usually built with discontinuous segments, such as Hampel's three-part estimator. M regression estimators can classify based on the mathematical features of the influence function as: Non robust, quasi robust, robust monotonous, robust soft - redescending estimators and robust hard redescending estimators. De Menezes et al. (2021) has reviewed 48 different types of M regression estimators in his paper.

1.1.3 Regression based on generalised M Estimator

Traditional M estimators are not "qualitatively robust" (Hampel (1971) and Hampel (1974)). This is because they exhibit a large asymptotic bias when the joint distribution of $(\mathbf{x}, y)$ follows the model 1.0.1 only approximately. This is as a result of the fact that they have unbounded influence function. Also, M estimators are vulnerable to the leverage points. To overcome this generalized M regression estimator were introduced. The influence of extreme $\mathbf{x}_i$ values were bounded by weight functions. It was Mallows (1975) proposed $\sum_{i=1}^{n} w(\mathbf{x}_i)\psi(r_i/\hat{\sigma})\mathbf{x_i} = 0$. In this estimation method, outliers are handled using the usual M estimator and leverage points are down weighted by using appropriate weight function. The general class of GM regression estimators is

$$\sum_{i=1}^{n} w_i(\mathbf{x_i})\psi\left\{\frac{e_i}{v(\mathbf{x_i})\hat{\sigma}_e}\right\}\mathbf{x_i} = 0, \tag{1.1.7}$$

where ψ the score function generally used is Huber or bi-weight function. This type of GM estimator is called as Schweppe type. The model matrix $\mathbf{x}$ determines the weights w_i and v_i initially. The initial estimate is OLS and the scale estimate is found by scaling the median of the absolute value of the OLS residuals. Final estimates are obtained by iterative procedure. Mallow's proposed another GM regression estimator. Mallow's method down weight both tiny and large residuals which is the main difference between the two estimators. Schweppe's down weight minor residuals only. The current generation of GM estimators has a breakdown point of at most $1/p + 1$ where p is the dimension of $\mathbf{x}_i$. Various GM estimators developed over different periods, more frequently discussed are Krasker (1980), Krasker & Welsch (1982) and Marazzi (1993).

After the development of robust M regression estimators and their generalized version, a question raised whether high breakdown point estimators could develop, as an answer to the question, repeated median regression estimator proposed by Siegal, with a 50% breakdown point. Repeated median estimator is a variation of Theil - Sen estimator. Theil (1950) advocated using the median of pairwise slopes as a slope estimator. Sen (1968) improved this estimator by adding the ability to handle ties. TSE is a robust estimator with a breakdown point of 29.3% and a bounded influence function. In addition, they have a high asymptotic efficiency. Theil - Sen regression estimator is only formulated for simple linear model. Many authors like Oja & Niinimaa (1985), Zhou & Serfling (2008) have made their efforts to extend TSE to multiple regression model. However, it is technically difficult and causes the analysis of the properties to be delayed. Thus, Siegel (1982) proposed highly efficient repeated median regression estimator and is defined as:

For any n observations $(x_{i1}, y_1), (x_{i2}, y_2), \ldots, (x_{in}, y_n)$ the objective is to find the parameter that fits these points exactly. The j^{th} coordinate of the parameter vector is represented by $\beta_j(i_1, \ldots, i_n)$. Then the regression estimator is described as:

$$\hat{\beta}_j = \text{med}_{i_1}(\ldots(\text{med}_{i_{p-1}}(\text{med}_{i_p}\beta_j(i_1, \ldots, i_p)))\ldots). \tag{1.1.8}$$

The estimator requires the consideration of complete subset of p observations and requires a lot of time to compute. It is easy to apply in problems with small p. For linear $\mathbf{x}$ - value transformations, the approach is not equivariant. When the contamination in the data is considerable, the repeated median regression estimator fails to distinguish between good and bad points of the data, despite having a breakdown point of 50%. Following the failure of repeated median regression estimators, Rousseeuw develop two high breakdown simple estimators Least Median Square estimator and Least Trimmed Square estimator.

1.1.4 Least Median Square (LMS) regression estimator

P. J. Rousseeuw (1984) develop Least Median of Squares (LMS) regression estimator, an L estimator which is calculated by replacing the objective squared residuals sum in OLS by median of the squared residuals. Rousseeuw's this proposal was based on the proposal of Hampel (1975)-p.380). The estimates are found by

$$\min \operatorname{med}(y_i - \sum x_{ij}\beta_j)^2 = \min \operatorname{med}(e_i^2), \tag{1.1.9}$$

where med denotes the median. LMS regression estimator always has a unique solution. P. Rousseeuw & Yohai (1984) in his paper has stated that, if p >1, and the observation are in general position, then the breakdown point of LMS estimator based regression method is $([n/2] - p + 2)$, The main concept of LMS regression estimator is to reduce the dispersion of the residuals. It has a breakdown point of 0.5. Even though LMS estimator has a high breakdown point, it has low efficiency because of its slow convergence rate. P. J. Rousseeuw & Croux (1993a) in their paper have shown that LMS estimator has a relative efficiency of 37%. P. Rousseeuw & Yohai (1984) in his paper itself proposed methods to overcome the slow convergence rate of LMS regression estimator by one step M estimators in which initial estimates are obtained by LMS regression estimator or by making use of an objective function other than that of LMS. Thus, least trimmed square regression estimators developed as an alternative to LMS estimator.

1.1.5 Least Trimmed Square (LTS) regression estimator

Least Trimmed Square regression estimator (LTS) is yet another L estimator based method developed by P. Rousseeuw (1985). The method is extended from trimmed mean. The method minimizes the sum of the trimmed squared residuals. i.e.,

$$\min \sum_{i=1}^{q} e_{(i)}^2, \tag{1.1.10}$$

where q have to satisfy $\frac{n}{2} \leq q \leq n$. The constant q determines the breakdown point of the LTS regression estimator. LTS estimator definition implies that $n-q$ observations with largest residual will not affect the estimator. The minimization of the objective function 1.1.10 implies, choose a sub sample of size q observations and find β minimizing the sum of squared residuals of the sub sample. Repeating this for every sub sample, we get nC_h candidates for LTS regression estimate and among them one that gives the smallest value of the objective function is the final estimate. LTS regression consists of finding β belongs to R^p such that sum of q smallest squared residuals is minimum. As mentioned above the value of the constant q determines the breakdown point of the estimator, the choice depends mainly on the purpose for which we use LTS regression estimator. The breakdown point of LTS regression estimate reach the upper bound $([(n-p)/2]+1)/n$ for regression equivariant estimators if the trimming constant $q = [n/2] + [(p+1)/2]$. Setting $q = (n/2) + 1$ ensures the estimator a breakdown of 0.5. It is always better to evaluate LTS regression estimator for a wide range of trimming constant values and compare how the estimator change with increasing value of the constant. Such an analysis provides us an insight on amount of contamination and estimate from the data. LTS regression estimator is scale and affine equivariant. Víšek (2006) in his paper has shown the asymptotic normality and Víšek (2006) $\sqrt{n}$ consistency of LTS estimator in a general linear regression model with continuously distributed disturbances. One of the drawbacks of LTS regression estimator is non-continuity of LTS objective function. Due to this LTS regression estimator sometime might possess highly sensitive to a change of one or several observations. LTS regression estimates efficiency varies depending on the trimming constant value and the outliers in the data. Stromberg et al. (2000) in their paper made clear that LTS regression estimator has a low efficiency of 8%. Its low efficiency makes it non desirable estimator. Though this method has low efficiency, the estimator is used in other robust regression methods as initial estimates. For example, LTS regression estimates are used as the initial estimate in GM based regression estimator proposed by Coakley & Hettmansperger (1993).

1.1.6 Robust regression based on S estimator

Since LTS and LMS regression estimators has slow convergence rate, P. Rousseeuw & Yohai (1984) introduced high breakdown value estimator that minimizes the dispersion of the residuals with high asymptotic efficiency and better convergence rate of objective function than LTS regression estimates. Robust regression based on S estimate is the solution to the smallest possible dispersion of the residuals. S based regression estimate is not simply minimizing the variance of the residuals; it minimizes the robust M estimate of residual scale. That is S regression estimator is obtained by minimization of the dispersion of the residuals. i.e., Minimize $s(r_1(\beta), r_2(\beta), \ldots, r_n(\beta))$ with respect to β, with final scale estimate:

$$\hat{\beta} = s(r_1(\hat{\beta}), r_2(\hat{\beta}), \ldots, r_n(\hat{\beta})).$$

The dispersion $s(r_1(\hat{\beta}), r_2(\hat{\beta}), \ldots, r_n(\hat{\beta}))$ is obtained as an answer of

$$\frac{1}{n}\sum_{i=1}^{n} \rho(\frac{r_i}{s}) = b,$$

where b is a specified constant as $b = \mathrm{E}_\Phi[\rho(e)]$ and Φ depicts the standard normal distribution. The ρ function should satisfy the conditions same as that in M estimator. S regression estimator is scale and affine equivariant. For any ρ function satisfying the above conditions along with the condition $\frac{b}{\rho(c)} = \frac{1}{2}$ has a breakdown of 50%. The major drawback of S regression estimator is its low efficiency. Generalised S (GS) based regression estimator proposed by Croux et al. (1994) to overcome the low efficiency of S regression estimator. The proposed estimator is calculated by finding a GM regression estimator of the scale of the residuals. Least quartile difference estimator is a unique case of GS regression estimator. Croux et al. (1994) p.1271 in his paper stated that although the estimator has high efficiency, but it possess "slightly increased worst case bias" which makes this regression estimator less acceptable. The fact that the objective function of GS regression estimator is independent of the intercept term is a key feature. When it comes to models with an asymmetric error distribution, GS regression estimator is ideal, since the objective function only depends on r_i through $|r_i|$, as a result, positive and negative residuals of same size are given the same weight. Yohai & Zamar (1988) proposed tau estimator obtained by minimizing estimate for the scale of the residuals.

1.1.7 Robust regresssion based on MM estimator

Regression based on MM estimator was proposed by Yohai (1987) which is most commonly and widely used regression estimator. The first step involves calculating an initial robust estimate of the regression parameters, which prioritizes a high breakdown point but not necessarily high efficiency. In the second phase, this initial estimator is used to compute a robust M estimate of the residuals' scale. The final stage focuses on obtaining an M estimate of the regression parameters, starting from the initial regression estimator. In practical applications, the initial estimators typically used are LMS or S regression estimators, paired with Huber or bi-square functions. MM regression estimation is a combination of high breakdown value estimator with high efficiency of approximately 95% relative to OLS estimates. The name itself hints that M estimation is used more than once to find the final regression estimates. MM regression estimator is affine equivariant. They have a breakdown of 50%. These inherits exact fit property from the initial estimate. MM regression estimator is consistent and asymptotic normal too. Mathematical proof of all these properties is provided by Yohai (1987). MM regression estimator needs a high breakdown estimator as initial estimate. By adjusting the constants required for the estimators across the three stages, MM regression estimates can achieve high efficiency without compromising their breakdown point. However, as noted by Yohai (1987), increasing the constant that controls efficiency makes the estimates more vulnerable to outliers. An MM regression estimation procedure is as follows:

- Initially the regression coefficient estimate and residuals are obtained using a high resistant regression estimator of 50% breakdown point; S estimate is commonly used for this purpose.
- Using residuals got from step 1, obtain the residual scale using M estimate.
- In the first iteration of weighted least squares, initial estimate of the residuals from step 1 and residual scale estimates from step 2 are utilized to determine regression coefficients using M estimate. $\sum_{i=1}^{n} \mathbf{w}_i(\frac{e_i^{(1)}}{\hat{\sigma}_e})\mathbf{x_i} = 0$.
- New weights $w_i^{(2)}$ are calculated based on the residuals obtained from initial weighted least squares in step 3.

- The above steps are reiterated until convergence

1.1.8 Robust regression based on covariance method

The OLS estimator 1.0.1 can be expressed in the following way too proposed by R. Maronna & Morgenthaler (1986). Let $\mathbf{z} = (\mathbf{x},\ y)$ represent the joint variable of the response and carriers. μ be the location and Σ be the scatter matrix of $\mathbf{z}$. Partitioning of μ and Σ with respect to $(\mathbf{x},\ y)$ be:

$$\mu = \begin{bmatrix} \mu_{\mathbf{x}} \\ \mu_y \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{bmatrix}. \tag{1.1.11}$$

The classical MLE of mean $\hat{\mu}$ and the covariance estimator $\hat{\Sigma}$ are traditionally used to estimate them. Specifically, the OLS estimators of β and α can be expressed as functions of the $\hat{\mu}, \hat{\Sigma}$ components as follows:

$$\hat{\beta} = \hat{\Sigma}_{xx}^{-1}\hat{\Sigma}_{xy}, \quad \hat{\alpha} = \hat{\mu}_y - \hat{\beta}^t\hat{\mu}_{\mathbf{x}}. \tag{1.1.12}$$

Major drawback of the above mentioned estimators is, classical estimators of location and scatter are sensitive to the presence of the outliers. Robustification of the classical estimators of scatter, location improve the performance of the regression estimator and leads to the formulation of a new robust regression estimator. Such developed regression estimators could be used as a robust alternative to classical OLS, since they inherits all the properties of location, scatter estimators and outperforms OLS. The literature presents, a numerous proposals for robust location and covariance estimators. R. Maronna & Morgenthaler (1986) explored multivariate M regression estimators, while Croux et al. (2003) examined S based multivariate regression estimators. Cabana et al. (2020) proposed reweighted RMDS based covariance approach in multiple regression. Lakshmi & Sajesh (2023) proposed S_n based covariance approach in multiple regression.

1.1.9 Robust and efficient weighted least square regression

Gervini & Yohai (2002) proposed the “robust and efficient weighted least squares” estimator (REWLSE). They demonstrate that this method achieves both breakdown point and full efficiency under Gaussian errors. The concept is similar to

weighted least squares, but the weights are determined using an initial robust estimator. The weighting scheme employs hard rejection (0 or 1), with the cut-off determined by the distribution of the standardized absolute residuals calculated from the initial robust estimators of the regression parameters and scale.
In conclusion, each of these alternatives of OLS estimate has their own drawbacks. While regression based on M estimation is robust to outliers in the response variable, it does not effectively handle outliers in the explanatory variables (leverage points). As a result, the estimator shares the same breakdown point as OLS estimate. Likewise, regression based LAD, R estimator, and GM estimator experience similar low breakdown point. To mitigate this lackness, alternatives such as robust regression based on LTS, LMS, S estimates, MM estimates, the covariance approach using S estimators, and REWLSE are viable options. Nonetheless, regression based LTS, LMS, and S estimates demonstrate low efficiency. Although the GS based regression estimator enhances efficiency compared to the regression based S estimator, it remains inadequate. On the other hand, the MM regression estimator, and the REWLSE estimator stand out as the best alternatives because they combine a high breakdown point with high asymptotic efficiency. It is important to highlight that although some of the estimators mentioned have a high breakdown point, their computation can be quite difficult, especially when working with large datasets or in high dimensions. This is why approximate algorithms are necessary for this task. However, the downside is that they result in reduced performance regarding consistency and breakdown point compared to the exact theoretical estimator. The problem worsens as the sample size n and/or the dimension p of the samples grows Stromberg et al. (2000), D. M. Hawkins & Olive (2002). Moreover, with all these estimators, decisions must be made regarding which tuning constant to select, which residual function to use, and which initial estimator to adopt. The situation becomes increasingly complex when dealing with real data, given the multitude of choices involved.

The specific objectives of our thesis is as follows:

- To examine and compare the existing robust linear regression estimators.
- To introduce a new robust regression estimator derived from existing robust

estimation technique.

- To propose a new robust estimator, extending to a robust measure and assess the properties of the proposed estimator.
- To asses the performance of the above proposed robust measure in real - life datasets.
- To introduce a new robust multiple regression estimator based on above mentioned robust estimator and evaluate the performance in benchmark datasets.
- To propose a new robust multivariate regression estimator based on above defined robust estimator and evaluate the performance in benchmark datasets.

1.2 Outline of the thesis

The structure of the thesis looks as follows: There are six chapters in this thesis. Current chapter forms first chapter of the thesis.

Chapter 2 contains review on some prominent robust outlier detection methods and their respective robust covariance matrices. A robust Shrinkage $Sh-_n$ covariance estimator developed and adapted to the Mahalanobis distance measure for application in multivariate datasets to detect outliers. The efficacy of proposed distance measures is evaluated extensively through monte carlo simulations. Not only distance measure, proposed robust covariance estimator's property also evaluated in this chapter. A real world dataset of Lung cancer patients from Malabar Cancer Centre (MCC) is taken for the assessment of distance measure. Patients Albumin, Absolute Neutrophil count, and survival days are used for assessment. Apart from real world dataset, seven benchmark datasets used to evaluate the distance measure.

Chapter 3, presents robust Shrinkage reweighted regression estimator in multiple regression using the robust covariance estimator proposed in chapter 2. The performance of reweighted estimator is assessed by comparing it with OLS and some of the other existing robust regression techniques. Extensive simulation study done. Also, checked the breakdown, affine equivariance property and sensitivity curve of proposed regression estimator empirically. Evaluated the performance of proposed estimator in benchmark datasets too.

Chapter 4, presents robust shrinkage two - step reweighted regression estimator in multivariate regression. The performance of two - step reweighted estimator is assessed with classical estimator and other existing robust estimators. Monte carlo simulation study done to check the robustness and efficiency of the estimator. Affine equivariance property of the proposed estimator is evaluated empirically. Investigated the performance in benchmark datasets too.

Last, chapter 5 gives the conclusion of the thesis and discussion on proposed estimators, and chapter 6 gives the scope and recommendations of our work.

Chapter 2

Robust estimation and outlier detection in multivariate data

2.1 Introduction

The estimation of a covariance matrix and sample mean are fundamental aspect of multivariate statistics, with widespread applications in fields such as econometrics, biostatistics, signal processing, neuroimaging, climatology, and many others. Classical estimates become unreliable in the presence of outliers. Outliers can influence the classical estimates $\hat{\mu}$, $\hat{\Sigma}$ and the estimates are highly influenced [PJ & Stahel (1986) and P. J. Rousseeuw & Van Zomeren (1990)]. This implies that an individual rogue observation or groups of observations that diverge from the main data structure can significantly affect these estimators and there by their respective application. To tackle this issue, it is always better to make use of some robust estimates of location and covariance matrix. To resolve the vulnerability of classical estimates, there are lot of robust reweighted alternatives available in literature. Multivariate outliers result from inconsistent measurements across multiple variables. When dealing with multivariate data, depending solely on univariate methods to detect extreme values within individual variables may neglect significant inter variable relationships. To effectively detect multivariate outliers, distance- based, depth- based, clustering- based, distribution- based techniques are available in the literature. Distance- based techniques were employed to detect outliers by considering the complete dimensional distance between a data point and its closest neighbors within a dataset. Mahalanobis Distance (MD), introduced by

Mahalanobis, is a commonly used distance metric for outlier detection in distance-based methods,that considers the deviation of an observation from the mean vector. In the case of multivariate Gaussian data, it is established, as per Gnanadesikan & Kettenring (1972) work, that the distribution of the squared Mahalanobis distance (MD^2) follows a chi-squared distribution with p degrees of freedom, where p represents the dimension of the data or the number of variables. Mahalanobis distance identifies multivariate outliers by considering deviation from mean and covariance. The Mean vector and Dispersion matrix are the primary components used in calculating the Mahalanobis distance (MD). Outliers can influence component estimates, making classical Mahalanobis distance inappropriate for detection. Robust Mahalanobis Distance (RMD) uses robust estimates of location and scatter parameters for improved detection.

2.2 Robust location and covariance estimates

2.2.1 M estimation

M estimation is an affine equivariant robust estimation procedure that is based on distances and is used for estimating location and scatter. This method was introduced by R. A. Maronna (1976), however, it was originally proposed by Huber (1992) for estimating a univariate location parameter. Later on,Huber & Ronchetti (2011) defined location vector $\mathbf{t}$ and covariance matrix $\mathbf{V}$ optimum solutions to the equations as follows:

$$\frac{1}{n}\sum_{i=1}^{n} u_1 \left[\left\{ (\mathbf{x_i}-\mathbf{t})^t \mathbf{V}^{-1} (\mathbf{x_i}-\mathbf{t}) \right\}^{1/2} \right] (\mathbf{x_i}-\mathbf{t}) = 0 \tag{2.2.1}$$

$$\frac{1}{n}\sum_{i=1}^{n} u_1 \left[\left\{ (\mathbf{x_i}-\mathbf{t})^t \mathbf{V}^{-1} (\mathbf{x_i}-\mathbf{t}) \right\}^{1/2} \right] (\mathbf{x_i}-\mathbf{t})(\mathbf{x_i}-\mathbf{t})^t = \mathbf{V}, \tag{2.2.2}$$

The u_1 and u_2 in the above equations are weight functions, generally used for down weighing the outlier effect in the data. The major drawback of the method R. A. Maronna (1976) is its least breakdown value, which makes it unsuitable for higher dimensional dataset.

2.2.2 MVE and MCD estimator

Minimum Volume Ellipsoid (MVE) and Minimum Covariance Determinant (MCD) were introduced by P. Rousseeuw (1985) and are widely accepted than M - estimation method for robust estimation owing to its high breakdown point. The Minimum Volume Ellipsoid (MVE) method finds the smallest ellipsoid that contains at least $h = [n/2] + 1$ data points, where n is the total number of samples. Here the center of the ellipsoid gives the location, and the shape of the ellipsoid represents the covariance. MCD is based on the search of smallest covariance determinant that accounts at least half of the data points. MCD and MVE have a breakdown of 50% which makes it suitable for detecting outliers in high contaminated datasets. The main disadvantage of robust methods is the complexity involved in finding the optimal subgroup which leads to high time consumption in getting results. The method required naive sub-sampling to minimize the objective function of the MCD. P. J. Rousseeuw & Leroy (2005) introduced a resampling algorithm that focuses on identifying a small set of good data points rather than bad points in which the mean and covariance subset of size $p + 1$ drawn from the data are calculated. The ellipsoid is modified by inflating or deflating it to include h observations, and this adjustment is repeated to preserve the minimum volume of the ellipsoid. P. J. Rousseeuw & Leroy (2005) proposed a reweighting technique to enhance the efficiency of the Minimum Volume Ellipsoid (MVE) method. In the reweighted MVE approach, the mean vector and covariance estimates are recalculated only for those samples whose Mahalanobis distance from the initial MVE mean vector and covariance matrix falls below a specified threshold, which is the quantile of the chi - square distribution with p degrees of freedom.

2.2.3 S estimator

S estimator is a method for multiple regression that minimizes a smooth and symmetrical function of residuals, proposed by P. Rousseeuw & Yohai (1984). Lopuhaä (1989) extended the definition of the S estimator to the mean vector and covariance matrix $(\mathbf{t}, \mathbf{V})$, where $\mathbf{V}$ positive semi definite matrix obtained by minimizing $|\mathbf{V}|$ subject to:

$$\frac{1}{n}\sum_{i=1}^{n} \nu \left[\left\{ (\mathbf{x_i} - \mathbf{t})^t \mathbf{V}^{-1} (\mathbf{x_i} - \mathbf{t}) \right\}^{1/2} \right] = b_0$$

In general, to ensure consistency under a normal distribution, the constant b_0 can be computed as $\mathrm{E}_{0,I}(\nu\|\mathbf{X_0}\|)$. S - estimators meet the first - order conditions of M - estimators and have a high breakdown point. Rocke (1996) discussed various options for ν - functions.

2.2.4 Hadi Forward Search Method

Hadi (1992) proposed Hadi forward search method, a non-affine equivariant MVE based method for multivariate outlier detection. This method addresses the limitations of the classical MVE resampling method and provides a more robust approach. In this method, a coordinate-wise median is calculated from the original data, and this median vector is then used to estimate the covariance. The observations corresponding to the $[(n+p+1)/2]$ smallest distances are chosen, and the classical mean vector and covariance matrix estimates from this subgroup are used to calculate the Mahalanobis distances for all observations in the original dataset. Once again, a subset of size $p+1$ with the smallest distances is selected, which is referred to as the basic subset.

2.2.5 Atkinson's Forward Search Method

A. Atkinson & Mulira (1993) proposed a forward search algorithm that is similar to Hadi's forward search method. The algorithm is based on the concept of the MVE resampling method. Atkinson's forward search algorithm starts by estimating the mean vector and covariance matrix from a randomly selected subset of size $m = p+1$. Next, the covariance is adjusted to include h observations from the original data, and the volume of the covariance is calculated. The resulting covariance is then used to compute the Mahalanobis distance for all observations. This process is repeated by selecting $m+1$ observations with the smallest distances, during which any observation whose squared Mahalanobis distance exceeds the critical value is identified as a potential outlier.

2.2.6 Hybrid Algorithm

Rocke & Woodruff (1993) proposed a computationally faster Hybrid algorithm. The method combines robust distance-based approaches for multivariate outlier detection using a hybrid of Hawkins' Forward Search Algorithm (FSA) and Atkinson's forward search method. It works in two phases. Phase I starts with Hawkins'

Forward Search Algorithm (FSA) to approximate a robust Minimum Covariance Determinant (MCD) estimate for location and shape. Then use Atkinson's forward search method to refine the estimates by identifying a subset of non- outlier observations. These are used to improve the estimates of the mean vector and covariance matrix. Phase II calculate Mahalanobis distances for all observations using the refined estimates. Then scale the distances and compare them to a critical value from the chi-square distribution to identify outliers. This hybrid method combines the global optimization of MCD with the efficiency of forward search algorithms for robust detection of multivariate outliers. It is worth noting that these algorithms require very large samples and long computation time to detect moderately large amounts of contamination from large dimensional datasets.

2.2.7 FAST - MCD estimator

EStimator (1999) proposed the Fast - MCD, since the MCD has greater theoretical advantages compared to the MVE, but it faces computational challenges in identifying the half-sample subset that minimizes the covariance determinant, a considerably more effective modification in the MCD method by improving its subsampling algorithm and is readily available in Matlab and R software. The C - step algorithm, starts by selecting a random subset of observations and calculating its mean vector and covariance matrix. Based on these estimates, all observations are ordered using Mahalanobis distance. A new subset is then chosen, consisting of the observations with the smallest Mahalanobis distance and results in a covariance determinant that is less than or equal to the previous subset's determinant. This process is repeated iteratively until the optimal MCD solution is reached, typically requiring only two iterations for convergence. Unfortunately Fast - MCD still takes a lot of time to run with higher dimensions because the number of possible solutions increases rapidly. This makes the method slow and computationally expensive, even for moderately sized problems.

2.2.8 BACON Method

To develop computationally efficient methods for robust estimation and multivariate outlier detection, Billor et al. (2000), proposed the Blocked Adaptive Computationally Efficient Outlier Nominator (BACON) method that uses iterative procedure without optimal conditions. The BACON method includes two versions of outlier detection algorithms: the first is affine - equivariant with a breakdown point of 20%, and the second is approximately affine - equivariant with a breakdown point of 40%. The BACON method is based on Hadi's forward search method Hadi (1992) and begins by selecting an initial outlier - free subset. This initial subset can be chosen in one of two ways: first, it can consist of $p+1$ observations that have the smallest Mahalanobis distance, calculated using the mean vector and covariance matrix of the entire dataset. In the second case, the initial subset includes $p+1$ observations with the smallest Mahalanobis distance, based on the component - wise median vector and its covariance matrix. The second algorithm provides a more robust and less affine - equivariant estimator because the component - wise median is not affine - equivariant. The mean vector and covariance matrix from the initial subset are then used to calculate the Mahalanobis distance for all observations.

2.2.9 Kurtosis estimator

Peña & Prieto (2001) developed the projection based Kurtosis algorithm for robust estimation. This algorithm is based on analyzing the projections of the sample points onto a specific set of directions that are obtained by maximizing and minimizing the Kurtosis coefficient of the projections. The kurtosis coefficient indicates the peakedness of a distribution. Outliers from the symmetric contamination model elevate the Kurtosis coefficient. A small number of asymmetrical outliers also results in a higher Kurtosis coefficient. However, when there are many asymmetrical outliers, the Kurtosis coefficient decreases to a very low value. The algorithm uses random directions generated by a stratified sampling scheme, combining both specific and random directions to provide an effective solution. It also has limitations in high dimensional datasets and correlated samples.

2.2.10 Orthogonalized Gnanadesikan - Kettenring estimator

The Orthogonalized Gnanadesikan - Kettenring (OGK) estimator was developed by R. A. Maronna & Zamar (2002), applying a general method to the pairwise robust scatter matrix from James & Stein (1992). The authors of the paper have explored the use of reweighted and non - reweighted scatter matrices based on the Orthogonalized Gnanadesikan - Kettenring (OGK) estimator in different scenarios. They have proposed a robust Mahalanobis distance for outlier detection, with different quantiles used as cutoffs to determine the performance of the proposed distance. The authors have mentioned that the non - reweighted OGK - based Mahalanobis distance performs better in simulation studies. They have also found that using reweighted scatter has improved the equivarince property of the measure, and in real - life datasets, reweighted scatter - based Mahalanobis distance outperforms other compared measures. Furthermore, the authors have employed eigenvalue orthogonalization to make the matrix positive definite.

2.2.11 Comedian estimator

Sajesh & Srinivasan (2012) proposed Comedian covariance estimator and based on that a robust distance measure. Falk (1997) introduced a non-positive semi-definite robust covariance estimator called the comedian, which generalizes the Median Absolute Deviation (MAD) introduced by Huber & Ronchetti (2011). They have used a similar orthogonalization technique as that of Maronna and Zamar to solve the lack of positive definiteness of the covariance estimator.

Kunjunni & Abraham (2022) introduced a robust Mahalanobis distance based on multidimensional S_nCov covariance estimator which is an adapted method and shown that their proposed estimator is better than other existing estimators. Similar to the repeated median measure in regression, authors proposed S_nCov estimator. As an application of proposed covariance estimator, they made use of S_nCov covariance estimate and component wise median as location estimate.

Cabana et al. (2021) proposed a robust Mahalanobis distance (RMDS) based on robust shrinkage comedian covariance estimate. The authors have proposed a number of robust distance measures with different combinations of location and scatter estimates in their paper. A combination of shrinkage comedian covariance estimate with location parameter as shrinkage L_1 - median is developed and the efficiency of proposed estimates investigated by applying to robust distance measure in their work

Section 3 of this chapter propose the shrinkage estimator and examined the efficiency, sensitivity curve of estimate, introduce a robust Mahalanobis distance measure, a wide simulation study of the proposed measure done along with sensitivity analysis and investigated the properties of the distance measure empirically. Section 4 shows the application of distance measure based in real - life datasets. Section 5 includes the summary of this chapter.

2.3 Robust Mahalanobis distance based on Shrinkage estimators

Detecting outliers is crucial, as they significantly impact estimation and inferential processes. Misidentifying normal observations as outliers can also impact parameter estimation. Accurate identification and addressing of outliers are paramount in statistical analysis. Various techniques are mentioned in the literature Barnett & Lewis (1994) for detecting outliers. The Mahalanobis distance is a commonly used quadratic distance - based metric that consider the covariance structure of the data and location of the data. Mahalanobis distance (MD) of a $n \times p$ $\mathbf{X}$ matrix is defined as:

$$\mathrm{MD}(\mathbf{X}) = (\mathbf{X} - \mathbf{C})^{\mathbf{t}}\mathbf{V}^{\mathbf{-1}}(\mathbf{X} - \mathbf{C}).$$

Here $\mathbf{V}$ is the sample covariance matrix of the data and $\mathbf{C}$ is the classical location of the data. Observations with large MD value is usually considered as an outlier. It is important to note that outliers do not necessarily need to have higher values for MD due to the masking problem. Similarly, not all observations with large MD values are necessarily outliers due to the swamping problem. It is because of the impact of outliers on classical location and scale estimators, masking and swamping happens.

Thus, the estimation of dispersion matrices is a crucial issue in multivariate analysis. The sample covariance matrix $1/n \sum_{i=1}^{n} x_i x_i^t$, where $x_1, x_2, \ldots, x_n \epsilon\ \mathbf{C}^P$ are data samples, is a very convincing option since it is structurally simple and it asymptotically converges to the population covariance matrix as $n \longrightarrow \infty$ with fixed p. However sample covariance matrix has some major drawbacks like, it is sensitive owing to the presence of outliers, it is not invertible for $n < p$, and it is a subpar estimate of the true covariance matrix whenever a sample of n observations and p variables are of same. Several strategies have been put out to deal with the $n < p$ problem, the most prevalent being the linear shrinkage method. Robust covariance estimators can effectively mitigate the sensitivity of traditional sample covariance matrices to

the presence of outliers. But conventional robust covariance estimators, such as the M, MCD and MVE estimators, which were created for non-Gaussian samples are ineffective for solving issues with large dimensional and small sample size. Leodit - Wolf proposed shrinkage covariance matrix by shrinking the sample covariance matrix towards a scaled identity matrix and proposed a shrinkage coefficient which is asymptotically optimal for any distribution. Thus, robust covariance estimators focus on robustness again outliers and non - Gaussian observations whereas shrinkage covariance estimators focus on improving the accuracy of covariance estimate, especially in higher dimensional with limited sample size.
The principle behind shrinkage estimation lies in the notion of ”shrinking” an estimator $\hat{\mathbf{E}}$ towards a target estimator $\hat{\mathbf{T}}$, which serves to effectively diminish estimation errors. By carefully calibrating the level of shrinkage, represented by the shrinkage intensity η, the resulting shrinkage estimator can surpass $\hat{\mathbf{E}}$ in terms of reducing estimation errors, provided certain general conditions are metJames & Stein (1992).

$$\hat{\mathbf{E}}_{\mathbf{Sh}} = (1-\eta)\hat{\mathbf{E}} + \eta\hat{\mathbf{T}}. \tag{2.3.1}$$

Utilizing a shrinkage estimator offers a significant benefit by balancing bias and variance. This technique can be employed to estimate both central tendency and dispersion parameters. When applied to covariance matrices, shrinkage estimation provides an added advantage: it yields a well - conditioned positive definite estimate. This attribute is particularly valuable in scenarios where the estimate needs to be inverted to calculate a Mahalanobis distance.

2.3.1 Location estimator

Let $\mathbf{x}$ be $n \times p$ matrix with size n, number of variables p and each $\mathbf{x_j} \,\epsilon\, \mathbf{R^n}$ $(j = 1, 2, \ldots, p)$. A highly efficient and robust estimator of central tendency known as the L_1-median, denoted as $\hat{\boldsymbol{\mu}}_{MM}$, has been extensively studied by Lopuhaä & Rousseeuw (1991), Oja (2010) and it is widely considered in various applications. It is defined as:

$$\hat{\boldsymbol{\mu}}_{MM} = argmin_{\mathbf{x_m},\ m\epsilon\{1,2,\ldots,n\}} \frac{1}{n} \sum_{i=1}^{n} \|\mathbf{x_m} - \mathbf{x_i}\|_1. \tag{2.3.2}$$

DeMiguel et al. (2013) in their work, put forward a shrinkage estimator that incorporates the sample mean and adjust it towards a scaled vector of ones. Similarly, our work focuses on exploring shrinkage estimator designed for the L_1 - median. The shrinkage estimator based on L_1 - median is defined as:

$$\hat{\boldsymbol{\mu}}_{Sh(MM)} = (1-\eta)\hat{\boldsymbol{\mu}}_{MM} + \eta\nu_{\boldsymbol{\mu}}\mathbf{e}, \tag{2.3.3}$$

where $\nu_{\boldsymbol{\mu}}\mathbf{e}$ is the shrinkage target matrix, $\mathbf{e}$ is a vector of ones with p - dimension and $\hat{\boldsymbol{\mu}}_{MM}$ is the L_1-median from the samples. Scaling factor $\nu_{\boldsymbol{\mu}}$ and the shrinkage intensity η should be such that, they minimize the expected quadratic loss, i.e.,

$$\min_{\nu_{\boldsymbol{\mu}},\eta} \quad \mathrm{E}\left[\|\hat{\boldsymbol{\mu}}_{Sh(MM)} - \boldsymbol{\mu}\|_2^2\right], \text{ s.t. } \hat{\boldsymbol{\mu}}_{Sh(MM)} = (1-\eta)\hat{\boldsymbol{\mu}}_{MM} + \eta\nu_{\boldsymbol{\mu}}\mathbf{e}, \tag{2.3.4}$$

where $\|x\|_2^2 = \sum_{i=1}^{p} x_i^2$. Consider the function to be minimized in our problem:

$$\begin{aligned}
\mathrm{E}\left[\|\hat{\boldsymbol{\mu}}_{Sh(MM)} - \boldsymbol{\mu}\|_2^2\right] &= \mathrm{E}[\|(1-\eta)\hat{\boldsymbol{\mu}}_{MM} + \eta\nu_{\boldsymbol{\mu}}\mathbf{e} - \boldsymbol{\mu}\|_2^2] \\
&= (1-\eta)^2\mathrm{E}[\|\hat{\boldsymbol{\mu}}_{MM} - \boldsymbol{\mu}\|_2^2] + 2\mathrm{E}[\langle(1-\eta)(\hat{\boldsymbol{\mu}}_{MM} - \boldsymbol{\mu}), \eta(\nu_{\boldsymbol{\mu}}\mathbf{e} - \boldsymbol{\mu})\rangle] \\
&+ \eta^2\|\nu_{\boldsymbol{\mu}}\mathbf{e} - \boldsymbol{\mu}\|_2^2.
\end{aligned}$$

The asymptotic distribution for the L_1 - median was studied by Bose & Chaudhuri (1993), Bose (1995), and Möttönen et al. (2010). Let us consider a p - variate x random vector with CDF F, density function f and $p > 1$ satisfying the following assumptions:

- The density function of x is bounded and continuous
- x has a unique zero spatial median.

Möttönen et al. (2010) and Devlin et al. (1981) shown in their paper that under the above assumptions $\hat{\boldsymbol{\mu}}_{MM}$ approximately follows $\mathrm{N_p}(\boldsymbol{\mu}, \frac{1}{\mathrm{n}}\hat{\mathrm{A}}^{-1}\hat{\mathrm{B}}\hat{\mathrm{A}}^{-1})$ where $\hat{\mathrm{A}}(\mathbf{x_i}) = \frac{1}{\|\mathbf{x_i}\|_2}(\mathrm{I_p} - \frac{\mathbf{x_i}\mathbf{x_i^t}}{\|\mathbf{x_i}\|_2^2})$ and $\hat{B}(\mathbf{x_i}) = \frac{\mathbf{x_i}\mathbf{x_i^t}}{\|\mathbf{x_i}\|_2^2}$, with $\mathbf{x_i} \in \mathbf{R^p}$, for each $i = 1, 2, \ldots, n$. Thus second term in the RHS of (2.3.5) reduces to zero and minimization reduces to:

$$\mathrm{E}\left[\|\hat{\boldsymbol{\mu}}_{Sh(MM)} - \boldsymbol{\mu}\|_2^2\right] = (1-\eta)^2\mathrm{E}[\|\hat{\boldsymbol{\mu}}_{MM} - \boldsymbol{\mu}\|_2^2] + \eta^2\|\nu_{\boldsymbol{\mu}}\mathbf{e} - \boldsymbol{\mu}\|_2^2. \tag{2.3.5}$$

Our problem is to minimize the objective function in relation to $\nu_{\boldsymbol{\mu}}$ and η, where the parameter $\nu_{\boldsymbol{\mu}}$ appears only on the second term of the above equation. Thus, the associated first - order optimality criterion with respect to $\nu_{\boldsymbol{\mu}}$ be:

$$2p\nu_{\boldsymbol{\mu}} - 2\langle \mathbf{e}, \boldsymbol{\mu} \rangle = 0.$$

Therefore,

$$\hat{\nu}_{\boldsymbol{\mu}} = \frac{1}{p}\sum_{i=1}^{p} \boldsymbol{\mu}_i = \frac{\boldsymbol{\mu}\mathbf{e}}{p}.$$

Since we are shrinkaging $\boldsymbol{\mu}_{MM}$ here, we replace $\boldsymbol{\mu}$ by $\hat{\boldsymbol{\mu}}_{MM}$. The first-order condition (2.3.5) for optimality with respect to shrinkage intensity η is obtained as:

$$2(1-\eta)E[\|\hat{\boldsymbol{\mu}}_{MM} - \boldsymbol{\mu}\|_2^2] + 2\eta\|\nu_{\boldsymbol{\mu}}\mathbf{e} - \boldsymbol{\mu}\|_2^2 = 0,$$

$$\text{and therefore} \quad \hat{\eta} = \frac{\mathrm{E}[\|\hat{\boldsymbol{\mu}}_{MM} - \boldsymbol{\mu}\|_2^2]}{\mathrm{E}[\|\hat{\boldsymbol{\mu}}_{MM} - \nu_{\boldsymbol{\mu}}\mathbf{e}\|^2]}. \tag{2.3.6}$$

We have the approximate distribution of $\hat{\boldsymbol{\mu}}_{MM}$ as normal, above mentioned, and

$$\begin{aligned} \mathrm{E}[\|\hat{\boldsymbol{\mu}}_{MM} - \boldsymbol{\mu}\|_2^2] &= \mathrm{E}\Big[\sum_{i=1}^{p}(\hat{\boldsymbol{\mu}}_{MMi} - \boldsymbol{\mu}_i)^2\Big] \\ &= \sum_{i=1}^{p}\sigma^2_{\hat{\boldsymbol{\mu}}_{MMi}}. \end{aligned}$$

Thus the numerator of (2.3.6) be approximated by trace($\frac{1}{n}\hat{\mathrm{A}}^{-1}\hat{\mathrm{B}}\hat{\mathrm{A}}^{-1}$). Consequently, the shrinkage parameters are estimated as:

$$\hat{\nu}_{\boldsymbol{\mu}} = \frac{\hat{\boldsymbol{\mu}}_{MM}\mathbf{e}}{p} \text{ and } \hat{\eta} = \frac{\mathrm{trace}(\frac{1}{n}\hat{\mathrm{A}}^{-1}\hat{\mathrm{B}}\hat{\mathrm{A}}^{-1})}{\|\hat{\boldsymbol{\mu}}_{MM} - \hat{\nu}_{\boldsymbol{\mu}}\mathbf{e}\|^2}.$$

2.3.2 Shrinkage S_n (Sh-S_n) covariance estimator

The Median Absolute Deviation (MAD) is similar as the median, is a reliable estimate of dispersion for a random variable X. P. J. Rousseeuw & Croux (1993b) offer a more effective substitute for MAD with a 50% breakdown and have established S_n, a consistent and unbiased estimator for the relevant population functional.

$$S_n = c.\mathrm{med}_i\mathrm{med}_j|x_i - x_j|. \tag{2.3.7}$$

The constant ($c = 1.1926$) in the above equation is chosen as the consistency factor for normal distributions and to make the estimator unbiased too. Also, S_n is location - free that it does not use any location estimator. S_n estimator of scale assures bounded influence function and is more applicable due to its low gross error sensitivity. A comedian - based robust substitute for the sample covariance between two random variables X and Y was put out by Falk (1997), also, he has mentioned the idea of extending S_n scale estimator as a robust counterpart to covariance between two random variables, same is developed here in multivariate version. S_n covariance of two random variables X and Y be:

$$S_n(\mathrm{X}, \mathrm{Y}) = 1.4304.(\mathrm{med}_i[\mathrm{med}_{j \neq i}\{(x_i - x_j)(y_i - y_j)\}]).$$

Let $\mathbf{X}$ be $n \times p$ matrix with sample size n, number of variables p and $\mathbf{X_j}$ (j $= 1, 2, \ldots, p$) be the columns of the matrix. Then the covariance matrix of $\mathbf{X}$ based on S_n would be:

$$\hat{S_n} = S_n(\mathbf{X_i}, \mathbf{X_j}). \tag{2.3.8}$$

The trace of the above equation provides a robust scale estimator,

$$\mathrm{trace}(\hat{S_n}) = \sum_{j=1}^{p} S_n(\mathbf{X_j}, \mathbf{X_j}) = \sum_{j=1}^{p} 1.4304.\ S_n^2(\mathbf{X_j}) = \sum_{j=1}^{p} \sigma^2_{\mathbf{X_j}}. \tag{2.3.9}$$

Here $\hat{S_n}$ is an unbiased estimator and high breakdown estimator (Siegel (1982)), but it is non - positive definite. S_n estimator will remain bounded whenever more than $[(n\text{-}1)/2]$ points of n observations are confined fixed while the remaining points are arbitrarily moved. Let us consider, $\theta = (x_i - x_j)(y_i - y_j)$, be a class of function, taken over a set of n integers where more than half of the observations are bounded. S_n is nothing but median of median, in which, inner median of θ is calculated across n - 1 observations which would be bounded, since we are taking over a set of observations in which more than half observations are bounded. The outer median is taken over the inner medians. Since the inner median is bounded and we are taking a median, the outer median will be bounded. Thus, proposed S_n covariance estimator possess high breakdown, but it is non positive definite.

Instead of employing the conventional methods to ensure positive semi - definiteness of the covariance matrix, our approach involves shrinkage estimation over the empirical covariance matrix $\hat{S_n}$. This estimation leads to a well - conditioned, positive semi - definite matrix, serving as the shrinkage dispersion matrix defined below:

$$\hat{\Sigma}_{\text{Sh}} = (1-\eta)\hat{\text{E}} + \eta\hat{\text{T}}, \tag{2.3.10}$$

where $\hat{\text{E}} = \hat{S_n}$. Various options for the shrinkage target $\hat{\text{T}}$ have been proposed in the literature. For instance, Ledoit & Wolf (2003b) used weighted average of the sample covariance matrix and a single - index covariance matrix as the shrinkage target. Ledoit & Wolf (2003a) in their another work chose the shrinkage target as a "constant correlation matrix" with correlations are equal to the average of all sample correlations. Using a scaled multiple of the identity matrix as the shrinkage goal, as suggested by Ledoit & Wolf (2004), ensures a well - conditioned shrinkage covariance matrix even if the sample covariance matrix is not. DeMiguel et al. (2013) have introduced an alternative approach to estimating the covariance matrix and its inverse. According to their proposal, a shrinkage estimator can be constructed by taking a convex combination of the sample covariance matrix and a scaled shrinkage target. The same approach is executed for a sample covariance inverse too in their paper. DeMiguel consider the scaled identity matrix as target, same as that of Ledoit & Wolf (2004). In our procedure too we use shrinkage target $\hat{\text{T}} = \nu_{\Sigma}\text{I}$. Thus, equation 2.3.10 becomes:

$$\hat{\Sigma}_{\text{Sh}} = (1-\eta)\hat{\text{E}} + \eta\nu_{\Sigma}\text{I}, \tag{2.3.11}$$

where $\hat{\text{E}} = \hat{S_n}$. We need to estimate η and ν_{Σ}. The parameters are selected such that minimizing the expected quadratic loss:

$$i.e., \min_{\nu_{\Sigma},\eta} \text{E}\Big[\|\hat{\Sigma}_{\text{Sh}} - \Sigma\|^2\Big], \text{ s.t. } \hat{\Sigma}_{\text{Sh}} = (1-\eta)\hat{S_n} + \eta\nu_{\Sigma}\text{I}, \tag{2.3.12}$$

where$||A||^2 = \text{trace}(AA^T)/p$. Consider the above function to be minimized:

$$\begin{aligned}\mathrm{E}\Big[\|\hat{\Sigma}_{Sh}-\Sigma\|^2\Big] =& \mathrm{E}\Big[\|(1-\eta)\hat{S}_n+\eta\nu_\Sigma \mathrm{I}-\Sigma\|^2\Big]\\ =& \mathrm{E}\Big[\|(1-\eta)\hat{S}_n+\eta\nu_\Sigma \mathrm{I}-\Sigma+\eta\Sigma-\eta\Sigma\|^2\Big]\\ =& (1-\eta)^2\mathrm{E}\Big[\|\hat{S}_n-\Sigma\|^2\Big]+\eta^2\|\nu_\Sigma \mathrm{I}-\Sigma\|^2+\\ & 2\mathrm{E}\Big[\langle(1-\eta)(\hat{S}_n-\Sigma),\eta(\nu_\Sigma \mathrm{I}-\Sigma)\rangle\Big].\end{aligned}$$

Let our associated inner product is $\langle \mathrm{A}_1, \mathrm{A}_2\rangle = \mathrm{trace}(\mathrm{A}_1\mathrm{A}_2)^{\mathrm{t}}/p$. The latter element in the above expression is zero as $\mathrm{E}(\hat{S}_n) = \Sigma$ which is shown above. As a result, the above minimization expression reduces to:

$$\mathrm{E}\Big[\|\hat{\Sigma}_{Sh}-\Sigma\|^2\Big] = [(1-\eta)^2\mathrm{E}\Big[\|\hat{S}_n-\Sigma\|^2\Big]+\eta^2\|\nu_\Sigma \mathrm{I}-\Sigma\|^2. \tag{2.3.13}$$

Parameter ν_Σ presents only on the right side element in the above expression. Thus minimizing right element gives the optimum value of ν_Σ. Also, $\|\nu_\Sigma \mathrm{I}-\Sigma\|^2 = \nu_\Sigma^2\|\mathrm{I}\|^2+\|\Sigma\|^2-2\nu_\Sigma\langle \mathrm{I},\Sigma\rangle$. Thus the first - order minimization condition with respect to ν_Σ be:

$$\begin{aligned}2\nu_\Sigma-2\langle\Sigma,\mathrm{I}\rangle&=0,\\ \nu_\Sigma&=\mathrm{trace}(\Sigma)/p.\end{aligned} \tag{2.3.14}$$

Since Σ is unknown, we propose to estimate with $\hat{S}_n$, thus $\hat{\nu}_\Sigma = \mathrm{trace}(\hat{S}_n)/p$. The first - order optimal condition of equation 2.3.13 with respect to η be:

$$2(1-\eta)\mathrm{E}\Big[\|\hat{S}_n-\Sigma\|^2\Big]+2\eta\|\nu_\Sigma \mathrm{I}-\Sigma\|^2,$$

$$\text{and there by}\quad \hat{\eta} = \frac{\mathrm{E}\Big[\|\hat{S}_n-\Sigma\|^2\Big]}{\mathrm{E}\Big[\|\hat{S}_n-\Sigma\|^2\Big]+\|\nu_\Sigma \mathrm{I}-\Sigma\|^2} \tag{2.3.15}$$

where,

$$\|\hat{S}_n - \Sigma\|^2 = \left\langle \hat{S}_n, \hat{S}_n \right\rangle - 2\left\langle \hat{S}_n, \Sigma \right\rangle + \langle \Sigma, \Sigma \rangle\,,$$

$$\text{and} \quad \mathrm{E}\left[\|\hat{S}_n - \Sigma\|^2\right] = \mathrm{E}\|\hat{S}_n\|^2 - \|\Sigma\|^2. \tag{2.3.16}$$

The right element in denominator of η in (2.3.15) can be expressed as:

$$\|\nu_\Sigma \mathrm{I} - \Sigma\|^2 = \nu_\Sigma^2 I - 2\langle \nu_\Sigma I, \Sigma \rangle + \|\Sigma\|^2. \tag{2.3.17}$$

From 2.3.16 and 2.3.17 it is clear that denominator of η in (2.3.15) is $\mathrm{E}\|\hat{S}_n - \nu_\Sigma I\|^2$.

$$\hat{\eta} = \frac{\mathrm{E}\left[\|\hat{S}_n - \Sigma\|^2\right]}{\mathrm{E}\left[\|\hat{S}_n - \nu_\Sigma \mathrm{I}\|^2\right]}. \tag{2.3.18}$$

Substituting $\hat{\eta}$ and $\hat{\nu_\Sigma}$ in 2.3.11 gives Shrinkage S_n (Sh-S_n) estimator.

A robust squared Mahalanobis distance (RMD) could be defined based on above - defined location 2.3.3 and covariance estimates 2.3.11 as follows:

$$\mathrm{RMD}(x_i, \hat{\boldsymbol{\mu}}_{Sh(MM)}, \hat{\mathrm{E}}_{\mathrm{Sh}}) = \mathrm{RMD}_i = (x_i - \hat{\boldsymbol{\mu}}_{Sh(MM)})^t \hat{\mathrm{E}}_{\mathrm{Sh}}^{-1} (x_i - \hat{\boldsymbol{\mu}}_{Sh(MM)}).$$

At this point, the distances of n observations are compared to each other, either through graphical evaluation, a predefined criterion, or a fixed threshold, to identify which observations are considered outliers. Several fixed thresholds have been proposed in literature; however, the simplest approach involves calculating the critical value from the χ_p^2-distribution with p degrees of freedom, [Gnanadesikan & Kettenring (1972)] based on a given significance level, α, usual consideration of α is 0.025. Here it is important to remember that chi-square is just an approximate distribution of squared Mahalanobis distance, not necessary, when the data is from non - normal distributed sample. Thus, finding exact cutoff values for identifying outlying distances is a challenging problem that has received a lot of attention, as no universally applicable method exists. $\chi^2_{p,0.975}$ quantile is often used as the threshold for detecting outliers in robust distance cases too, but this method may have some limitations. Reimann et al. (2005) in his paper mentioned that cutoff should be adapted with respect to sample size. R. A. Maronna & Zamar (2002) used a cutoff other than the usual chi-square quantile and determined that suits the RMD better than chi-square quantile. Here we use the cutoff same as R. A. Maronna & Zamar (2002) for outlier detection: $cv = \frac{\chi^2_{0.95,p} \times \text{median}(\text{RMD}_i)}{\chi^2_{0.5,p}}$. Any observation for which $\text{RMD}_i > cv$ is considered as an outlier. Sensitivity analysis of the chosen cutoff has done in the upcoming section 2.3.5. Using this cutoff value and the robust Mahalanobis distance, a weight function can be defined to derive robust estimates for location and scatter. These estimates are positive definite and approximately affine equivariant.

2.3.3 Sensitivity curve of Sh-S_n estimator

The influence function (IF) of an estimator, as introduced by Hampel (1974), is the asymptotic counterpart of the sensitivity curve. It provides an approximation of the behavior of the estimator when the sample includes a small proportion of identical outliers. The influence function (IF) can be viewed as a "limit case" of the sensitivity curve in the following context: When a new observation x_0 is added to the sample $x_1, \ldots, x_n$ the contamination fraction becomes $1/n + 1$. Consequently, the standardized sensitivity curve (SC) is defined as:

$$\text{SC}_n(x_0) = (n+1)\left(\hat{\theta}_{n+1}(x_1, \ldots, x_n, x_0) - \hat{\theta}_n(x_1, \ldots, x_n)\right).$$

It can be expected that if the x_i's are *i.i.d.* with distribution F , then for large n, the standardized sensitivity curve $\text{SC}_n(x_0)$ will approximate the influence function $IF(x_0, F)$. In order to check the sensitivity of our proposed covariance estimator, we consider the metric Frobenius norm of $\left(\hat{\theta}_{n+1}(x_1, \ldots, x_n, x_0) - \hat{\theta}_n(x_1, \ldots, x_n)\right)$ with respect to proposed Sh-S_n estimator. And then we plot the values. We have considered $n = 100$ and 500. Sample observations are taken from multivariate normal with mean zero vector and dispersion identity matrix. Observation x_0 is taken from multivariate normal $\text{N}(\lambda \times \mathbf{1}_p, 0.01 \times \mathbf{I}_{p \times p})$. Here λ take values (-100, -95, -90,. . . , 0 ,. . . ,90, 95, 100). The plots below, 2.1, 2.2, shows classical covariance and Sh-S_n estimator performance. From the graphs, it is very clear that our proposed estimator remains bounded irrespective of change in λ or n. The plot of classical covariance shows us, how an addition of observation influence the estimator. Sh-

S_n estimator sensitivity curve remains bounded throughout all samples considered. And this gave us assurance on empirical finite influence function property of our proposed estimator.

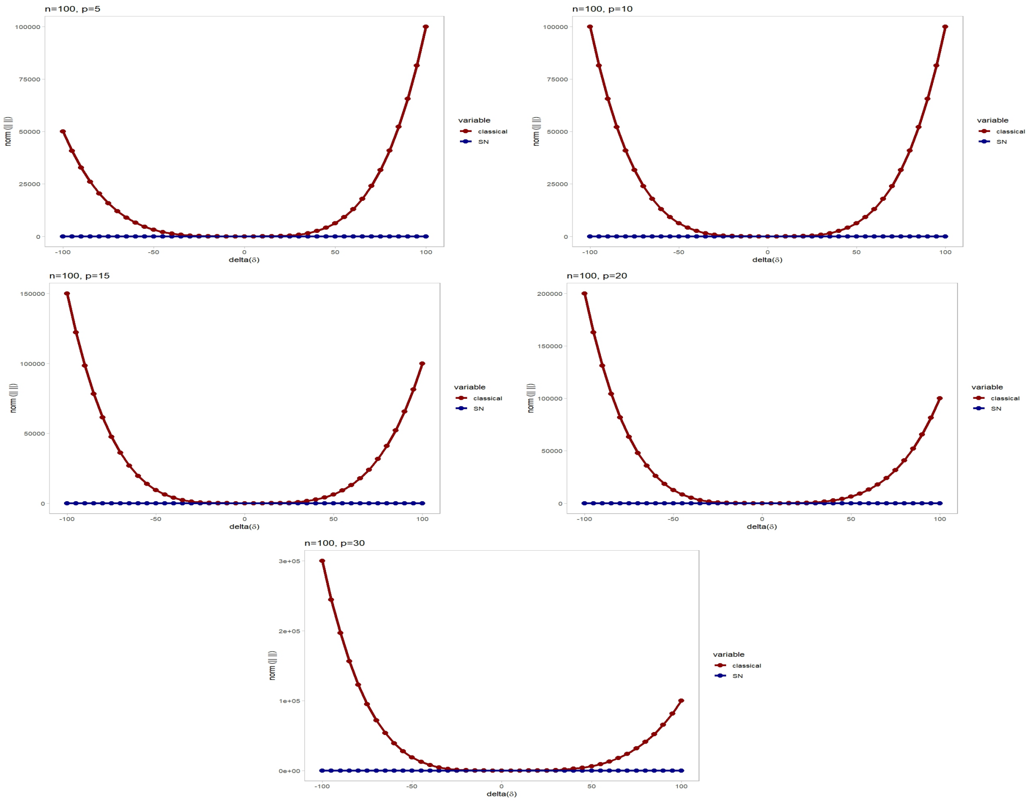


Figure 2.1: Sensitivity curve of Sh-S_n estimator and classical estimator for $n = 100$

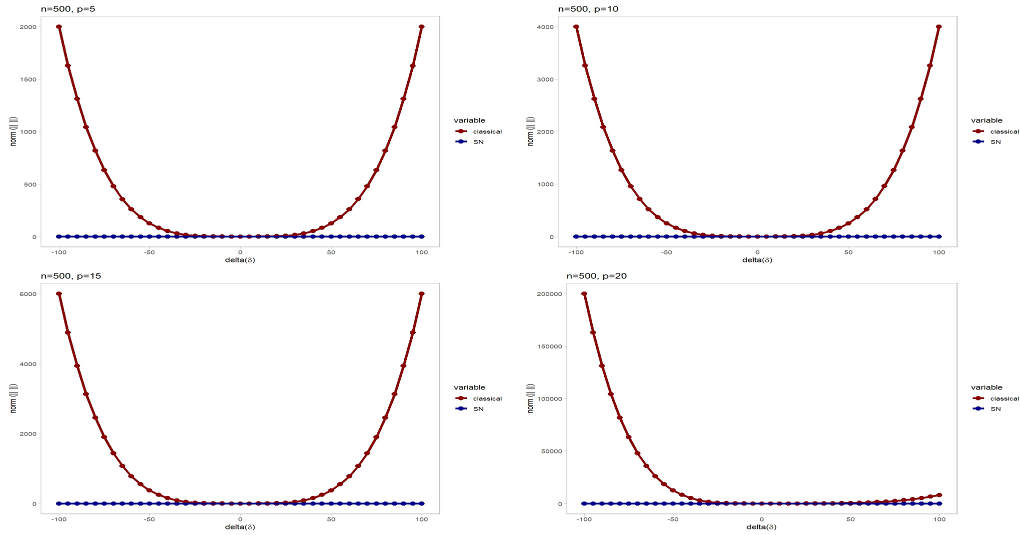


Figure 2.2: Sensitivity curve of Sh-S_n estimator and classical estimator for $n = 500$

To check the efficiency of Sh-S_n estimator along with other robust estimators in non contaminated data, we generated data from multivariate normal distribution with mean zero vector and covariance identity matrix. 1000 times data generated and each time trace of our proposed robust shrinkage estimator along with classical estimator identified. We found the trace in each replication of the data. And finally found median of ratio of the trace as our efficiency metric.

$$\text{Eff} = \text{median}\left(\frac{\text{trace}(\textit{Robust method})}{\text{trace}(\textit{Classical method})}\right)$$

Sample size considered for this study are n = 20, 30, 50, 100 and dimensions considered are p = 5, 10, 15, 20. Except for sample 20, we have considered all aforementioned dimensions, because MCD estimator function in MASS package used, demands sample size strictly greater than dimension. The results are tabulated and given below in table 2.1. The results shows that our proposed covariance estimator works almost similar to that of classical covariance in terms of variability of the estimate. In general, trace of a matrix can explain the possible total variance of the matrix. The ratio of trace could indicate the efficiency associated with the matrices. Ratio value near one indicates both the variance of the matrices in comparison are somewhat near. Apart from $Sh\text{-}S_n$ estimator, we have considered MCD, Comedian and S_n estimators. The tabulated results assure the efficiency of $Sh\text{-}S_n$ estimator.

2.3.4 Simulation Study

A random variable X is considered in the study, which follows a contaminated multivariate normal distribution. The distribution of X is mixture of normal, where the majority component is represented by $(1 - \alpha)$ times the standard multivariate normal distribution with mean 0 and identity covariance matrix $(\mathrm{N}(0, \mathrm{I}))$, and the contaminated component is represented by α times a normal distribution with mean δe and covariance matrix λ times the identity matrix $(\mathrm{N}(\delta e, \lambda\mathrm{I}))$. Here, e denotes a p - dimensional vector of ones. The model described is similar to the one employed by several authors, including EStimator (1999), Peña & Prieto (2001), R. A. Maronna & Zamar (2002), Sajesh & Srinivasan (2012), and Kunjunni & Abraham (2022). The experiment was carried out across various dimensions of the sample space, specifically p = 5, 10, 15, 20, 30 and 50 and the sample sizes chosen were n = 100, 500 and 1000. The experiment involved varying contamination levels

Table 2.1: Table values on efficiency of Sh-S_n estimator along with others

n	p	Sh-S_n	MCD	Comedian	S_n
20	5	0.9553157	0.7495664	0.99995	0.8397958
	10	0.9903444	0.9844151	0.959030	0.7977359
	15	0.9581188	0.9920467	0.98887	0.8543333
30	5	0.9516421	0.9326292	0.968482	0.7427073
	10	0.9276487	0.8064676	1	0.7714724
	15	0.9965764	0.9045927	0.956633	0.7183989
	20	0.9918758	0.9684	0.99876	0.8168414
50	5	0.9745917	0.8322049	0.99601	0.7315996
	10	0.9687162	0.7823275	0.99965	0.725625
	15	0.9564507	0.9249257	0.97765	0.7341021
	20	0.9768504	0.9344656	0.96654	0.7842835
100	5	0.9643644	0.8269753	0.98109	0.6684087
	10	0.9856587	0.8871585	0.9760979	0.7096787
	15	0.9704157	0.9131435	0.9931914	0.7622983
	20	0.9823935	0.9186862	0.9884738	0.7334713

represented by $\alpha = 0, 0.1, 0.2$, and 0.3. The outliers were positioned at distances $\delta = 5$ and 10, and the concentration of contamination was set at $\lambda = 0.01$, 0.1 and 1. 1000 random samples were generated and repeated for each combination of these values. In our simulation study to evaluate the effectiveness of our distance measure, we incorporated not only the multivariate normal distribution but also the contaminated multivariate t_3 and the multivariate exponential distribution. The contamination scheme and sample sizes were same across all three distributions.

For the evaluation purpose, specific measures were examined, namely the true positive rate (TPR) and the false positive rate (FPR). Let TO represent the actual number of outliers and NO denote the actual number of non - outlying observations. Let FP (False Positives) be the observations incorrectly diagnosed as outliers by the respective method. Let TP (True Positives) be the observations correctly diagnosed by the respective method as outliers. Then TPR and FPR can be defined as:

$$\text{TPR} = \frac{TP}{TO} \quad \text{and} \quad \text{FPR} = \frac{FP}{NO}$$

TPR reflects the model's effectiveness in accurately detecting true outliers. A higher TPR signifies a stronger capability to identify actual outliers. FPR indicates the percentage of normal data points that are mistakenly classified as outliers. A lower FPR is preferable to reduce the occurrence of false alarms. Both metrics are crucial in assessing a model's performance, and they are often analyzed together. The challenge lies in balancing TPR and FPR to optimize the model's effectiveness for the specific application. Throughout our study we make use of TPR and FPR as measures to compare. In order to evaluate our proposed distance measure, we compare it with several existing non reweighted covariance - based robust Mahalanobis distance (RMD), including FAST - MCD covariance proposed by EStimator (1999), RMD - based on OGK covariance estimator introduced by R. A. Maronna & Zamar (2002), RMD - based on S_n covariance estimator introduced by Kunjunni & Abraham (2022) and RMD - based on shrinkage comedian (RMDS) covariance estimator proposed by Cabana et al. (2021). All the simulations are done in R software. MCD and OGK estimators are available in different packages in R. Tables 2.2, S1 in Appendix A shows the FPR results on 0% contamination in multivariate normal distribution. It is apparent from the tables our proposed method has minimal FPR in 0% contamination for all sample sizes considered and provides an assurance of using the proposed method in usual cases too. Tables S2 - S11 in Appendix A shows the results of TPR and FPR in different contamination for $n = 1000$, 500, 100. For $n = 1000$, instead of table, we have plotted the FPR A.1 and exhibited in supplementary material. In our simulation scenarios, we considered a range of λ and δ values from small to large. When δ and λ are large, the contamination becomes quite evident, and in such cases, almost all robust methods successfully identify the outliers. Taking this into consideration, we considered small deviations of δ, λ. For $\delta = 10$, $\lambda = 1$ and for all sample size considered, Our proposed distance measure possess 100% TPR in outlier detection along with RMD - based OGK, and S_n. However, when we consider the results across varying sample sizes, our distance measure achieves a 100% TPR for almost all dimensions. Additionally, we observe that RMD - based S_n and OGK perform best with a 100% TPR for larger values of δ, λ, but their TPR degrades for smaller deviations of δ, λ and higher contamination. Especially with 30% contamination RMD based OGK is attaining zero TPR for small deviated δ, λ irrespective of the sample size. RMD based S_n shows improved TPR, but not better than our distance measure, irrespective of the parameters. Consequently, our proposed distance measure has an advantage over the other methods when dealing with higher contamination and small δ, λ, as considered in the study. Table S12 in Appendix A shows the FPR results under zero contaminated multivariate t_3 distribution. The results show that all RMD exhibits comparatively high FPR for multivariate t_3 distribution. Among the high FPR scenarios, our proposed method exhibits the smallest FPR along with RMD S_n than others. Tables S13 - S22 in Appendix A shows the results of respective contaminated distribution FPR and TPR for n = 1000, 500, 100. For n = 1000, FPR is plotted and exhibited in the supplementary material A.2. The results show that our proposed method shows 100% TPR for higher as well as lower contamination for any dimension and sample sizes with large δ, λ considered in the study along with the RMD - based S_n and OGK method. But for small deviated δ, λ included sample, irrespective of sample size, RMD - based OGK and S_n fails to possess a better TPR, especially for higher contamination, but our proposed measure outperforms here. Furthermore, our proposed measure along with RMDS exhibit low FPR compared to the other methods. In terms of FPR, our proposed method and RMDS shows better performance. But RMDS miserably fails to perform in terms of TPR in majority of cases considered. Also, we can see all method shows decreasing FPR for increasing levels of contamination in all dimensions and sample sizes. Thus, our proposed distance measure has an advantage in terms of TPR for samples with higher contamination and small outliers, over other methods in samples of multivariate t_3. Table S23 in Appendix A shows FPR of multivariate exponential distribution under zero contamination level. Respective results shows that even for multivariate exponential distribution our introduced method exhibits the minimal FPR than all others in a non contaminated data. Also, for higher dimensions we can see 0% FPR for our distance measure which asserts it could be used in normal

Table 2.2: Zero contaminated multivariate normal distribution FPR table for n = 1000

p	$Sh\text{-}S_n$	RMDS	S_n	MCD	OGK
5	0.0219	0.0272	0.1689	0.03201	0.0909
10	0.0168	0.0267	0.0005	0.03691	0.10561
15	0.0156	0.0351	0.0519	0.03712	0.1029
20	0.0119	0.0299	0.0501	0.04389	0.10363
30	0.0099	0.0377	0.0591	0.05176	0.10161
50	0.0051	0.0299	0.0615	0.07697	0.10398

situations too. To check the performance in respective contaminated distribution, the contamination is defined as a mixture of $(1-\alpha)$ times the standard multivariate exponential distribution with a mean zero,unit covariance and α times the multivariate exponential distribution with mean δe, λI. Tables S24 - S33 in Appendix A show the results of the Multivariate Exponential distribution. And we have plotted and presented the FPR for $n = 1000$ A.3. Similar to multivariate normal and t_3 distribution, FPR and TPR are compared here for $n = 1000$, 500, 100. All the methods generally possess a low FPR and high TPR for this particular distribution than other two considered in the study. Our proposed distance measure, RMD - based S_n attains 100% TPR for largely deviated outlier included sample. And for smaller δ, λ, only our proposed measure attains higher TPR, which is an advantage of our method. FPR of our proposed method and RMDS is lower than others. But RMDS terribly fails to perform in terms of TPR and even attain 0% TPR for higher contamination in many situations. And we can see a decreasing FPR as contamination percentage increases across the dimensions for all methods. The findings demonstrate the effectiveness of our proposed technique in terms of TPR across all assessed scenarios, especially in higher level of contamination.

For smaller samples in non-multivariate normal, our proposed measure shows an advantage over other methods. And for multivariate normal distribution our proposed method posses an advantage of having high TPR than others, even for 30% contamination level in larger samples. Even though our method excels in achieving a higher TPR, it demonstrates a slightly lesser degree of effectiveness in controlling the FPR when compared to the RMDS, albeit by a negligible margin. In all examined scenarios both, the proposed technique and the RMDS method exhibit a low level of FPR. Furthermore, for the case of 30% contamination, both methods demonstrate 0% FPR in many situations. However, in terms of TPR, our technique is better than RMDS irrespective of all parameters considered in the study. In terms of TPR RMD - based S_n and OGK exhibit a better TPR similar to our method in many scenarios. But if we closely observe, RMD based S_n and OGK fails to perform in situations like higher contamination with small deviated outliers irrespective of sample size. Also, FPR of S_n is always high compared to our proposed distance measure. Our proposed method performs well and possesses advantages over them in certain scenarios. Even when dealing with non - normal data, our proposed measure is performing similarly to that of multivariate normal data.

2.3.5 Sensitivity analysis

Usually classical quantile $\chi^2_{0.975,p}$ will be used as cutoff for outlier detection in distance measure. But it has its own disadvantages. One of the major disadvantages is it has the tendency to detect more non-outlying observations as outliers R. A. Maronna & Yohai (1995). In real-life application we can see this disadvantage in distance measures using chi - square quantile as cutoff. Also, the difficulty in setting this threshold value in the case of robust distance is highly subjective because there is no proof of the real distribution of the squared robust Mahalanobis distance. We have used the same cutoff as that of R. A. Maronna & Zamar (2002). In this section we have conducted the sensitivity analysis with respect to the parameters $\delta = (1, 2, 3,\ldots,10)$, $\lambda = (0.01, 0.1, 1)$ and $\alpha = (10\%, 20\%, 30\%)$ to check the performance of cutoff. We have included only the graphs with respect to $\alpha = 30\%$ below. Figures 2.3 - 2.7 show after a small δ and λ deviations, our proposed measure attains 100% TPR with respect to our cutoff even for 30% contamination. The result from graphs acknowledge utilizing chosen cutoff for our distance measure is suitable for outlier detection.

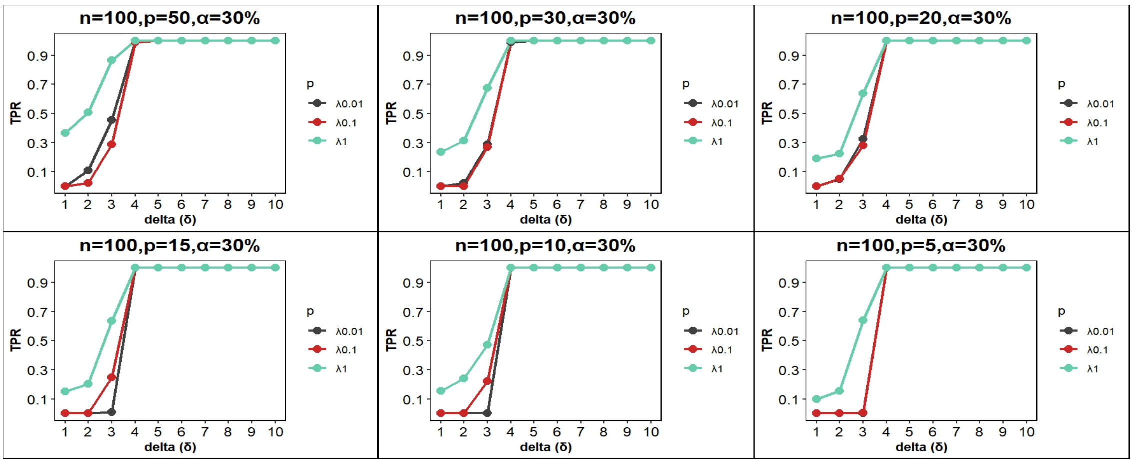


Figure 2.3: TPR across different dimension for $n = 100$, $\delta = (1, 2, \ldots, 10)$ and $\alpha = 30\%$

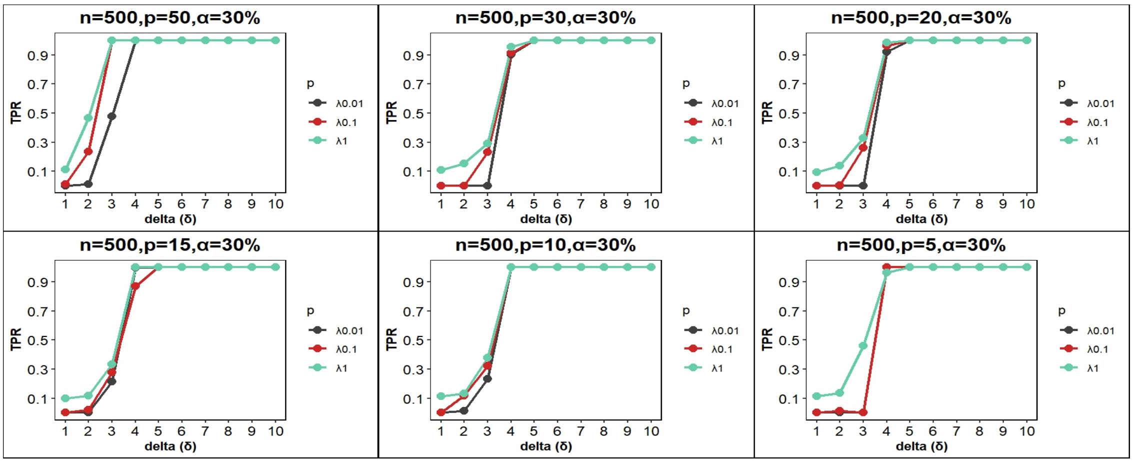


Figure 2.4: TPR across different dimension for $n = 500$, $\delta = (1, 2, \ldots, 10)$ and $\alpha = 30\%$

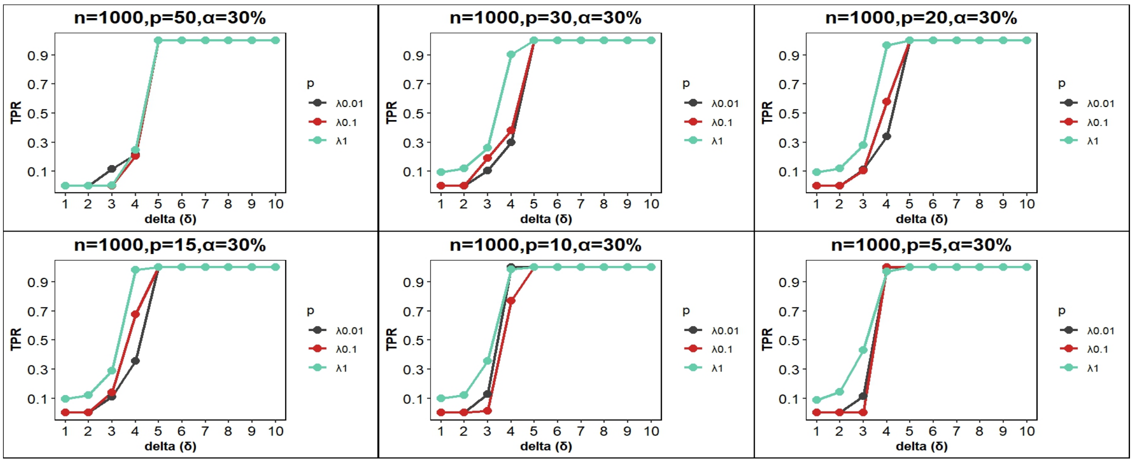


Figure 2.5: TPR across different dimension for $n = 1000$, $\delta = (1, 2, \ldots, 10)$ and $\alpha = 30\%$

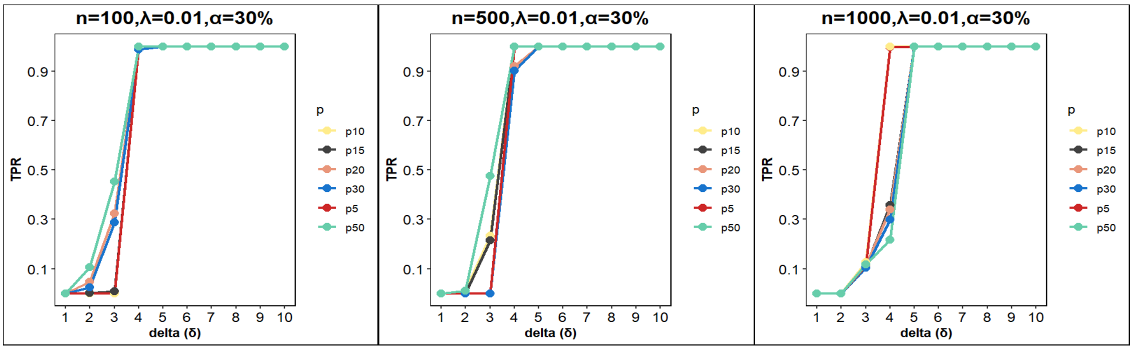


Figure 2.6: TPR in relation to varying sample size for $\lambda = 0.01$, $\delta = (1, 2, \ldots, 10)$ and $\alpha = 30\%$

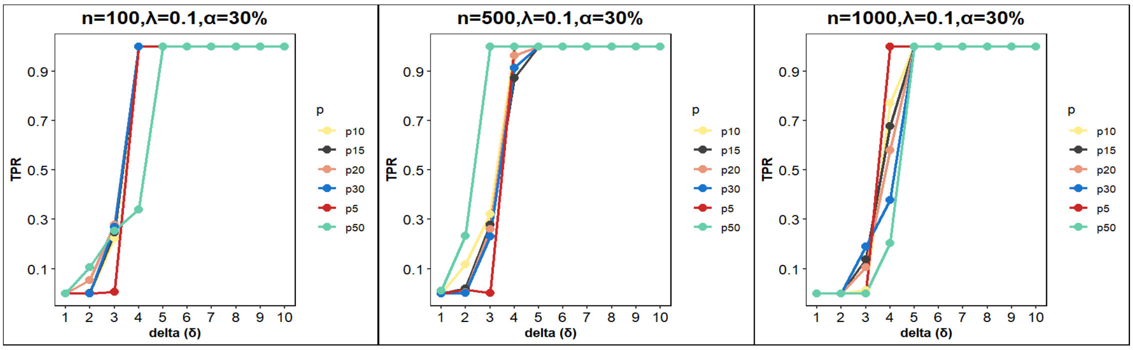


Figure 2.7: TPR in relation to varying sample size for $\lambda = 0.1$, $\delta = (1, 2, \ldots, 10)$ and $\alpha = 30\%$

2.3.6 Affine Equivariance property

In this section, we will be examining the equivariance property of proposed estimator through a simulation. Consider a $n \times p$, X matrix where $\mathrm{x}_i^t \; \epsilon \; \mathrm{R}^\mathrm{p}$ $(i = 1, 2, 3, \ldots, n)$ be the rows of the matrix and a pair of multivariate center and dispersion (c, V). Let $X_A = A\mathrm{x}_1, \ldots, A\mathrm{x}_\mathrm{n}$ where A be a $p \times p$ non-singular matrix. Then the location and scatter estimates are said to be affine equivariant if,

$$\mathrm{c}_A = \mathrm{c}(\mathrm{X}_A) = A\mathrm{c}(\mathrm{X}) \text{ and } \mathrm{V}_A = \mathrm{V}(\mathrm{X}_A) = A\mathrm{V}(\mathrm{X})A^t.$$

Affine equivariance of a location and scatter matrix is identical to the corresponding robust Mahalanobis distance's affine equivariance. Thus,

$$\mathrm{RMD}(A\mathrm{x}_\mathrm{i}, \mathrm{c}_A) = (A\mathrm{x}_\mathrm{i} - \mathrm{c}_A)\mathrm{V}_A^{-1}(A\mathrm{x}_\mathrm{i} - \mathrm{c}_A)^t.$$

The proposed shrinkage RMD is ultimately based on not affine equivariant estimators which are the L_1-median (Lopuhaä & Rousseeuw (1991)) and the $Sh - S_n$ covariance estimates. L_1 median is orthogonal equivariant with any orthogonal matrix A and it implies that the L_1-median transforms appropriately under all transformations that preserve euclidean distances. S_n matrix proposed is location invariant, symmetric and scale invariant (Falk (1997)). R. A. Maronna & Zamar (2002) probed the effect of lack of equivariance of OGK estimates with transformed data through simulations. Similar to that method, here we generate random matrix A = TD where T is a random orthogonal matrix and D is a diagonal matrix with diagonal entries are independent observations generated from $\mathrm{U}(0, 1)$. We tested the performance of our proposed method by examining its ability to detect outliers in transformed data. To do this, we first generated a data matrix X, which was then affine-transformed by multiplying it with a previously defined random matrix A. Next, we applied the proposed distance measure to the transformed data matrix X to identify any outliers. The contaminated data scheme for affine equivariance consists of data generated from a mixture of normal $100(1-\alpha)\mathrm{N}(0, \mathrm{I}) + \alpha\mathrm{N}(\delta e, \lambda\mathrm{I})$. The dimension $p = 5$, 10, 15, 20, 30 with contamination level $\alpha = 10$, 20, 30% is considered here. The distance of outliers $\delta = 5$, 10 and the contamination concentration $\lambda = 0.01$, 0.1, 1 considered here. Based on the study results presented in tables S34 of supplementary material gives the evidence our proposed method is capable of detecting almost all the outliers even after undergoing transformation. As the sample size increased the method showed improved performance with an increase in dimension. We have even checked using reweighted covariance estimator-based RMD too. The results are not shown here, reweightedness has improved the equivariance of the RMD measure. Thus we have shown empirically our measure is affine under transformations. Investigating the behavior of the proposed measure under correlated data is crucial as it is not affine equivariant. In their study, Devlin et al. (1981) utilized a correlation matrix P to generate monte carlo data from various distributions with a moderate dimension of $p = 6$. The matrix has the form:

$$P = \begin{bmatrix} P_1 & 0 \\ 0 & P_2 \end{bmatrix} \text{ with } P_1 = \begin{bmatrix} 1 & 0.95 & 0.30 \\ 0.95 & 1 & 0.10 \\ 0.30 & 0.10 & 1 \end{bmatrix} \text{ and } P_2 = \begin{bmatrix} 1 & -0.499 & -0.499 \\ -0.499 & 1 & -0.499 \\ -0.499 & -0.499 & 1 \end{bmatrix}$$

The matrix P was chosen because of its large dimension, which makes it suitable for studying multivariate estimators, and because of the wide range of correlation values it offers. This enables to examine, difference in the abilities of various distance measure to detect outliers from highly correlated data. Dataset of $n = 100$, generated from asymmetric contaminated normal mixture distribution such that : $(1-\alpha)$ $\mathrm{N}(0, P) + \alpha\mathrm{N}(5e, P)$. The contamination levels considered are: $\alpha = 10\%$, 20%, 30%. Table 2.3 shows that our proposed distance measure possesses 100% TPR and low false detection rate even in correlated data with high contamination.

Table 2.3: Table showing results on correlated data

	Sh-S_n		FAST MCD		OGK		RMDS		S_n	
α	TPR	FPR	TPR	FPR	TPR	FPR	TPR	FPR	TPR	FPR
0.1	1	0.0367	1	0.0226	1	0.0736	1	0.0158	1	0.0088
0.2	1	0.0171	0.8565	0.0062	0.9792	0.0533	1	0.0033	1	0
0.3	1	0.0003	0.1238	0.0614	0.478	0.046	0.9973	0.0001	0.9103	0.0004

2.3.7 Breakdown property

The breakdown value of an estimator refers to the maximum percentage of outliers that can be present in a dataset without causing the estimator to provide a false estimate. Essentially, this means that the estimator can handle a certain amount of modified data before its accuracy becomes compromised. Here we have studied the breakdown property of our proposed distance measure empirically. To find the breakdown value of the proposed measure we generate symmetric and asymmetric contaminated distribution. First we generate n observations from $\mathrm{N}_p(0, \mathrm{I})$. For symmetric contamination the i^{th} observation multiplied by $100i$ and for asymmetric contamination the i^{th} observation multiplied by $(100i)\mathbf{e}$, $i = 1, \ldots, 100\alpha$, where $\mathbf{e} = (1, \ldots, 1)^t$. Here we consider p = 10, 30, 50, 80, 100, n = 100, 500, 1000 and α = 10, 20, 30, 40, 45 for the study to identify empirical breakdown. The results are displayed in S35 table of Appendix A. The results shows that our proposed measure exhibit a stable performance of 100% TPR with extremely low FPR for nearly every combination of samples and contamination level considered. Notably, our proposed measure stands out as it consistently maintains a perfect TPR across all combinations of contamination levels and sample sizes. Additionally, it achieves a flawless zero FPR for higher contamination scenarios. The highest observed false detection rate, both in symmetric and asymmetric cases, is merely 0.0101, occurring with dimension five and 10% contamination level. These results offer empirical evidence that our proposed measure excels in detecting outliers at high contamination levels with remarkable accuracy.

2.4 Real - Life application of distance measure

2.4.1 Bushfire dataset

This section explains the efficacy of the suggested method in real-world datasets Campbell (1989) acquired bushfire data for exploring real data applications, which consisted of satellite measurements on five frequency bands, each corresponding to 38 pixels. R. A. Maronna & Yohai (1995) examined the data and determined that observations 7-11 are outliers that may be easily recognized using several robust methods.The observations 32-38 are masked by the initial outliers and Stahel Donoho projection estimator by R. A. Maronna & Yohai (1995) determined those observations too without any fail. From figure 2.8 we can see that RMD - based on Sh-S_n detected the outliers exactly. RMD - based MCD, OGK and S_n detect the outliers and in addition, they identify more nearby observations as outliers.

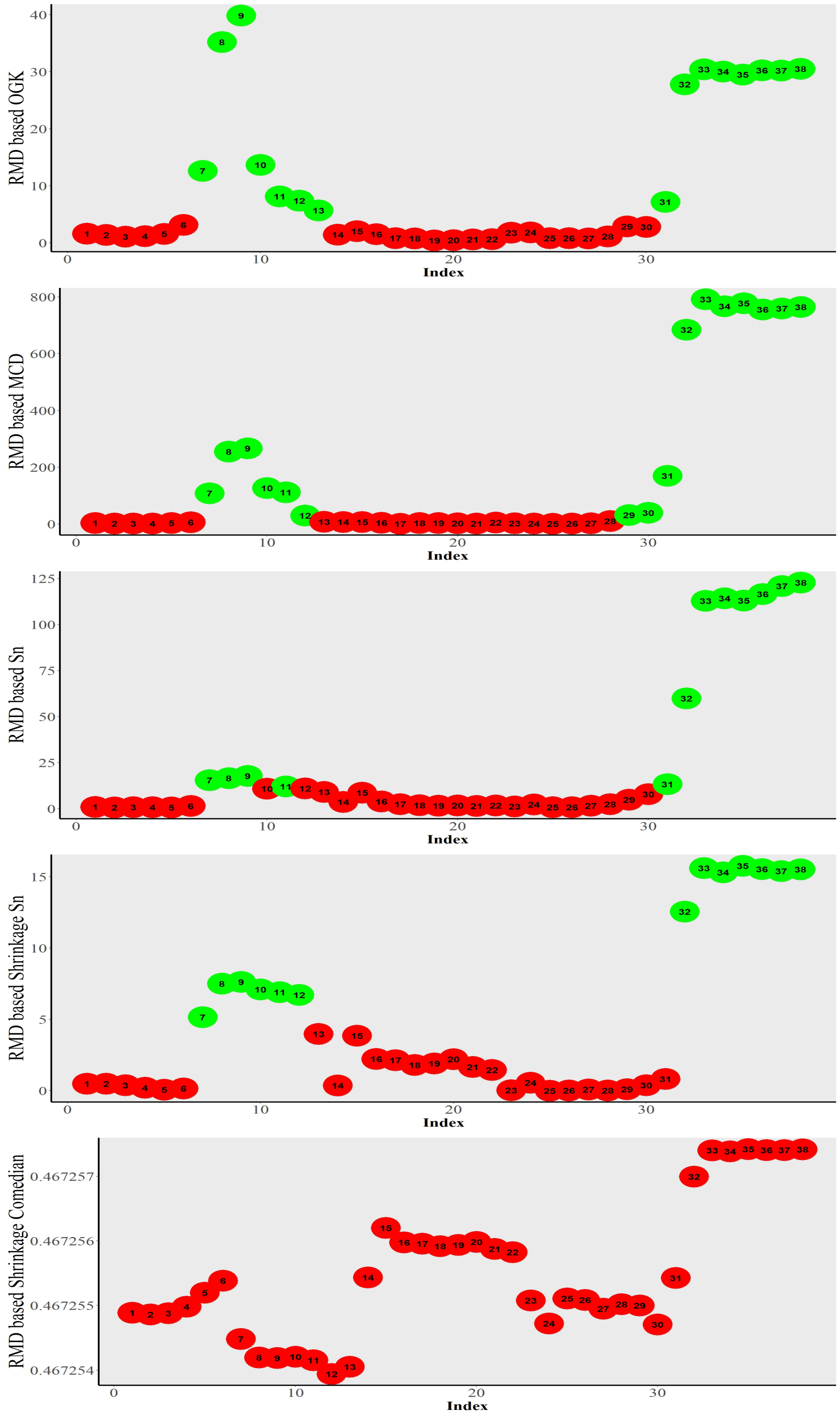


Figure 2.8: Bushfire data Index Plot of Robust Mahalanobis distance based on OGK, MCD, S_n, Sh -S_n, RMDS

2.4.2 Milk dataset

Daudin et al. (1988) provided data readings on the composition of 86 containers of milk, contrary to the misrepresented claim in their data description that it was based on 85 containers. A. C. Atkinson (1994) and Rocke & Woodruff (1996) analyzed the same dataset. Atkinson recommended excluding observation 70 from the analysis due to anomalous findings observed during the preliminary analysis. Additionally, observations 63 and 64 are identical duplicates (which could potentially account for the inconsistency between the 86 observations listed in Table 1 of Daudin et al. and the stated 85 observations contained within it). The removal of the duplicate point 64 and observation 70 from the analysis necessitates a renumbering of the subsequent observations after 63 which we have done and reflected in the graph present too. But the comparison of our results with A. C. Atkinson (1994) and Rocke & Woodruff (1996) is done in terms of the original number of the observations. Both forward and hybrid algorithms identified observations 1, 2, 3, 12, 13, 14, 15, 41, 44 and 74(72^{nd} point in graph) as the most extreme outliers, with additional observations 11, 17, 27 and 77(75^{th} point in graph) flagged as potential outliers. Our proposed measure identifies all these outliers and the corresponding index plots given in figure 2.9 . RMD - based OGK and MCD methods also identify majority of these outliers.

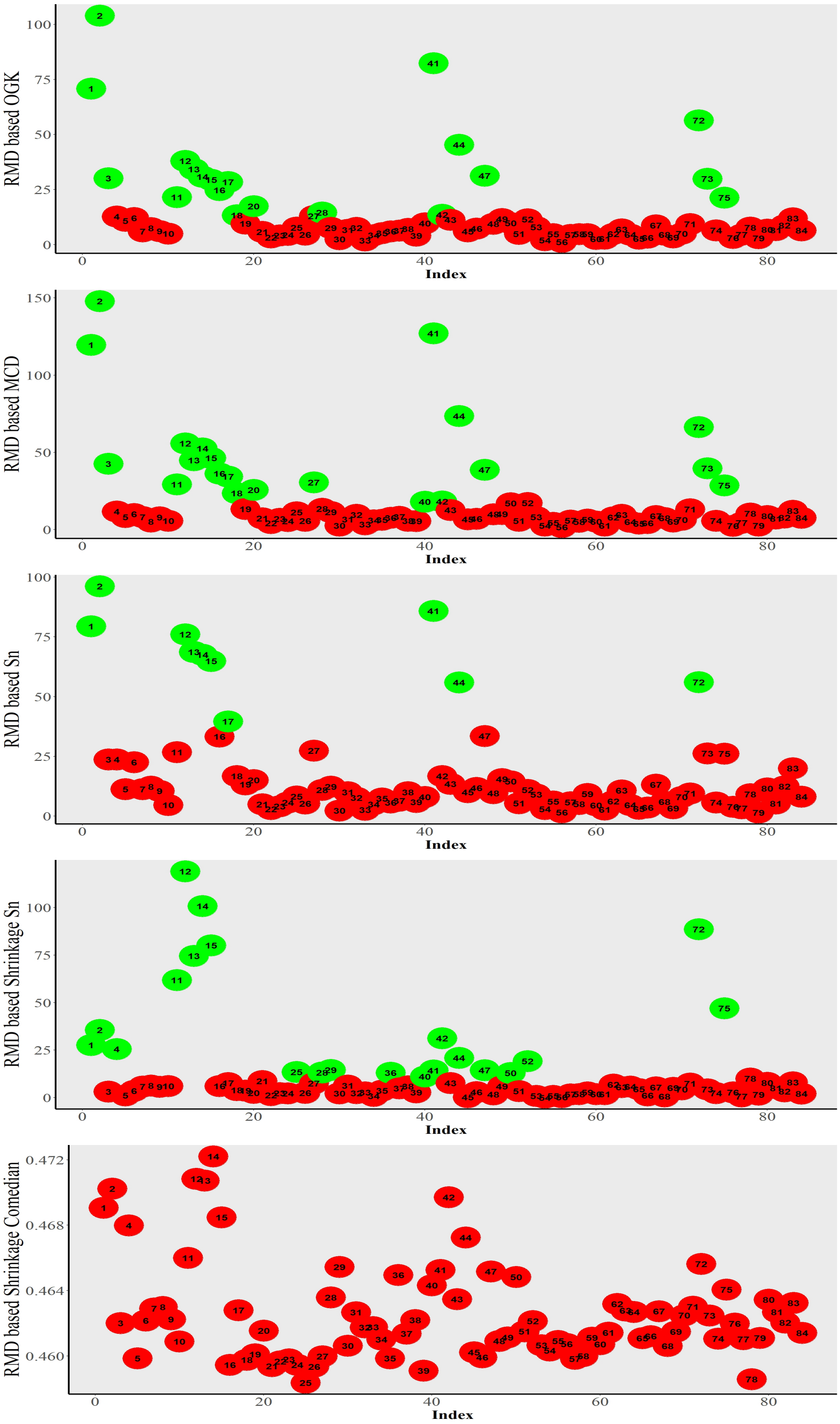


Figure 2.9: Milk data Index Plot of Robust Mahalanobis distance based on OGK, MCD, S_n, Sh - S_n, RMDS

2.4.3 Stackloss dataset

The Stackloss dataset (Brownlee (1965)) comprises 21 observations regarding the percentage of unconverted ammonia escaping from an oxidation plant. It was provided by P. J. Rousseeuw & Van Zomeren (1990), a detailed analysis of the dataset and identified the observations 1, 2, 3 and 21 as outliers. He also mentioned that observation 17 is also a far away point in the data. Figure 2.10 shows the index plot for Stackloss data. RMD based on Sh-S_n identifies 1, 2, 3 and 17 as outliers. RMD based MCD and S_n detects 1, 2, 3, 17, 21 as outliers, along with these observations they detect few more observations as outliers.

2.4.4 Hawkins - Bradu - Kass dataset

The hbk is an artificial dataset generated by D. M. Hawkins et al. (1984) comprises 75 observations across four dimensions, featuring one response variable and three explanatory variables. Notably, the dataset provides a good example of masking effect. The first 14 points in the dataset are outliers (figure 2.11), grouped in two sets (1 – 10 and 11 – 14). Classical methods may fail to identify outliers beyond observations 12, 13 and 14 due to the masking effect which are masked. Robust distances can overcome the drawback of masking. P. J. Rousseeuw & Van Zomeren (1990) studied the dataset in detail in their paper.

2.4.5 Brain and Weight dataset

The data consists of brain weight in grams and body weight in kilograms of 28 animals. P. J. Rousseeuw & Van Zomeren (1990) discussed and analyzed this particular dataset in their paper graphically, along comparing the robust distance based MCD used in it with classical Mahalanobis distance. They have found 6, 14, 16, 17 and 25 as outliers in their work using distance measure and graphically analyzing it. We used same data here which is obtained from P. J. Rousseeuw & Leroy (2005) and analyzed using the distance methods compared in this chapter. The results below in figure 2.12 shows that MCD detects 2, 6, 7, 9, 12, 14, 15, 16, 24 and 25 as outliers, MCD distance we have used is inbuilt function in R. OGK detects 2, 5, 7, 9, 12, 14, 15, 16, 24 and 25 as outliers. S_n detects 2, 6, 7, 9, 12, 14, 15, 16 and 25 as outliers. Sh-S_n detects 6, 7, 14, 15, 16 and 25 as outliers. Here also RMDS fails to detect any observation as outlier. We can definitely confirm

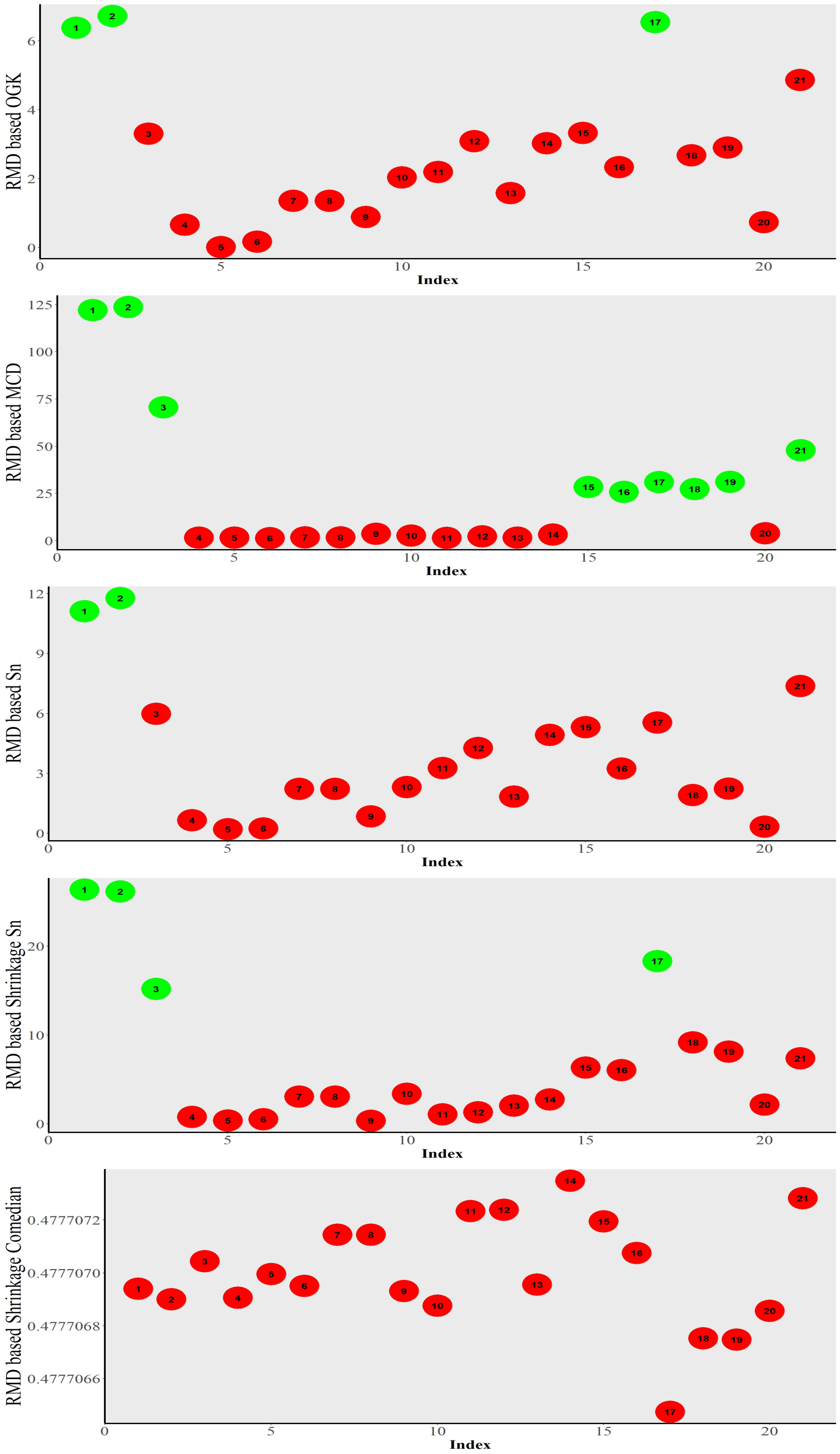


Figure 2.10: Stackloss data Index Plot of Robust Mahalanobis distance based on OGK, MCD, S_n, Sh - S_n, RMDS

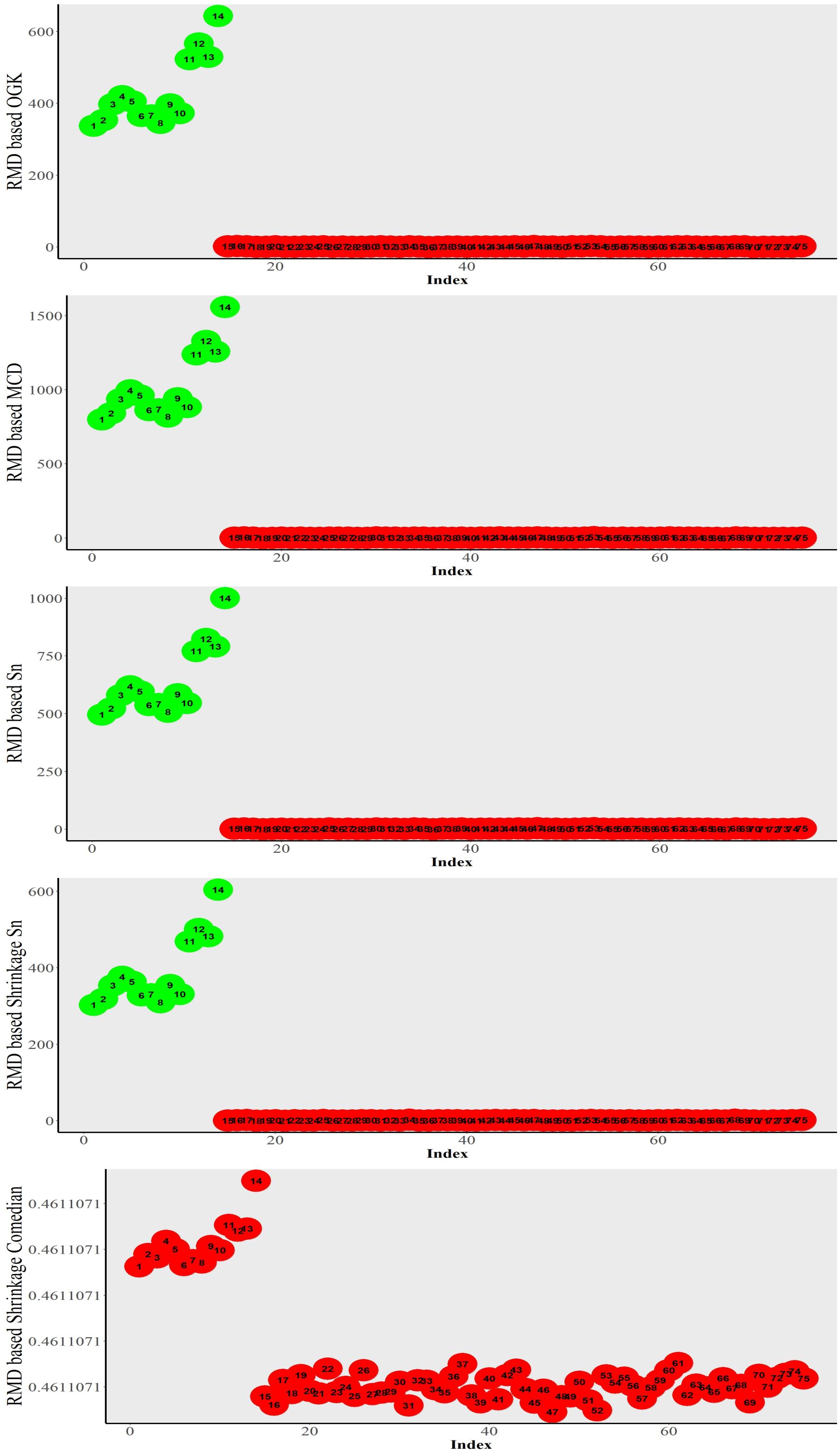


Figure 2.11: hbk data Index Plot of Robust Mahalanobis distance based on OGK, MCD, S_n, $Sh - S_n$, RMDS

that observations 6, 7, 14, 15, 16 and 25 as outliers in this particular data, because all the methods in our study detect those observations as outliers.

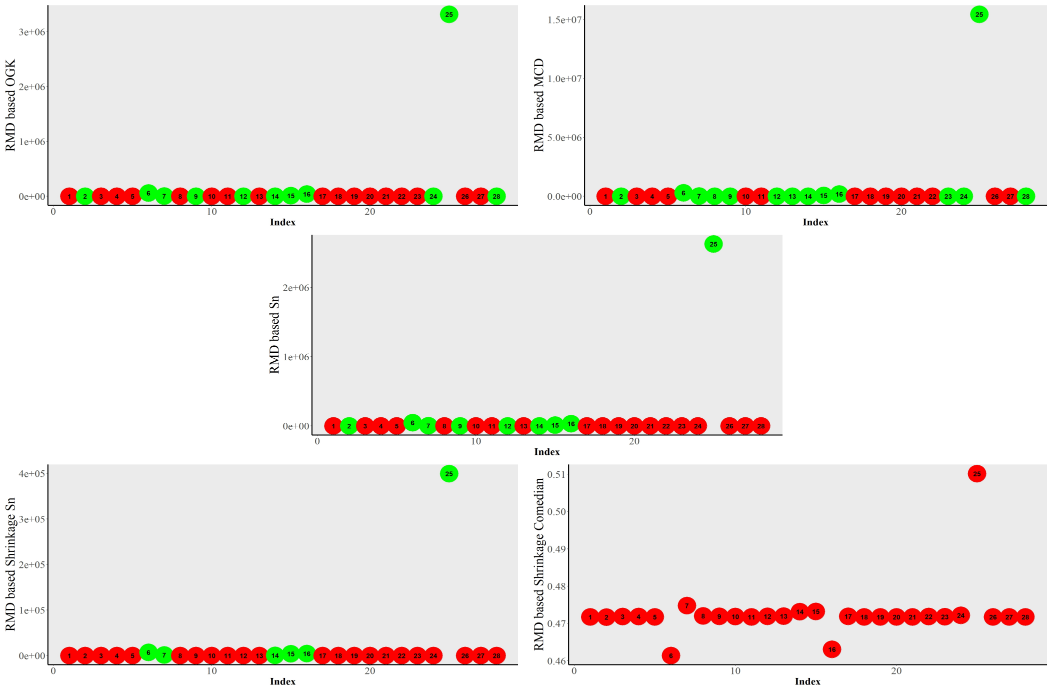


Figure 2.12: Brain data Index Plot of Robust Mahalanobis distance based on OGK, MCD, S_n, $Sh - S_n$, RMDS

Table 2.4: Summary Table on RMD Sh-S_n performance in considered benchmark datasets

Dataset Name	Observations considered to be outliers in the dataset	Observations detected by RMD Sh-S_n as outliers	FPR	TPR
Bushfire Data	7,8,9,10,11,32,33,34,35,36,37,38	7,8,9,10,11,12,32,33,34,35,36,37,38	0.03846	1
Milk Data	1,2,3,11,12,13,14,15,17,27,41, 44,74(72^{nd} point) and 77(75^{th} point)	1,2,4,11,12,13,14,15,25,27,28,29, 30,40,41,42,44,47,50,52,74,77	0.14285	0.8571
Stackloss Data	1,2,3,21	1,2,3,17	0.0588	0.75
hbk Data	1,2,3,4,5,6,7,8,9,10,11,12,13,14	1,2,3,4,5,6,7,8,9,10,11,12,13,14	0	1
Brain-Weight Data	6,14,16,17,25	6,7,14,15,16,25	0.0869	0.8
Salinity Data	3,5,16	3,4,6,5,16	0.0869	1
US-Judge Lawyers Data	5,7,8,12,13,14,23,35	5,7,8,14,17,20,22,23,35,41	0	0.375

2.4.6 US Judge Lawyers dataset

Sajesh & Srinivasan (2012) in their paper discussed the Lawyers rating of state judges in the US superior court dataset. The data consists of 43 observations on 12 variables. The observations 5, 7, 8, 12, 13, 14, 23 and 35 (total of 8 observations) are labeled as outliers in this dataset. We can see the results below in the graph

Table 2.5: FPR and TPR of different methods in benchmark dataset

	FPR					**TPR**				
Dataset Name	**Fast MCD**	**OGK**	**RMDS**	S_n	Sh-S_n	**Fast MCD**	**OGK**	**RMDS**	S_n	Sh-S_n
Bushfire Data	0.1538	0.1154	1	0.0385	0.0384	1	1	0	0.9167	1
Milk Data	0.1	0.1	1	0	0.1288	1	0.9286	0	0.7143	0.8571
Stackloss Data	0.2941	0.0588	1	0.1765	0.0588	1	0.5	0	0.75	0.75
hbk data	0	0	1	0	0	1	1	0	1	1
Brain-Weight	0.3043	0.2609	1	0.3043	0.0869	0.8	0.8	0	0.2174	0.8
Salinity Data	0.2	0.04	1	0.04	0.08	0.6667	0.6667	0	0.6667	1
US-Judge Data	0.2571	0.1714	1	0.08	0	0.75	0.75	0	0.25	0.375

Table 2.6: Summary table on benchmark datasets, observations detected by different method

		Observations detected by different methods as outliers				
Dataset	**Observations considered to be outliers**	**MCD**	**OGK**	**RMDS**	S_n	Sh-S_n
Bushfire	7,8,9,10,11,32,33,33,34,35,36,37,38	7,8,9,10,11,12,29,30,31,32,33,34,35,36,37,38	7,8,9,10,11,12,13,31,32 ,33,34,35,36,37,38		7,8,9,11,31,32,33,34,35,36,37,38	7,8,9,10,11,12,32,33,34,35,36,37,38
Milk Data	1,2,3,11,12,13,14,15,17,27,41,44,74(72^{nd} point),77(75^{th} point)	1,2,3,11,12,13,14,15,16,17,18,20,27,40,42,41,44,47,72,73,75	1,2,3,11,12,13,14,15,16,17,18,20,28,42,41,44,47,72,73,75		1,2,12,13,14,15,17,41,44,72	1,2,4,11,12,13,14,15,25,27,28,29,36,40,41,42,44,47,50,52,74,77
Stackloss	1,2,3,21	1,2,3,15,16,17,18,19,21	1,2,17		1,2	1,2,3,17
hbk data	1,2,3,4,5,6,7,8,9,10,11,12,13,14	1,2,3,4,5,6,7,8,9,10,11,12,13,14	1,2,3,4,5,6,7,8,9,10,11,12,13,14		1,2,3,4,5,6,7,8,9,10,11,12,13,14	1,2,3,4,5,6,7,8,9,10,11,12,13,14
Brain-Weight	6,14,16,17,25	2,6,7,8,9,12,13,14,15,16,23,24,25,28	2,6,7,9,12,14,15,16,24,25,28		2,6,7,9,12,14,15,16,25	6,7,14,15,16,25
Salinity Data	3,5,16	5,11,15,16,17,23,24	5,16,23		5,16,23	3,4,6,5,16
US Judge	5,7,8,12,13,14,23,35	5,7,8,10,14,18,19,20,22,23,29,31,34,35,43	5,7,8,10,14,20,23,29,31,32,35,43		5,7,20,23	5,7,8,14,17 20,22,23,35,41

2.13. The results show that MCD detects almost 15 observations as outliers, OGK detects almost 12 observations as outliers, among them, observations 5, 7, 8, 14, 23 and 35 previously mentioned also includes. Sh-S_n detects 5, 14 and 23 as outliers, and S_n detects 5, 7, 20 and 23 as outliers. RMDS doesn't detect any observation as outlier.

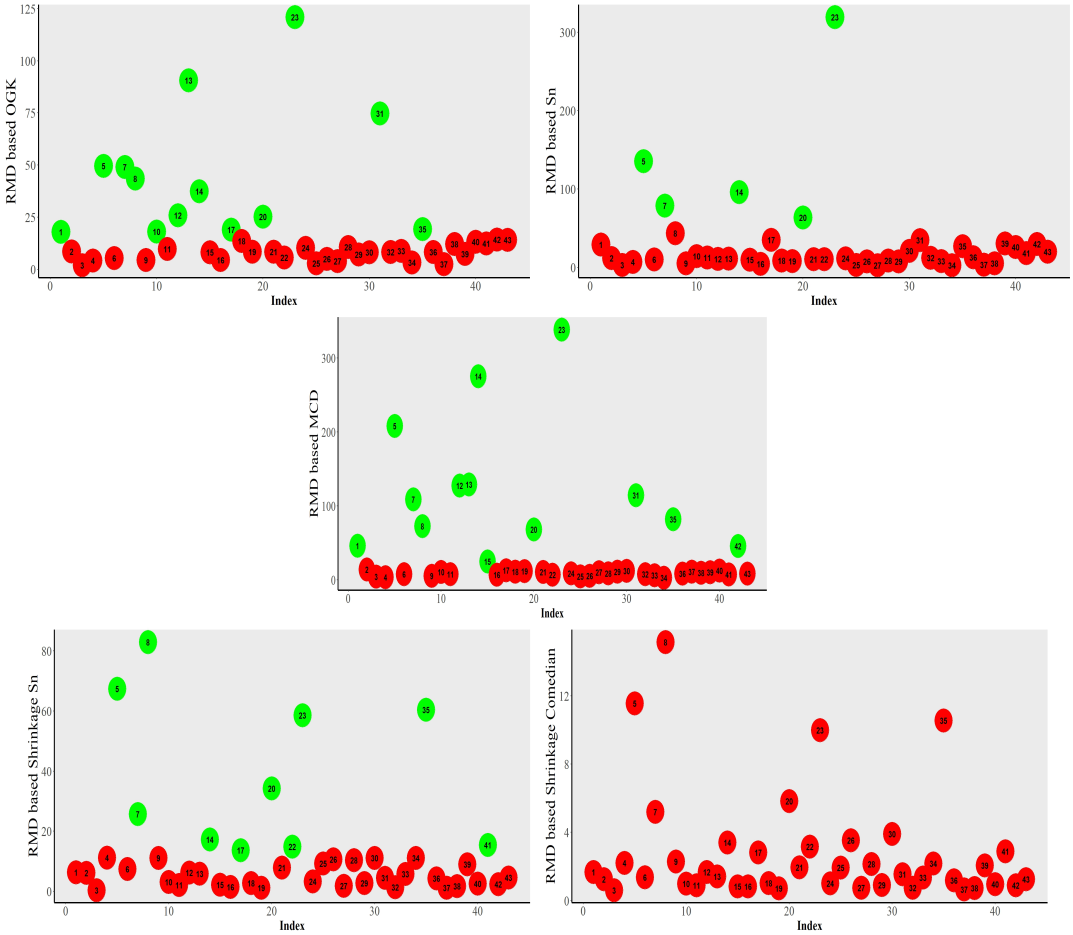


Figure 2.13: US judge data Index Plot of Robust Mahalanobis distance based on OGK, S_n, MCD, $Sh - S_n$, RMDS

2.4.7 Salinity dataset

This dataset comprises measurements of water salinity (salt concentration) and river discharge in North Carolina's Pamlico Sound. It records bi - weekly averages from March, April, and May during the years 1972 to 1977. The dataset was documented by Ruppert & Carroll (1980). In Carroll & Ruppert (1985), the physical background of the data is explained. They noted that observations 5 and 16 represent periods of very heavy discharge. They demonstrated that the outlier in observation 5 was concealed by observations 3 and 16; it was only after removing these observations that the influential nature of observation 5 became apparent. This dataset serves as a prime example of the masking effect. Results in the figure 2.14 shows that OGK and S_n detects 5, 16, 23 as outliers. Apart from the observations 5, 16, 23, MCD detects 11, 15, 17 and 24 also as outliers. All these three didn't detect the third observation as outlier. Our proposed measure detects 3, 5, 16 and 4, 6 as outliers. That is, our proposed measure additionally detects two observations as outliers. RMDS again fails to detect any observation as outlier.

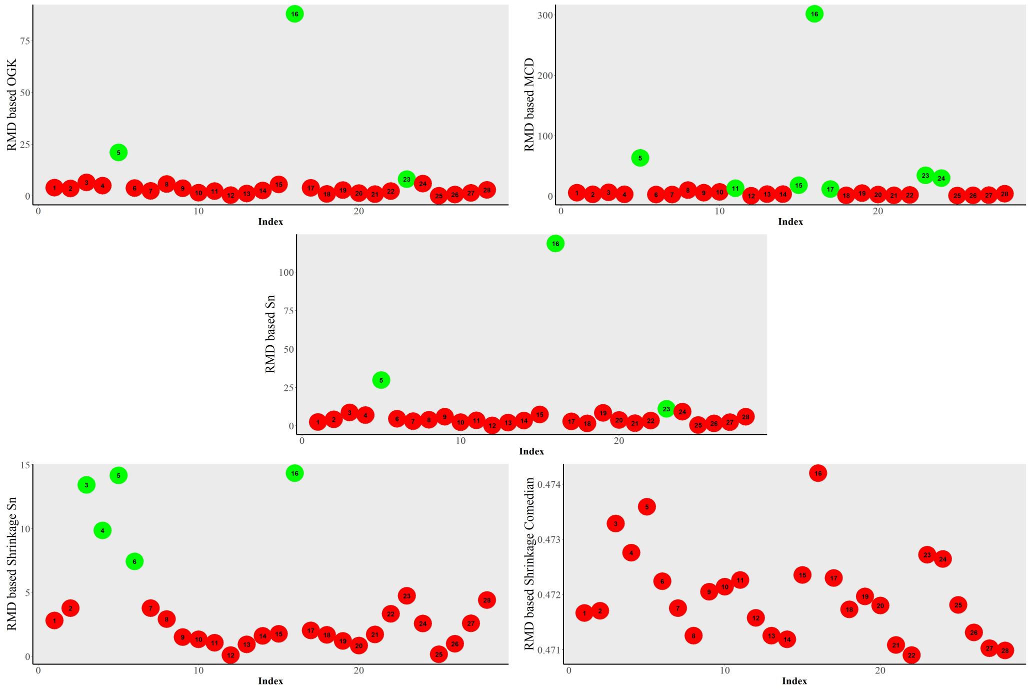


Figure 2.14: Salinity data Index Plot of Robust Mahalanobis distance based on OGK, MCD, S_n, $Sh - S_n$, RMDS

2.4.8 Lung cancer data from Malabar Cancer Centre (MCC), Thalassery

For the evaluation of distance measure performance, we collected data from Malabar Cancer Centre, Thalassery, Kannur, Kerala. The data constitutes information of Lung cancer patients of MCC in different stages with $n = 168$. The dataset comprises 29 variables, most of which are categorical. We focused specifically on the variables Albumin, Absolute Neutrophil Count (ANC), and patient survival days, excluding the categorical variables from our analysis. Missing observations were removed prior to analysis. Our proposed measure, along with other comparative measures, was applied to the selected variables. The results in the figure 2.15 indicate that observations 6, 21, 67, 147, and 153 are consistently identified as outliers by all the measures except RMDS. This gives an affirmation on necessity of checking outliers, even in medical dataset. It would be the discretion of the researcher to determine, what cause the patients to act as outliers in the dataset or how the variables considered in the study affects the patients overall characteristics.

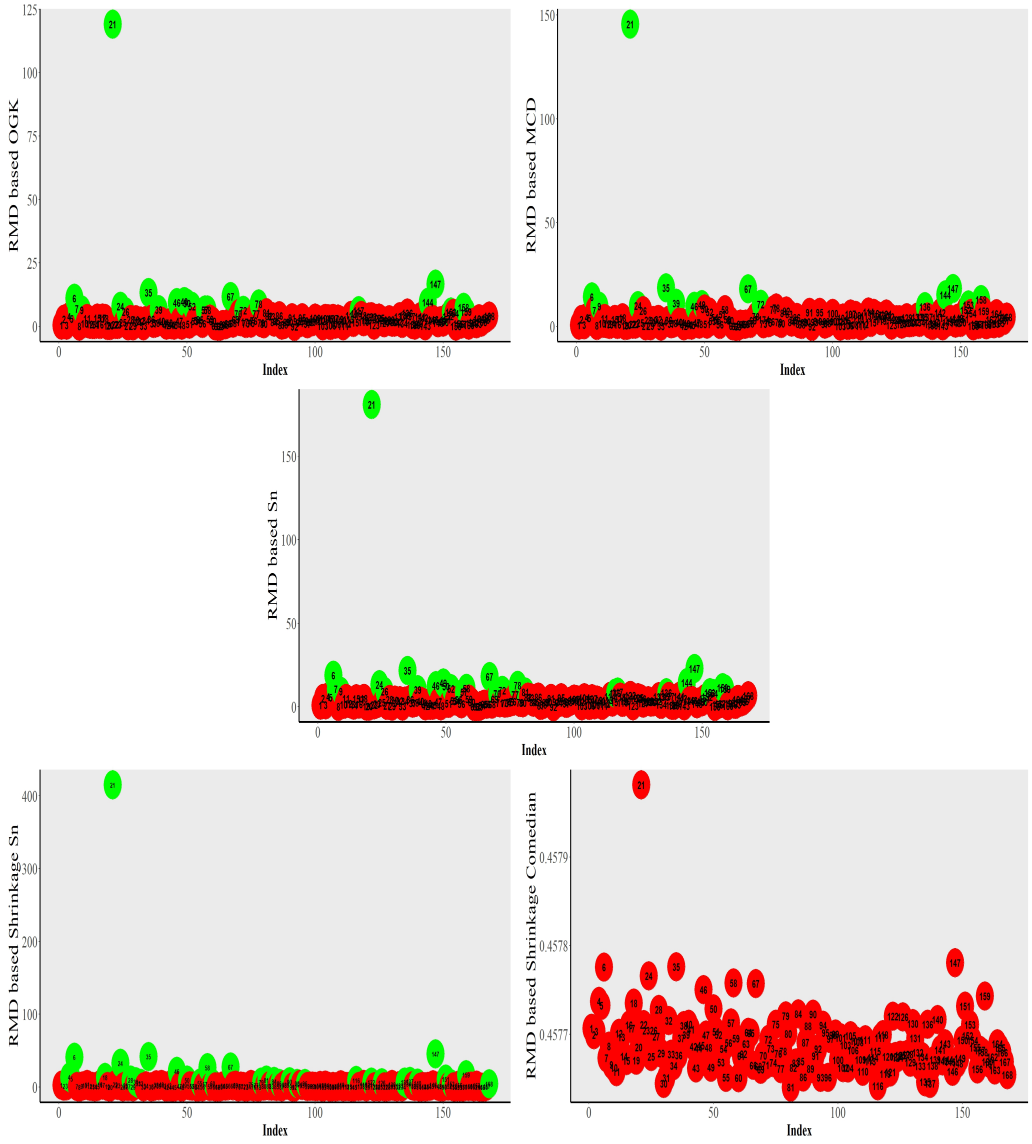


Figure 2.15: Lung cancer data Index Plot of Robust Mahalanobis distance based on OGK, MCD, S_n, $Sh - S_n$, RMDS

2.5 Summary

To address the sensitivity of classical sample mean and covariance estimates, we propose a robust alternative. We have assessed the efficiency and sensitivity curve of the $Sh\text{-}S_n$ estimator, and results are in favour to $Sh\text{-}S_n$ estimator. Mahalanobis distance is a commonly employed metric for detecting outliers in multivariate scenario, a fundamental and crucial aspect in data analysis. Several variants of Mahalanobis distance can be formulated based on the selection of covariance and location estimators, and a range of estimators are accessible in existing literature. In this chapter, we introduce a robust Mahalanobis distance based on shrinkage $Sh\text{-}S_n$ covariance estimator and shrinkage L_1-median location estimator. We have evaluated the performance of $Sh\text{-}S_n$ estimator based distance measure along with FAST-MCD, OGK, RMDS, RMD based on S_n through an extensive simulation study. The results of the simulation study highlight RMD $Sh\text{-}S_n$ proficiency in outlier detection. The FPR result comparison under 0% contamination in data shows that RMD $Sh\text{-}S_n$ exhibits lowest FPR than all other methods in all cases considered in the study. This offers us empirical assurance in the effectiveness and applicability of RMD $Sh\text{-}S_n$, due to its low FPR in 0% contaminated datasets. Through our study, we have found that the RMD based $Sh\text{-}S_n$ displays a commendable performance in all levels of contamination considered, as evidenced by its high TPR. In addition to our proposed measure, it has also been observed that the RMD based on S_n and OGK exhibits a high TPR similar to us, in many of the cases like large deviated outlier included samples considered in the study. Regarding FPR, RMD $Sh\text{-}S_n$ generally maintains a lower rate compared to the other methods except RMDS that were used in our study. It is worth noting that in many cases, the RMDS measure did a negligible improved performance in terms of FPR, but fails in terms of TPR and even attained 0% TPR for higher contamination. Thus, RMD $Sh\text{-}S_n$ measure exhibits a good performance along with other methods in terms of TPR and have some advantages over others considered, which we discussed in the simulation section. To evaluate the performance in collinear data of proposed measure, we generate highly correlated data within a specified dimension. And the results substantiate the ability of RMD $Sh\text{-}S_n$ to detect outliers in collinear data. Empirically we have assessed the affine equivariance property of the measure, which demonstrates that

RMD Sh-S_n can effectively identify outliers within transformed data. We have also studied the empirical breakdown of RMD Sh-S_n. The results provide an assurance that RMD Sh-S_n possesses substantially low breakdown value and exhibits resilience in the presence of outliers. RMD Sh-S_n is applied to a number of benchmark and a real-life dataset and in all datasets the measure is capable of detecting outliers. The results are summarized and tabulated (2.4 ,2.5 and 2.6). All of these findings provide a strong assurance that proposed estimator is suitable to use in multivariate analysis.

Chapter 3

Robust multiple regression based on Sh-S_n estimator

3.1 Introduction

Regression Analysis is extensively utilized across various academic disciplines, including social science, health science, engineering, and physical science. Its primary reliance on the OLS estimator makes it vulnerable, particularly in the presence of outliers. Outliers can be characterized as observations that significantly deviate from the majority of data points within a dataset. Consequently, robust regression was developed as a superior and more effective alternative to OLS estimates when dealing with contaminated points in the dataset. A variety of robust regression algorithms are available, some of which also exhibit resistance to outliers. In this chapter, we examine several popular and effective robust regression estimators applicable to multiple linear regression models with contaminated data. Furthermore, the covariance matrix representation of classical normal equations provides a straightforward approach to executing multiple regression. Here, we offer a robust reweighted regression (RSh-S_n) based on Sh-S_n covariance estimator, along with a discussion of robust regression estimators. The primary objective is to employ simulated datasets to compare different estimators, depending exclusively on comprehensive empirical simulations to determine the characteristics of the proposed estimator. In the following section, we will briefly review the OLS approach and emphasize the importance of robust methodologies, as well as the significant robust estimators that have emerged over time. A new reweighted regression estimator based on a robust covariance matrix technique is proposed. In this chapter, the suggested estimator is compared with other robust regression estimators using a range of simulation techniques. Additionally, the properties of the proposed estimator are thoroughly examined through extensive simulations. The chapter concludes with applications of these estimators to real - world data.

Classical linear regression is about estimating the unknown regression parameters. Consider a linear multiple regression model for a sample of size n:

$$y = \alpha + \mathbf{x}_1\beta_1 + \cdots + \mathbf{x}_p\beta_p + \epsilon \tag{3.1.1}$$

where, α the unknown intercept, $\beta = (\beta_1, \ldots, \beta_p)^t$ the unknown regression parameter vector of size $(p \times 1)$, $\epsilon = (\epsilon_1, \ldots, \epsilon_n)^t$, is a $n \times 1$ vector of error components which follows $i.i.d.$ Normal distribution with mean zero and constant variance, and $\mathbf{x}=\begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{2p} \\ 1 & x_{31} & x_{32} & \cdots & x_{3p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}$ is a $n \times p$ matrix of n observations on each p variables where $\mathbf{x}_i$ represents the $i^{th}, (i = 1, \ldots, p)$ column of $\mathbf{x}$, $y = (y_1, \ldots, y_n)^t$ is a $n \times 1$ vector of n observations on dependent variables. The classical approach OLS, which minimizes the sum of squared residuals, is the traditional method for estimating the parameters of the model 3.1.1 is defined as:

$$\hat{\beta} = \min_{\beta} \sum_{i=1}^{n} (y_i - \mathbf{x_i^t}\beta)^2. \tag{3.1.2}$$

When the data meet these classical estimators' assumptions, they perform well. Solving 3.1.2 gives us $\hat{\beta}_{\text{OLS}} = (\mathbf{x^t x})^{-\mathbf{1}}\mathbf{x^t}y$ and the expected value of y_i would be $\hat{y_i}$=$x_i^t\hat{\beta}_{\text{OLS}}$.

The least square estimator performs effectively when the error terms are well - behaved or when the underlying assumptions are satisfied. Despite its mathematical elegance and computational simplicity, the OLS estimator lacks robustness, as it is sensitive to outliers in the data. Identifying leverage points and outliers in a dataset can be highly challenging without thorough examination. Detecting outliers in simple linear regression using a scatter plot is usually straightforward, but this approach becomes limited in multiple linear regression. Outliers in either the $\mathbf{x}$ or y directions pose a considerable risk to OLS estimator. To identify these outliers, one can employ statistical or graphical methods. Mahalanobis distance is a statistical method employed to identify outlying $\mathbf{x}$ points. We cannot say Mahalanobis distance as a perfect method, as it fails to detect the outliers in the y direction. Regression diagnostics methods involve removing one observation at a time to see how it affects the regression coefficients, helping to identify influential points or observations that significantly impact the model. Once identified, these points can either be removed or corrected before further analysis. The diagnostics, known as deletion statistics, include measures like Cook's distance, Studentised residuals, DFFITS, DFBETAS and Jacknife residuals which quantify individual data points' influence. However, calculating these statistics can become complex when multiple outliers are present, complicating the assessment of their collective impact. Robust regression estimation is alternative strategy for handling outliers. Robust regression methods aim to create estimators that are immune to outliers. Consequently, many robust estimators have been developed to address the limitations of the OLS in such cases.

3.2 RSh-S_n multiple regression estimator

In this chapter, we propose to use robust shrinkage estimator instead of the classical estimator in 3.1.2 which was proposed in previous chapter. A detailed study on properties of these estimates done there. The principle behind shrinkage estimation lies in the notion of "shrinking" an estimator $\hat{E}$ towards a target estimator $\hat{T}$, which serves to effectively diminish estimation errors. This approach capitalizes on the fact that, while the shrinkage target $\hat{T}$ may exhibit bias, it typically displays lower variance compared to the estimator $\hat{E}$. By utilising shrinkage estimation, we introduced robust $Sh - S_n$ covariance estimator and location vector in previous chapter and make use of the same location estimate and covariance estimate in equation 3.1.2 and thereby propose a reweighted regression estimator. Consider $\mathbf{z} = (\mathbf{x}, y)$, the joint variable with location and covariance matrix μ, Σ, respectively. The associated squared Mahalanobis distance for each observation $\mathbf{z}_i, i = 1, 2, 3, \ldots, n$, based on $\hat{\mu}_{Sh}$ and $\hat{\Sigma}_{Sh}$ be:

$$\mathrm{RD}^2(\mathbf{z}_i) = (\mathbf{z}_i - \hat{\mu}_{Sh})^t \hat{\Sigma}_{Sh}^{-1} (\mathbf{z}_i - \hat{\mu}_{Sh}). \tag{3.2.1}$$

The weight function based on the above defined robust Mahalanobis distance is $w_i = w(\mathrm{RD}^2(\mathbf{z}_i))$, where a weight of 1 is assigned to the observations $(\mathbf{z}_i)$ with a Mahalanobis distance less than $\frac{\chi^2_{0.95,p} \times \mathrm{median}(\mathrm{RD}^2(\mathbf{z}_i))}{\chi^2_{0.5,p}}$. Thus the reweighted shrinkage location and S_n covariance matrix be defined as:

$$\hat{\mu}_{wSh} = \frac{\sum_{i=1}^n w_i \mathbf{z}_i}{\sum_{i=1}^n w_i}, \quad \hat{\Sigma}_{wSh} = \frac{\sum_{i=1}^n w_i (\mathbf{z}_i - \hat{\mu}_{wSh})(\mathbf{z}_i - \hat{\mu}_{wSh})^t}{\sum_{i=1}^n w_i}. \tag{3.2.2}$$

The regression estimates, namely reweighted shrinkage regression estimates based on the reweighted $\hat{\mu}_{wSh}$ and $\hat{\Sigma}_{wSh}$ be defined as:

$$\hat{\beta}_{RSh-S_n} = (\hat{\Sigma}_{wSh})_{xx}^{-1} (\hat{\Sigma}_{wSh})_{xy}, \quad \hat{\alpha}_{RSh-S_n} = (\hat{\mu}_{wSh})_y - \hat{\beta}^t_{RSh-S_n} (\hat{\mu}_{wSh})_{\mathbf{x}} \tag{3.2.3}$$

where $(\hat{\beta}_{RSh-S_n}, \hat{\alpha}_{RSh-S_n})^t$ be our reweighted shrinkage regression estimator (RSh-S_n). The scale estimate of errors based on above defined regression estimator is:

$$\hat{\sigma} = (\hat{\Sigma}^1)_{yy} - (\hat{\beta}^1)^t (\hat{\Sigma}^1)_{xx} \hat{\beta}^1.$$

The next step is to reweighting by taking into considerations the residuals based on the one step reweighted shrinkage based regression estimator:

$$r_i = y_i - (\hat{\beta}^1)^t \mathbf{x}_i - \hat{\alpha}^1,$$

The Mahalanobis distance for the above defined residuals be:

$$d(r_i) = ((r_i)^t (\hat{\sigma})^{-1} r_i)^{1/2}.$$

Let $wr_i = w(d^2(r_i))$ be the weighting function with respect to the residual Mahalanobis distance, where a weight of 1 be assigned to those residuals with a Mahalanobis distance less than $\chi^2_{1,0.99}$. Define $\mathbf{u}_i = (\mathbf{x}_i^t, 1)^t$, then:

$$\phi^{RSh-S_n} = ((\hat{\beta}^{RSh-S_n})^t, \hat{\alpha}^{RSh-S_n})^t = (\sum_{i=1}^n wr_i \mathbf{u}_i \mathbf{u}_i^t)^{-1} \sum_{i=1}^n wr_i y_i \mathbf{u}_i \tag{3.2.4}$$

be our two step reweighted regression estimator based on shrinkage S_n $(RSh - S_n)$.

3.3 Simulation Study

This section presents the results of a simulation study that compares the performance of the suggested RSh-S_n regression estimator with OLS and a few of the robust regression techniques that were previously discussed: MM, S, LTS, LMS. Simulations were done in R software. The lmrob.S function from robustbase package for S estimator, the rlm function for MM estimator in MASS package, the lmsreg function from MASS package for LMS and the lqs function from MASS package for LTS estimator utilised. We have used the in - built functions but changed the tolerance level, sub - samples required, and number of iterations of the respective functions according to our simulation study. Consider the linear model:

$$y = \alpha + \mathbf{x}\beta + \epsilon \tag{3.3.1}$$

where $\mathbf{x}$ is $n \times p$ matrix, y is $n \times 1$ vector of response variables, β is $p \times 1$ vector of unknown regression coefficients, α unknown intercept, and ϵ are *i.i.d.* error variables. For the study of simulation, sample sizes considered in different situations are n = 80, 100, 150, 200, 500, 1000 and dimensions considered are p = 5, 10, 15, 20, 30. 1000 times each of the simulation scenarios are repeated in our study. We have considered simulation scenarios similar to those found in the literature R. Maronna & Morgenthaler (1986), Gervini & Yohai (2002), Croux et al. (2003), P. J. Rousseeuw et al. (2004), Agulló et al. (2008), Yu & Yao (2017), Cabana et al. (2020). In the first scenario, we generate the response variables from a standard normal distribution with $\beta = 0$, $\alpha = 0$, and errors are considered to be standard Gaussian. Let the scenario be denoted as NE. The independent variable $\mathbf{x}$ follows multivariate normal distribution with mean vector $0_{p\times 1}$ and covariance matrix, a $p \times p$ identity matrix. In the second scenario, normal errors are considered as in NE but with a probability δ contamination in response and independent variables. The contaminants in independent variables are taken from $\mathrm{N}(\lambda\sqrt{\chi^2_{p,0.99}}, 1)$ and response variables from $\mathrm{N}(k\sqrt{\chi^2_{1,0.99}}, 1)$ where $\lambda, k = 0, 0.5, 1, 1.5, 2, 3, 4, 5, 6, 7, 8, 9, 10$. Let the respective scenario be denoted as NEO further in the article. The percentage of contamination considered are $\delta = 10\%, 20\%$. If $\lambda = 0$ and $k > 0$ we will get y

outliers; if $\lambda > 0$, $k = 0$ we will get good leverage points; and if $\lambda > 0$, $k > 0$ we obtain bad leverage points. Thus, our choice of λ and k gives data that ranges from extreme outliers to intermediate outliers. Under NE criteria, it is clear that OLS will be have maximum efficiency. To analyze the performance of the estimators, we have used the efficiency of the estimators as the metric. Let $\phi = (\beta^t, \alpha)^t$ be the joint $(p+1) \times 1$ vector of regression coefficients. The finite sample efficiency of any robust estimator R defined as in Gervini & Yohai (2002):

$$\text{Eff} = \frac{\frac{1}{N}\sum_{i=1}^{N} ||\hat{\phi}_{OLS}^{(i)} - \phi||_2^2}{\frac{1}{N}\sum_{i=1}^{N} ||\hat{\phi}_{R}^{(i)} - \phi||_2^2}. \tag{3.3.2}$$

Table 3.1 shows the simulation results of relative efficiency (NE) for the joint regression estimator ϕ, RSh-S_n, and other estimators used for the comparison. The results indicate that our proposed method is more efficient than the robust S, LTS, and LMS estimators. Additionally, $RSh-S_n$ along with the MM and SR estimators, achieves efficiency values close to one. The table values confirm the performance of our proposed estimator in non-contaminated data. The results indicate that, regardless of sample size and dimension, our estimator demonstrates higher relative efficiency compared to other robust methods. Furthermore, its efficiency improves as the dimension increases. In contrast, the LTS and LMS methods consistently perform poorly across all dimensions and sample sizes, highlighting their limitations for use in non-contaminated datasets. A robust method that matches the performance of OLS in clean data while outperforming it in contaminated cases is ideal, making it highly suitable for practical applications. The efficiency value of our proposed estimator remains consistently close to one, indicating that it performs well in non-contaminated data.

To assess the performance in robustness property, i.e., NEO criteria, Mean Square Error(MSE) of the estimated parameter $\phi = (\beta^t, \alpha)^t$ averaged over N simulation runs was identified. And then we considered the maximum MSE across different values of k for each λ, i.e.,

$$\text{MSE}_\lambda(.)_{\max} = \max_{k\ \epsilon\ 0,0.5,\ldots,10} \text{MSE}_{\lambda,k}(.),$$

Finally, the maximum of the MSE is considered as the metric for NEO performance

Table 3.1: Finite Sample efficiency in case of Normal errors

n	p	$RSh - S_n$	MM	S	LTS	LMS	SR
	5	0.90111	0.8901934	0.3133835	0.1687298	0.1870315	0.91122
80	10	0.91998	0.9222617	0.2686664	0.1470201	0.128014	0.92895
	15	0.933981	0.9306423	0.2520925	0.1427853	0.1423768	0.93391
	20	0.91189	0.8765167	0.2664371	0.1724709	0.1269676	0.96774
	30	0.98881	0.8412267	0.273741	0.1836285	0.1050502	0.91334
100	5	0.91198	0.9625692	0.3307681	0.1678611	0.1163469	0.92122
	10	0.95003	0.9477326	0.2681845	0.1338582	0.1469313	0.93895
	15	0.92289	0.9122369	0.2664602	0.1358261	0.1444049	0.93581
	20	0.94478	0.9410851	0.2956971	0.1368254	0.1305699	0.96994
	30	0.92339	0.8474065	0.3050582	0.167848	0.1229417	0.92293
150	5	0.93328	0.9304732	0.3096604	0.11059407	0.14979494	0.891122
	10	0.95001	0.9483747	0.3067942	0.13424255	0.10780639	0.938895
	15	0.94991	0.9346738	0.2526533	0.11922365	0.11141836	0.958391
	20	0.93391	0.9345712	0.2508206	0.11280803	0.10334599	0.967874
	30	0.93443	0.8910862	0.2851218	0.09904998	0.09132823	0.922234
200	5	0.96991	0.9410785	0.3035179	0.11344367	0.12879938	0.89012
	10	0.94411	0.9328904	0.2738827	0.0992274	0.09473856	0.93915
	15	0.94471	0.9434414	0.2659222	0.08857581	0.08025978	0.95691
	20	0.92287	0.9314504	0.2494056	0.08732885	0.09294969	0.93399
	30	0.92278	0.9114167	0.2663687	0.08309036	0.07708978	0.92224
300	5	0.99981	0.9836872	0.3414343	0.10978863	0.11293767	0.90012
	10	0.94441	0.9344473	0.2969611	0.08428847	0.07586824	0.92981
	15	0.96227	0.9637224	0.2685592	0.06195658	0.06536106	0.93441
	20	0.95619	0.9543658	0.2491195	0.05939695	0.05864878	0.97001
	30	0.95671	0.9463761	0.2569738	0.05703498	0.05549416	0.92881
500	5	0.98827	0.9707854	0.352695	0.08741342	0.08629351	0.92199
	10	0.95189	0.9504772	0.3079973	0.06210263	0.06096699	0.93391
	15	0.93891	0.9365855	0.2447813	0.04703221	0.04344129	0.94417
	20	0.95178	0.9402005	0.2720199	0.04044304	0.03248954	0.91887
	30	0.96881	0.9514596	0.2416872	0.03636119	0.03106638	0.93241
1000	5	0.98811	0.947679	0.3242796	0.05436184	0.05822269	0.92289
	10	0.95587	0.9488504	0.3353234	0.03039559	0.03082224	0.94805
	15	0.96111	0.9575694	0.2741748	0.02430555	0.02085405	0.96791
	20	0.94445	0.935774	0.2761304	0.0203216	0.01564593	0.97987
	30	0.95011	0.9417903	0.268742	0.01838522	0.01896355	0.97733

assessment and is defined as:

$$\mathrm{MSE}()_{\max} = \max_{\lambda \ \epsilon \ 0,0.5,\ldots,10} \mathrm{MSE}_{\lambda}(.)_{\max}.$$

The table 3.2 shows that our proposed method exhibit least $\mathrm{MSE}()_{\max}$ than other methods for both 10% and 20% contamination. Also our proposed regression estimator shows a decreasing $\mathrm{MSE}()_{\max}$ as the dimension increases. And even for higher contamination in dataset, our estimator possess least $\mathrm{MSE}()_{\max}$ than other compared estimators. Classical method OLS fails to have robustness and exhibit drastically high $\mathrm{MSE}()_{\max}$ value. Among other robust estimators, MM estimator shows minimum $\mathrm{MSE}()_{\max}$ and LMS shows worst performance in terms of $\mathrm{MSE}()_{\max}$ in robustness.

Table 3.2: $\mathrm{MSE}()_{\max}$ of estimates for checking robustness - NEO case

p	δ	OLS	$RS-hS_n$	MM	S	LTS	LMS	SR
5	10%	9.8874	0.00238	0.09585	0.05366	0.05885	0.05997	0.04857
10		8.9993	0.00266	0.03826	0.02715	0.06561	0.06356	0.02943
15		7.6712	0.00287	0.02485	0.02604	0.07844	0.07844	0.02253
20		6.8201	0.00313	0.02179	0.02253	0.09406	0.09397	0.01978
30		3.1178	0.00370	0.02709	0.01968	0.13581	0.13832	0.01598
5	20%	9.9991	0.02448	0.19517	0.14019	0.12295	0.13931	0.22584
10		9.0001	0.00645	0.09110	0.07951	0.13289	0.13031	0.12067
15		8.5619	0.00789	0.07720	0.05918	0.15259	0.15697	0.07964
20		7.2211	0.00710	0.09124	0.04359	0.23309	0.23358	0.06508
30		4.1871	0.00784	0.06429	0.03797	1.17839	1.55502	0.04654

3.4 Equivariance Property

Equivariance, breakdown, and robustness properties determine an estimator's actual usefulness rather than theoretical goodness. Three forms of equivariance are taken into consideration for regression estimators.

- Regression equivariant is equivalent to adding the coefficients of this linear function to the estimators if we convert the dependent variable by adding a linear function of independent variables.

- y - equivariant is defined as: The estimators transform correctly if the response variable is changed linearly. These two mentioned properties can be summarized as follows:

$$\hat{\phi}_{RSh-S_n}(\mathbf{x}, yb + \mathbf{x}g + u) = \hat{\phi}_{RSh-S_n}(\mathbf{x}, y)b + (g^t, u)^t \qquad (3.4.1)$$

where $b \epsilon \mathbf{R}$ be any non - zero constant, g be any $p \times 1$ vector and $u \epsilon \mathbf{R}$ is any constant. Keeping $\mathbf{x}$ same and transforming the dependent variable as $yb + \mathbf{x}g + u$, then the resulting transformed estimators are: $\hat{\beta}^{new}_{RSh-S_n} = b(\hat{\beta}_{RSh-S_n}) + g$ and $\hat{\alpha}^{new}_{RSh-S_n} = b\hat{\alpha}_{RSh-S_n} + u$.

- If the independent variables undergo a linear transformation, the equivalent transformed estimator can be defined as follows for $\mathbf{x}$ - equivariance: $\hat{\phi}_{RSh-S_n}(\mathbf{x}A, y) = ((\hat{\beta}_{RSh-S_n})^t(A^{-1})^t, \hat{\alpha}_{RSh-S_n})^t$. i.e., if the independent variables are transformed using any non - singular $p \times p$ matrix A, then the resulting new regression estimators are: $\hat{\beta}^{new}_{RSh-S_n} = A^{-1}\hat{\beta}_{RSh-S_n}$ and the intercept remain the same.

Since it is impossible to investigate all possible transformations, R. A. Maronna & Zamar (2002) and Sajesh & Srinivasan (2012) suggested in their papers that A matrices be randomly generated in order to check $\mathbf{x}$ - equivariance as $A = TD$, where D is a $p \times p$ diagonal matrix with diagonal entries that are uniformly and independently distributed and T is a random orthogonal matrix. We examined the $\mathbf{x}$ - equivariance property of the proposed estimator using the above described approaches. Also, we propose a random non - zero b, g and u for checking regression and y - equivariance. The approximate affine equivariance of the initial estimates $\hat{\mu}_{Sh}$ and $\hat{\Sigma}_{Sh}$ is demonstrated by means of extensive simulation in the previous chapter. In this work, our main concern is around the estimator $\hat{\phi}_{RSh-S_n} = (\hat{\beta}^t_{RSh-S_n}, \hat{\alpha}_{RSh-S_n})^t$. We have studied the equivaraince property of the proposed estimator on transformed data as described above. Similar to the simulation scenarios NE and NEO, we have considered $\delta = 0\%$, 10%, 20% contamination here. $\hat{\phi}_{RSh-S_n}$ is obtained from the non - transformed data and recorded. Then the data is transformed as in the explanations given above and record $\hat{\phi}^{new}_{RSh-S_n}$ from the transformed data. The MSE is calculated between $\hat{\phi}^{new}_{RSh-S_n}$ and the estimator should obtain if the equivariance property holds. Tables 3.4 and 3.3 shows $\text{MSE}_\lambda(\hat{\phi}^{new}_{RSh})_{\max}$ for each λ. For $\lambda = 0$ maximum MSE of y - equivariance increases with an increase in dimension and contamination. Also for higher dimensions, the error in y - equivariance is negligibly low across all contaminations. The same pattern we could see in $\mathbf{x}$ - equivariance too. The values give us an empirical confirmation regarding the equivariance property.

3.5 Sensitivity curve

The sensitivity curve of an estimator shows how an estimator performs when a small contamination replaces a single observation in the dataset. That is, it measures how the estimator responds to the local effect of a single observation. An estimator with a bounded sensitivity curve has an influence function that remains bounded. For the evaluation of sensitivity curve, normal errors are considered, as in NE case. The independent variables are taken from $\mathrm{N}(0_{p\times 1}, \mathbf{I}_{p\times p})$ and response variables from $\mathrm{N}(0,1)$. The contaminant, single observation in independent variable and dependent variable taken from $N(\lambda\sqrt{\chi^2_{p,0.99}}, 1)$ and $\mathrm{N}(k\sqrt{\chi^2_{1,0.99}}, 1)$ respectively, where $\lambda, k =$ -10, -8, -6, -4, -2, 0, 2, 4, 6, 8, 10.

$$\mathrm{SC}_n(y) = (n+1)\left(\hat{\theta}_{n+1}(y_1, \ldots, y_n, y) - \hat{\theta}_n(y_1, \ldots, y_n)\right).$$

Here $\hat{\theta}$ represents the regression coefficient of the respective estimators. Then, across λ, k values, we found the norm of the above - defined difference of the regression coefficients. Each value is obtained after 1000 simulations. After obtaining the norm values across λ, k values, we obtain the maximum norm over k at different values of λ and it is plotted. The maximum of norm across different values of λ, k helps to understand the resistance on contaminants of robust estimators. The figure 3.1 shows that RSh-S_n estimator remains bounded compared to other estimators, irrespective of dimension. This gives the assurance empirically on the bounded nature of the influence function of our proposed estimator.

3.6 Breakdown property

The maximum percentage of outliers that the estimator can safely accept is measured by the breakdown point. The breakdown point can have a maximum value of 50%. Considering high contamination levels, simulations such as those conducted by Sajesh & Srinivasan (2012) can be used to investigate the empirical breakdown

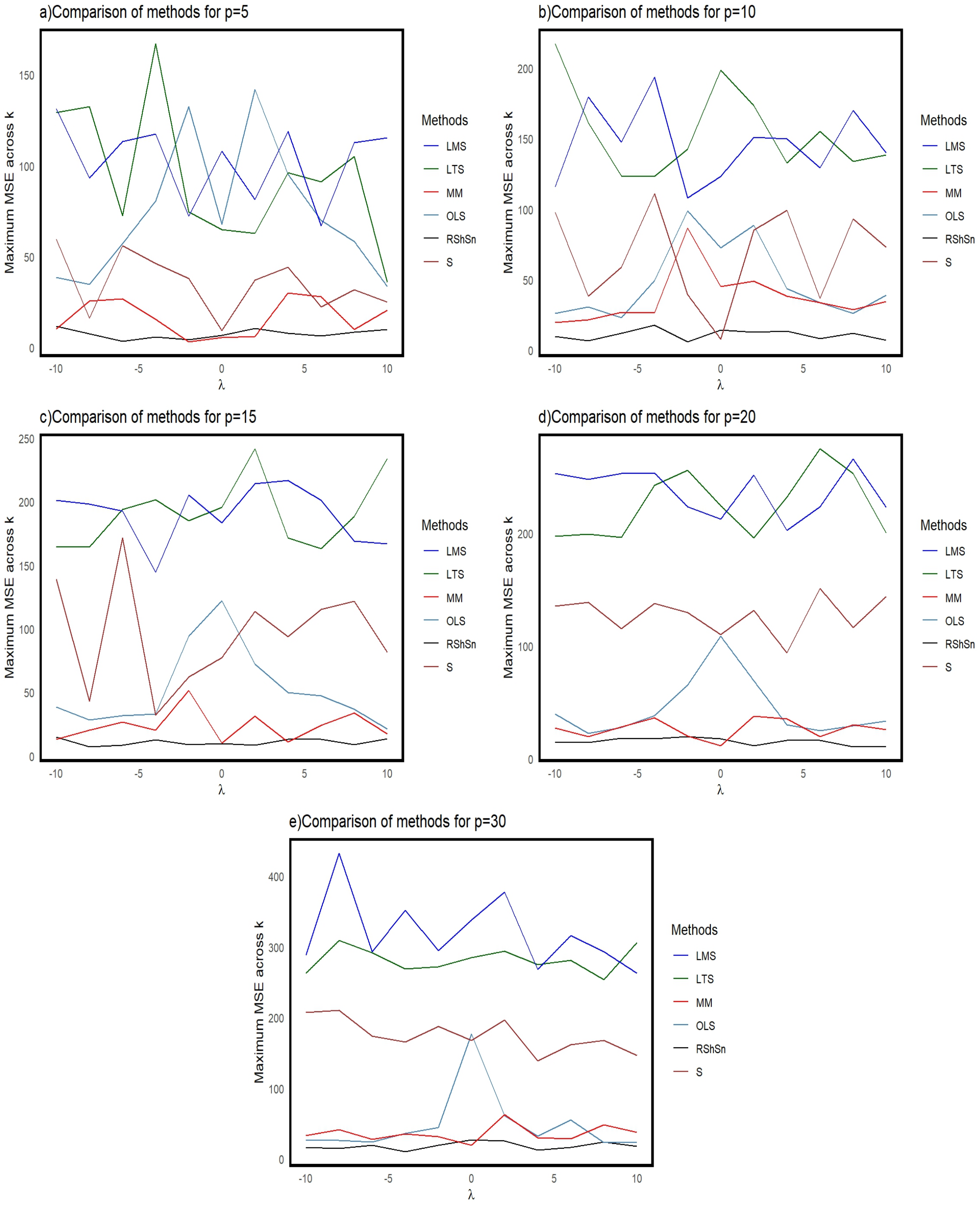


Figure 3.1: Sensitivity Curve of our proposed regression method along with other methods a)$p = 5$ b)$p = 10$ c)$p = 15$ d)$p = 20$ e)$p = 30$

Table 3.3: Affine y - equivariance and regression equivariance $\mathrm{MSE}_\lambda(.)_{\max}$ values

	$p = 5$			$p = 30$		
λ	$\delta = 0\%$	$\delta = 10\%$	$\delta = 20\%$	$\delta = 0\%$	$\delta = 10\%$	$\delta = 20\%$
0	0.01692	0.03357	0.08543	0.0006	0.02763	0.09313
0.5	0.01693	0.03212	0.03269	0.00065	0.00526	0.00292
1	0.01645	0.03259	0.02042	0.00061	0.00541	0.00266
1.5	0.01687	0.03309	0.02099	0.00058	0.00533	0.00272
2	0.01683	0.03275	0.02016	0.00061	0.0058	0.00271
3	0.01767	0.03195	0.02029	0.00061	0.00552	0.00248
4	0.01677	0.03449	0.02045	0.00063	0.00557	0.00294
5	0.01736	0.03307	0.02059	0.00062	0.00558	0.00263
6	0.01725	0.03275	0.02049	0.00064	0.00507	0.00254
7	0.01659	0.03275	0.02065	0.0006	0.00547	0.00262
8	0.01706	0.03267	0.02093	0.00063	0.0052	0.0025
9	0.01662	0.03249	0.02051	0.00059	0.00544	0.00279
10	0.01710	0.03332	0.02015	0.00061	0.0054	0.00284

Table 3.4: Affine $\mathbf{x}$ - equivariance $\mathrm{MSE}_\lambda(.)_{\max}$ values

	$p = 5$			$p = 30$		
λ	$\delta = 0\%$	$\delta = 10\%$	$\delta = 20\%$	$\delta = 0\%$	$\delta = 10\%$	$\delta = 20\%$
0	9.55e-27	0.17066	0.35649	0.13403	0.22271	0.12016
0.5	1.51e-6	0.94962	0.36883	0.00089	0.09106	0.26992
1	4.23e-24	0.36592	0.09488	0.00055	0.13569	0.12746
1.5	1.71e-27	0.35441	0.22789	0.00251	0.10637	0.08758
2	3.81e-26	0.08954	0.64404	0.00024	0.10065	0.12487
3	3.81e-26	0.51310	0.56414	0.00656	0.11819	0.14615
4	1.43e-27	0.48495	0.49300	0.00656	0.11531	0.13136
5	2.11e-23	0.73876	0.3392	0.00142	0.11602	0.11201
6	7.41e-25	0.92863	0.44410	0.00194	0.14555	0.16444
7	1.21e-27	0.48099	0.40906	0.00517	0.14801	0.11676
8	0.00016	0.26175	0.19499	0.03388	0.10334	0.12029
9	4.80e-7	0.59238	0.35093	0.00106	0.13357	0.12236
10	1.13e-5	0.34573	0.91354	0.00379	0.19618	0.11513

value. We suggest examining whether the error and the bias are controlled in these circumstances in order to assess the effectiveness of the suggested RSh-S_n estimator, even though low levels of contamination are assumed, making these scenarios less relevant in practical applications. We have considered NEO scenario for checking breakdown, with percentage of contamination $\delta = 30\%$, 40%, 45%. For evaluation, we considered $\text{MSE}()_{\max}$ as defined:

$$\text{MSE}()_{\max} = \max_{\lambda\epsilon 0,0.5,1,\ldots,10}\text{MSE}_{\lambda}(.)_{\max}$$

The results in the tables 3.5 show that RSh-S_n estimator possesses the least $\text{MSE}()_{\max}$ than others, especially for higher contamination like 40% and 45%. Also, we can see a decrease in $\text{MSE}()_{\max}$ with an increase in the dimension. Even for 45% contamination, our method, RSh-S_n, possesses least $\text{MSE}()_{\max}$, which gives us an empirical assurance of having high breakdown property.

Table 3.5: $\text{MSE}()_{\max}$ table for breakdown property

	$p = 5$			$p = 30$		
Method	$\delta = 30\%$	$\delta = 40\%$	$\delta = 45\%$	$\delta = 30\%$	$\delta = 40\%$	$\delta = 45\%$
OLS	3.43891	4.57897	3.58334	2.4692	4.5482	1.2738
$RSh - S_n$	0.22585	0.56321	0.96679	0.0113	0.0323	0.0607
MM	0.49268	1.3743	4.0716	3.3505	5.5076	5.5949
LTS	0.29355	0.84809	1.70897	0.5239	0.7636	0.8144
LMS	0.28607	0.88574	1.82799	0.6210	0.8764	0.9119
S	0.28507	0.86094	2.57749	0.1356	0.4699	0.6267
SR	0.3125	0.6782	1.0122	0.02118	0.02442	0.07023

3.7 Real - Life dataset

For the model comparison of the estimators in real - life datasets, we used Mean Square Error (MSE), Mean Absolute Percentage Error (MAPE), Mean Absolute Deviation (MAD) and Akakike Information criteria (AIC). We haven't considered the coefficient of determination. Initially we considered the same metric and determined. As we all know, the coefficient of determination can be determined in two ways. However, when we applied both ways to the same dataset, we obtained different values for the same metric in data with a high contaminated outliers. Thus, we dropped using R^2 due to its sensitivity to outliers. We compare the different estimators by looking for the one that get lowest of MSE, MAPE, MAD and AIC.

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$$

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{Y_i - \hat{Y}_i}{Y_i}\right|$$

$$\text{MAD} = \frac{1}{n}\sum_{i=1}^{n}|Y_i - \hat{Y}_i|$$

$$\text{AIC} = n \times ln\left[\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2\right] + 2p$$

3.7.1 Mineral data

Smith et al. (1984) in their paper studied the investigation of how various elements from the Golden Grove massive sulphide deposits disperse across the lateritic landscape. The authors conducted measurements of the concentrations (in parts per million) of 22 chemical elements in 53 rock samples from Western Australia. Two variables from the aforementioned data are studied by Maronna in his book. Here we use the same data given in Maronna. The data in question contains details on the Zinc (*study variable*) and Copper (*explanatory variable*) content deposits. And we here tried to explore the relationship of these two elements. For the purpose of model comparison, we consider the MSE, MAPE, and MAD of the study variables. Also, we consider AIC for comparison purposes. This particular dataset possesses outliers in it. Upon examining the fitted versus residual(figure 3.2), Q-Q plots (figure 3.4), it is clear the OLS exhibits, the respective lines attracted towards the point 15, but for robust estimator plot of RSh-S_n estimator along with others we couldn't find such a pattern. Similar false performance of OLS can be found in other datasets we recognized here. Observation 15, 2, 25, 3, 39 be detected as extreme lying observations by all the estimators in residual versus fitted plot. And if we take a closer look at these observations behavior in standardized residual plot (figure 3.3), one could find the smooth pattern of the points inhibited by the aforementioned observations. The Q-Q plot too substantiates the fact that the aforementioned observations cause non - normality in the data. The observations of data were supposed to be studied in detail and should have found the reasons for exhibiting these kinds of patterns by the observations. The coefficient and model parameter values are tabulated below in table 3.6.

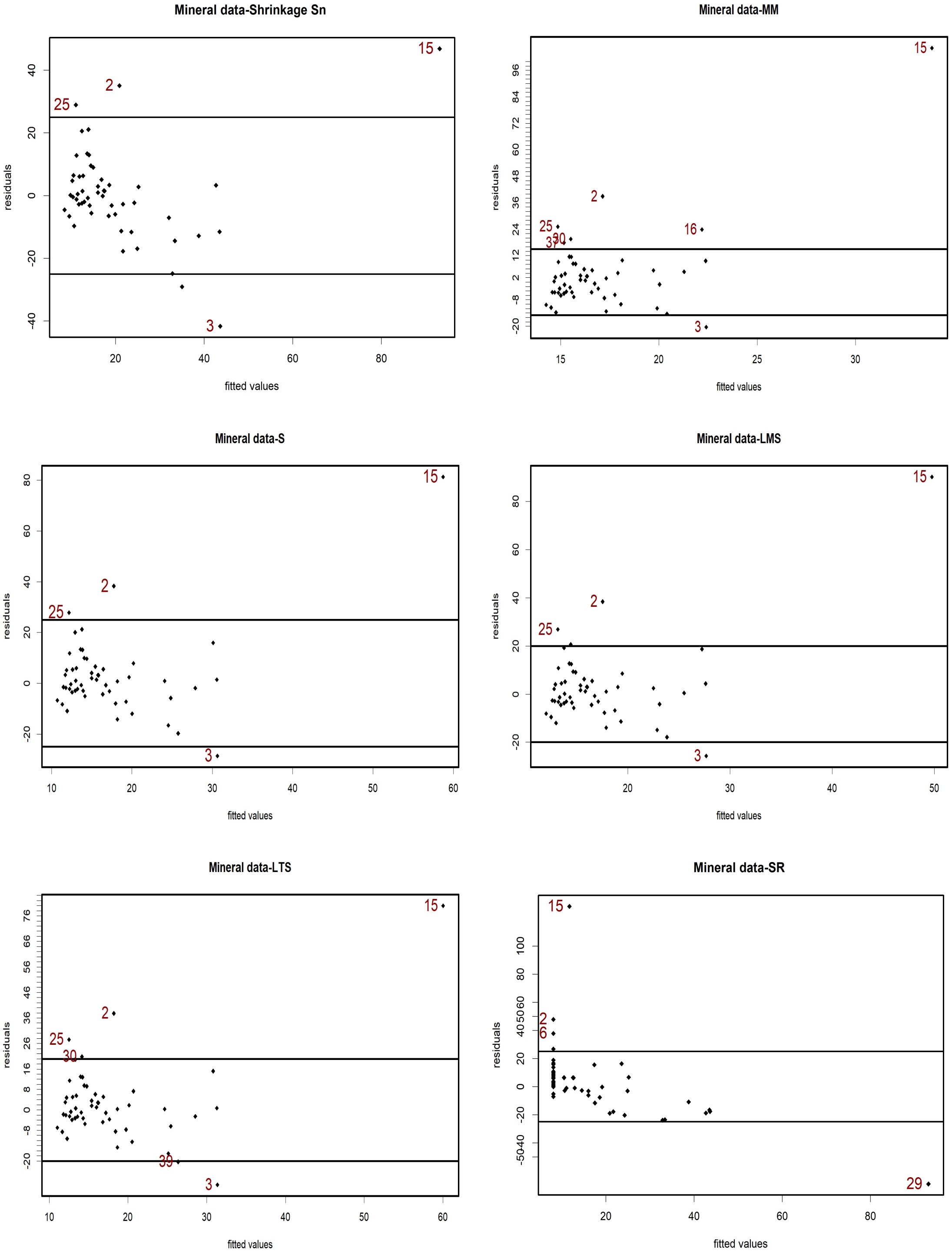


Figure 3.2: Mineral data Fitted values versus Residuals Plot of $RSh - S_n$, MM, S, LMS, LTS,SR

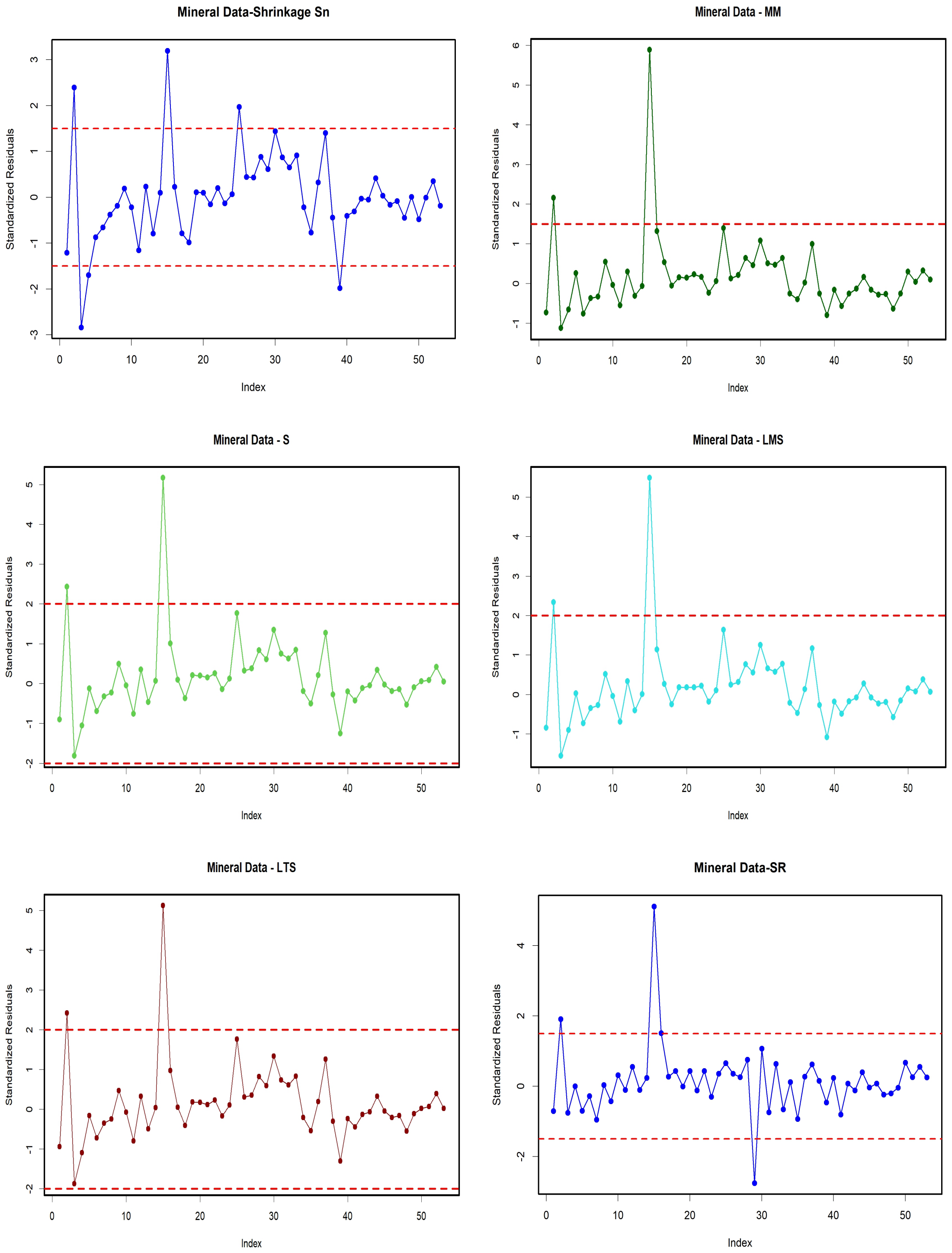


Figure 3.3: Mineral data Standardized Residuals Index Plot for outlier detection of $RSh - S_n$, MM, S, LMS, LTS ,SR

Table 3.6: Mineral Data

Method	Intercept	Slope	RSE	MSE	MAPE	MAD	AIC
$RSh - S_n$	10.24283	0.07720	18.03592	331.4453	1.48007	13.0142	301.5835
MM	14.13856	0.031207	18.03535	327.2739	1.017315	9.833758	308.9123
LTS	10.71525	0.07792	15.60709	245.5813	1.038663	9.28958	293.6923
LMS	11.79	0.06	16.45569	272.7899	1.013786	9.443962	299.2612
S	10.41003	0.076318	15.74303	249.8429	1.013658	9.324528	294.6041
OLS	11.4759	0.06313	7.90696	64.51997	0.789537	6.23915	204.0148
SR	7.96063	0.13457	25.07286	604.9256	1.082492	14.42829	343.4706

3.7.2 Learning data

The simulation data consisted of student motivation ($\mathbf{x}_1$), learning facilities ($\mathbf{x}_2$), and student learning outcomes in the cognitive domain (y). Both motivation and facilities were assessed through a questionnaire with scores ranging from 0 to 30, while learning outcomes were evaluated using a test with scores ranging from 0 to 100. Data is tabulated in Jana et al. (2023). 20 respondents participated in the study. Figures (3.5, 3.6, 3.7) show that data contains outliers in it, and the observations 7, 14, 15, and 18 were detected as outliers. If we closely observe the data, we can confirm these observations as outliers or observations require close surveillance. For example, if we look in to the observations 7, 18, the respective students express high student motivation and very good learning facility but the student learning outcome is very low. The low learning outcome of a student who felt good motivation in learning and good facilities contradicts, explicitly state the student should be monitored, this discrepancy suggests that despite favorable learning conditions, other unobserved factors may be influencing their academic performance. These cases give warning alarm to the researchers, for further investigation to understand the underlying causes and provide necessary interventions. We explored the data with all the estimators used in our comparison study, and the results are shown in tables. The tabulated results in table 3.7 and plots assure that RSh-S_n is reliable.

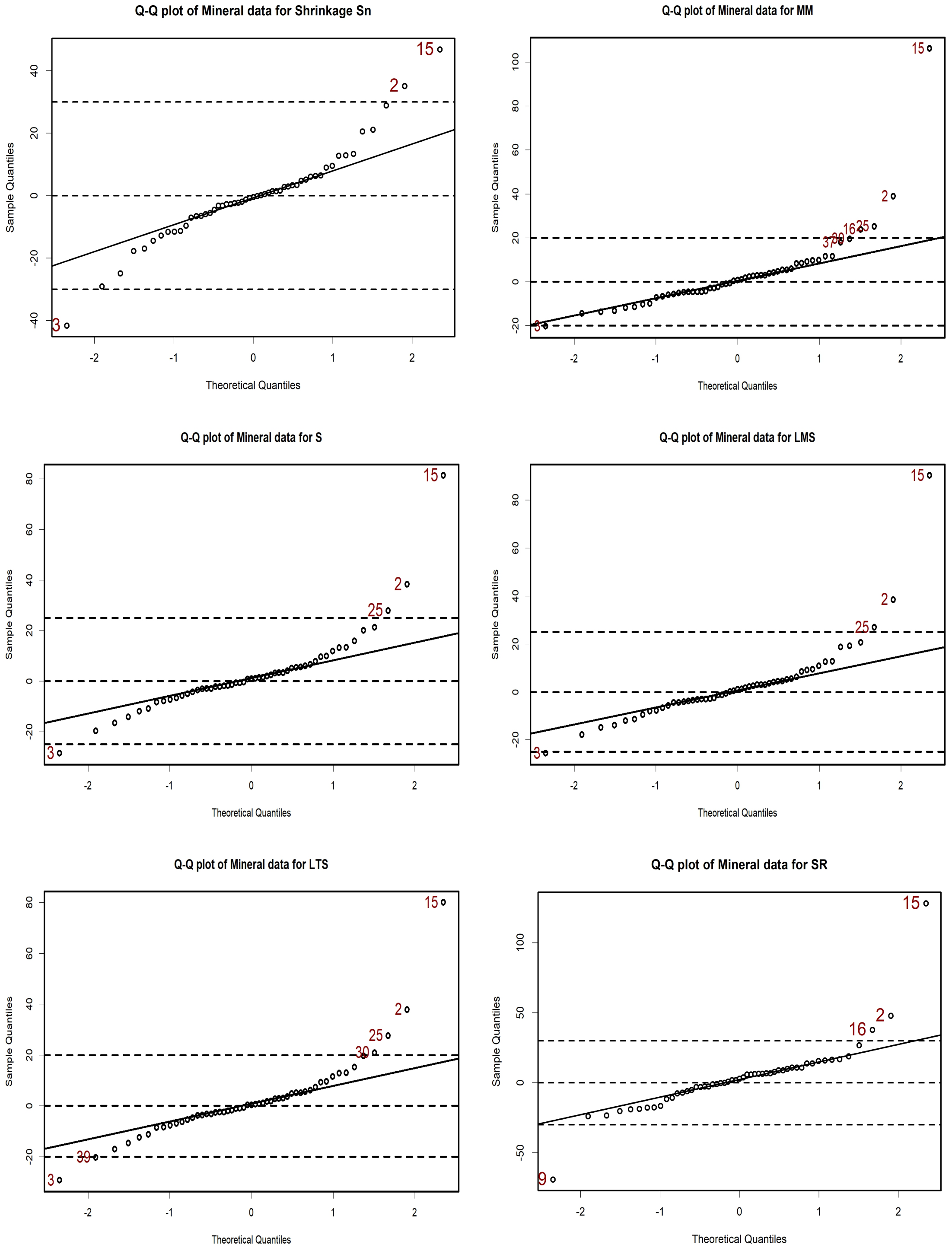


Figure 3.4: Mineral data Q-Q plot of Residuals of $RSh - S_n$, MM, S, LMS, LTS, SR

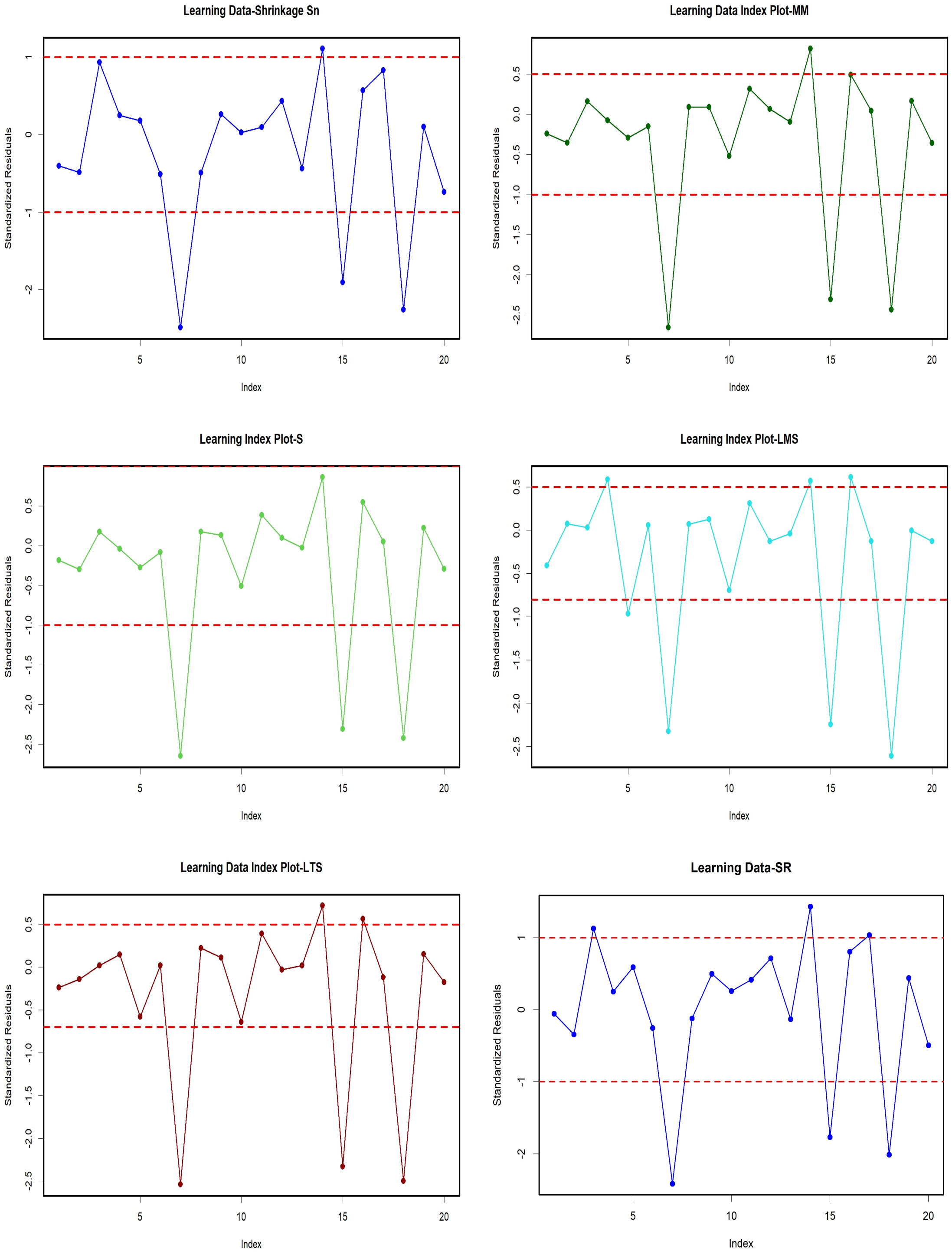


Figure 3.5: Learning data Standardized Residuals Index Plot for outlier detection of $RSh - S_n$, MM, S, LMS, LTS, SR.

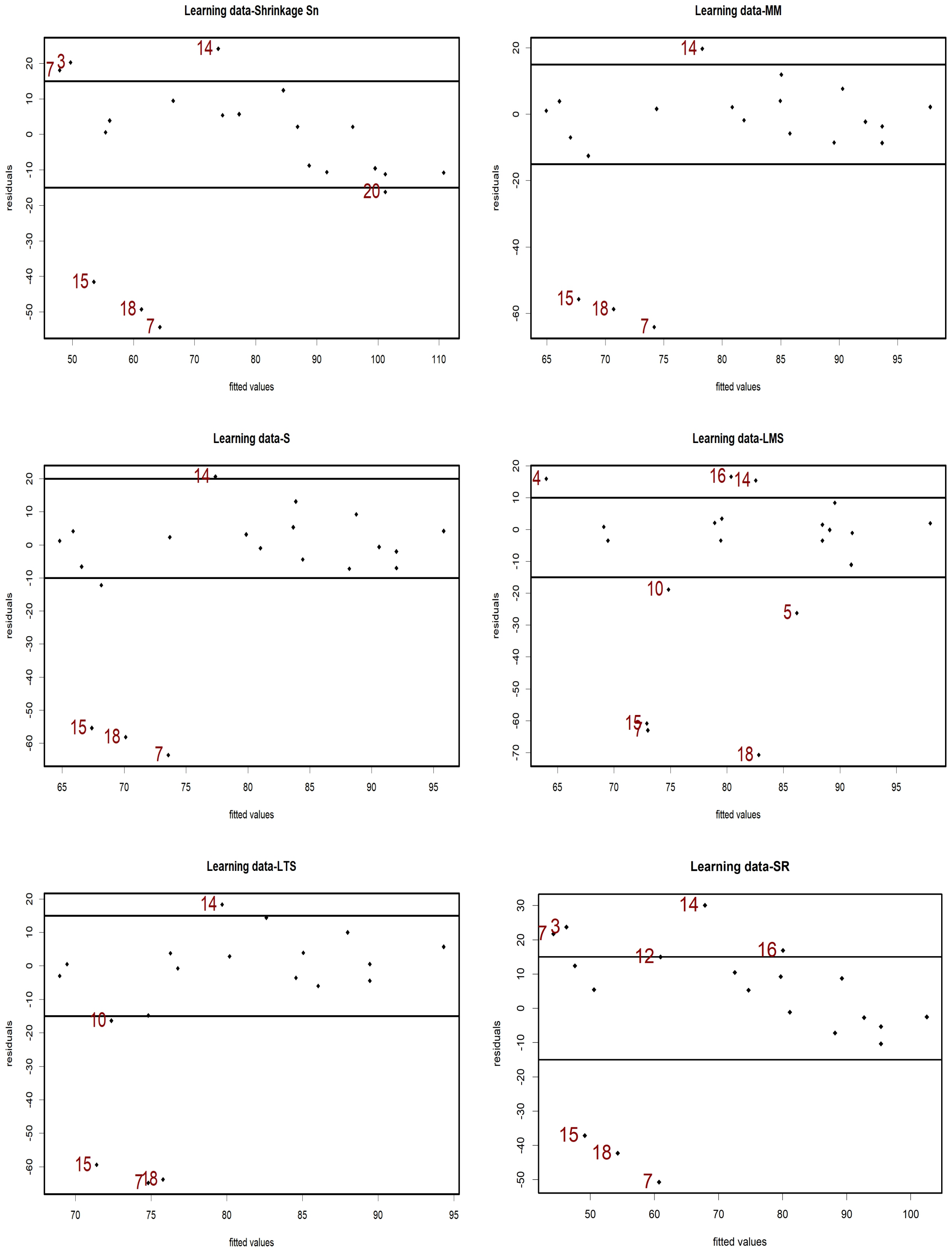


Figure 3.6: Learning data Fitted values versus Residuals Plot of $RSh - S_n$, MM, S, LMS, LTS, SR.

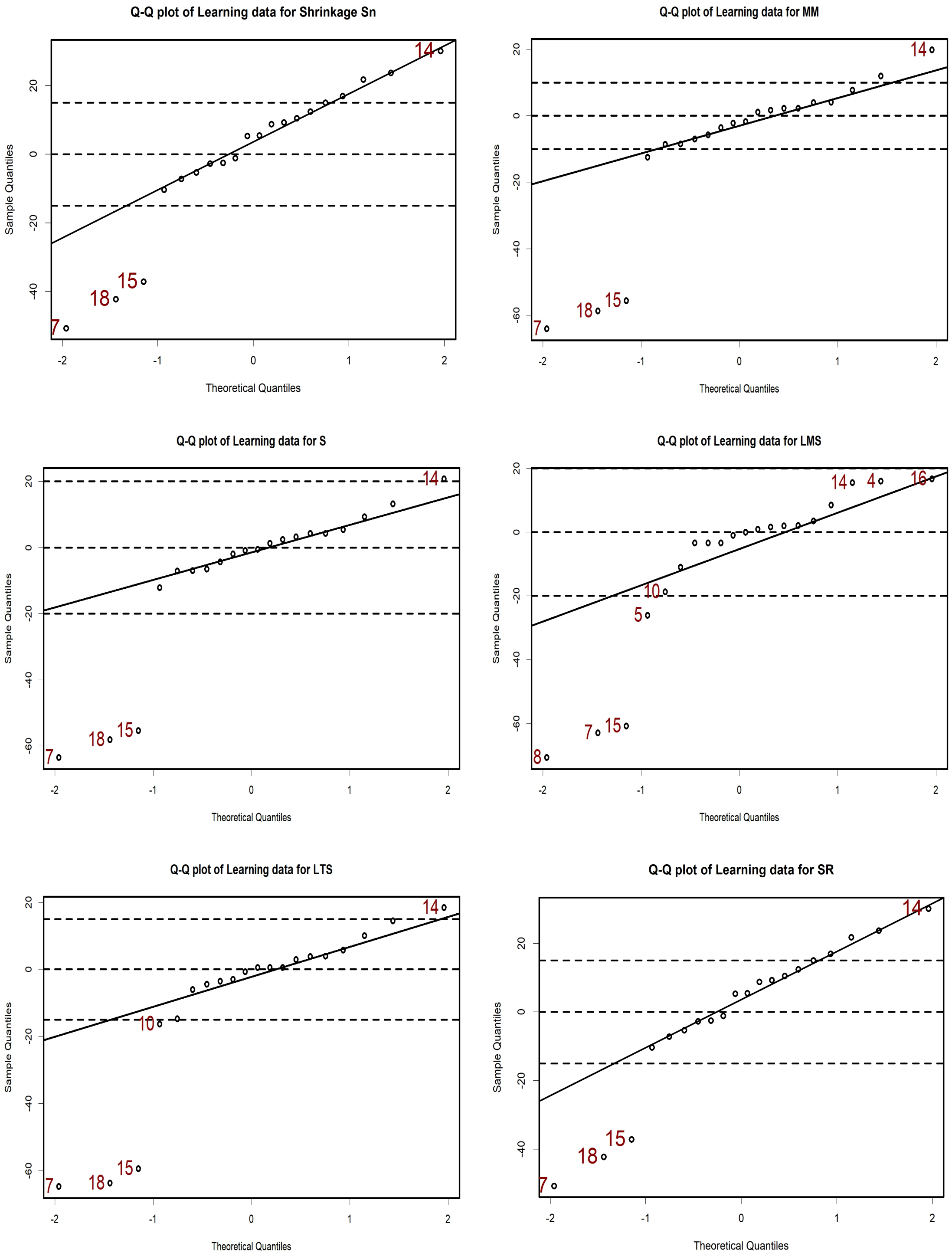


Figure 3.7: Learning data Q-Q plot of Residuals of $RSh - S_n$, MM, S, LMS, LTS, SR.

Table 3.7: Learning Data

Method	Intercept	Slope $\mathbf{x}_1$	Slope $\mathbf{x}_2$	RSE	MSE	MAPE	MAD	AIC
$RSh - S_n$	50.11493	1.91120	1.77704	21.7768	477.2288	0.75461	9.35787	129.91775
MM	39.30157	0.8221	1.12754	24.14265	585.8677	0.86269	14.16605	131.4619
LTS	50.4268	0.9756	0.4878	25.51272	653.8988	0.90765	14.87439	133.6591
LMS	51.6364	1.9091	-0.3636	27.10937	737.9178	0.95359	16.44091	136.0766
S	40.5031	0.7626	1.0814	21.29027	578.4258	0.85499	14.06815	131.2062
OLS	-1.789	1.9178	2.03	21.00165	444.0691	0.70329	15.9306	125.9196
SR	-1.373	1.433	2.0337	21.0275	445.0811	0.80911	16.7816	126.6918

3.7.3 Aircraft data

Aircraft data mentioned in Gray (1985) consists of 23 observations in 5 variables. The response variable is cost and four explanatory variables, namely, aspect ratio, life to drag ratio, weight of the plane, maximal thrust. The standardized residual plot (figure 3.8) shows some observations deviate and inhibits the smooth pattern of the data. And if we investigate the fitted versus residual plot (figure 3.9), the same observations as in the residual plot (figure 3.8) remain outlying. P. J. Rousseeuw & Van Driessen (2000) quoted some of the datasets, which include the Aircraft dataset and the outlying observations in the dataset. Similar to their findings, it is apparent, RSh-S_n estimator based plot detects observations 2, 3, 10, 11, 12, 16, 17, 18, 19, 20, 22 as outliers. All other robust estimator based plots detects these observations as outliers. The regression coefficients and model metrics are tabulated and given in table 3.8. We can affirm that RSh-S_n estimator successfully detects the outliers and performs better than OLS in the presence of outliers.

Table 3.8: Aircraft data

Method	Intercept	Slope				RSE	MSE	MAPE	MAD	AIC
$RSh - S_n$	6.0121	-2.0006	1.6189	0.00182	-0.0008	11.443	135.95	0.32531	6.1762	120.98
MM	6.1417	-3.2306	1.6711	0.00192	-0.0009	10.339	111.89	0.34389	5.5301	112.50
LTS	14.4816	-4.6712	1.9305	0.00198	-0.0013	11.643	140.56	0.35657	5.9568	121.75
LMS	5.7789	-2.0064	3.3645	0.00079	-0.0005	17.4118	308.17	0.29575	8.6583	139.81
S	13.3733	-4.0219	1.5413	0.00170	-0.0009	11.8193	144.69	0.33773	5.9442	122.42
OLS	-3.7913	-3.8529	2.4883	0.00349	-0.0019	7.0921	55.298	0.58333	5.6856	100.29
SR	4.62125	-3.2716	1.8726	0.00212	-0.0011	9.65777	98.2725	0.36041	5.40455	113.5181

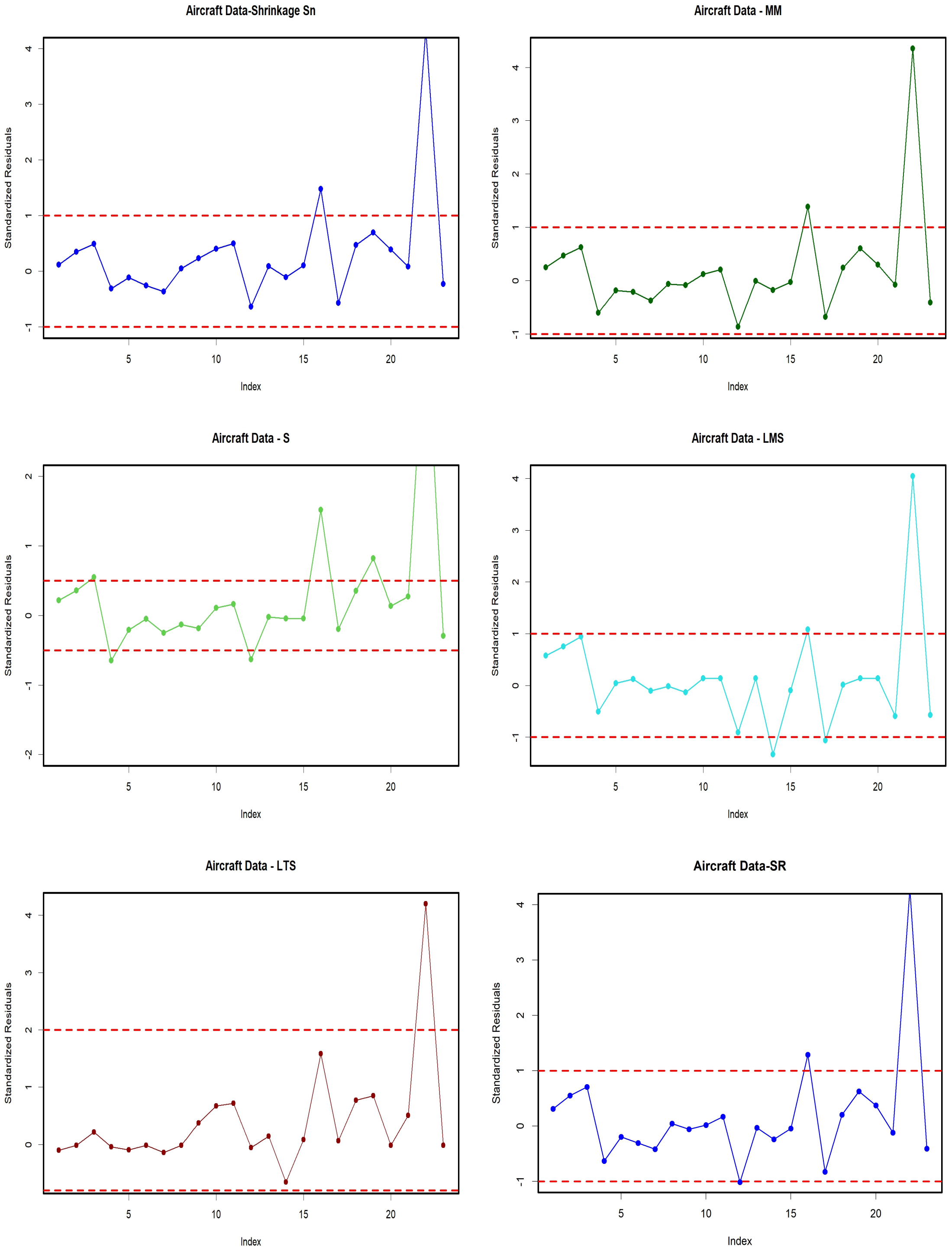
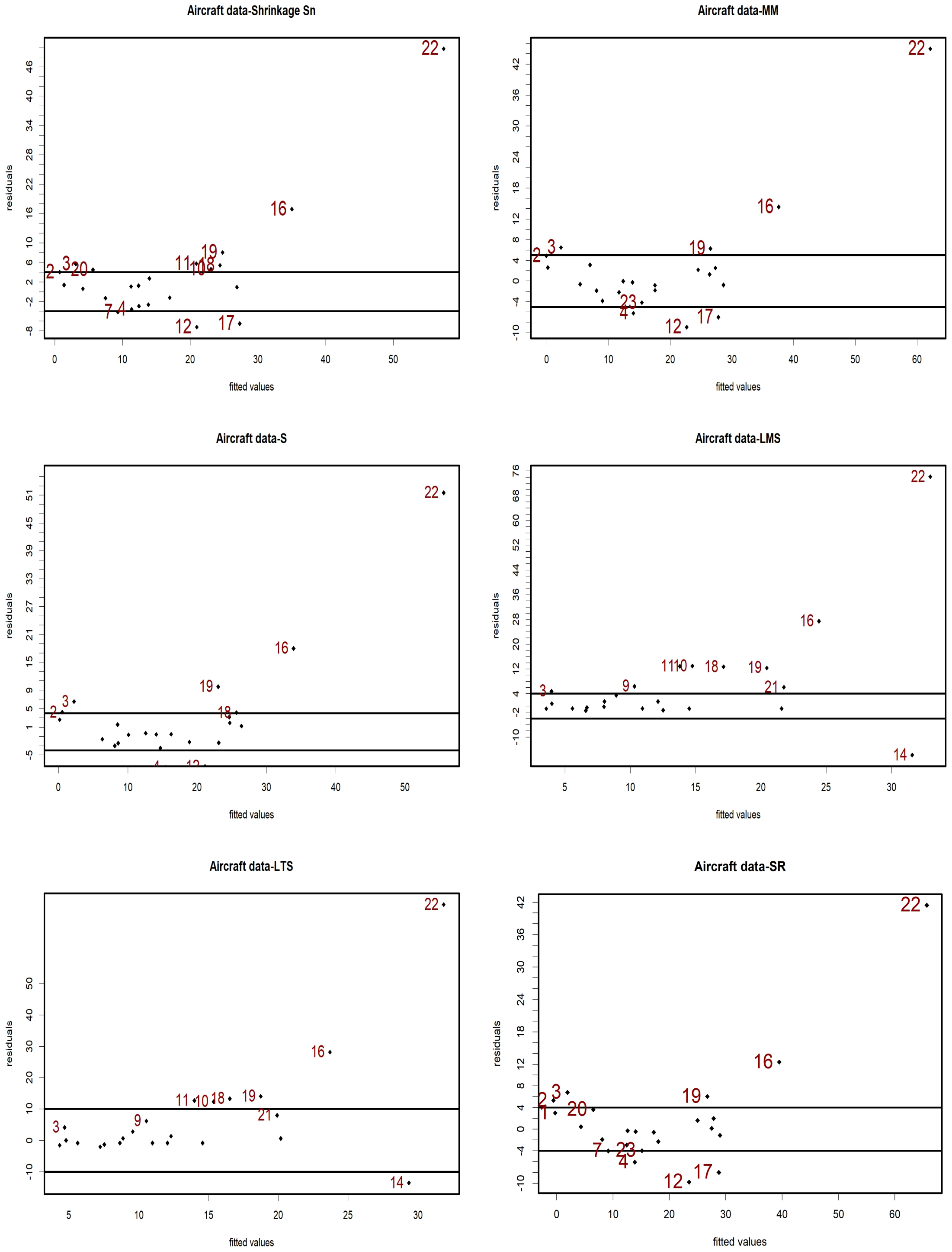


Figure 3.8: Aircraft data Standardized Residuals Index Plot for outlier detection of $RSh - S_n$, MM, S, LMS, LTS, SR



Aircraft data-Shrinkage Sn

Aircraft data-MM

Aircraft data-S

Aircraft data-LMS

Aircraft data-LTS

Aircraft data-SR

Figure 3.9: Aircraft data Fitted values versus Residuals Plot of $RSh - S_n$, MM, S, LMS, LTS, SR

Table 3.9: Belgium Phone Call data

Method	Intercept	Slope	RSE	MSE	MAPE	MAD	AIC
$RSh-S_n$	-5.0119	0.07490	5.7419	37.969	2.7439	4.9275	95.28
MM	-5.2423	0.11009	6.5926	48.462	0.30085	3.5316	101.12
LTS	-5.6162	0.1159	6.5856	48.3706	0.31328	3.5306	101.09
LMS	-5.5947	0.1155	6.5879	48.3989	0.31277	3.5316	101.11
S	-5.2732	0.1102	6.6042	48.6151	0.30198	3.5384	101.21
OLS	-26.006	0.5041	4.8966	28.9765	1.5236	4.2453	88.796
SR	-26.008	0.5042	25.4240	511.7187	10.718	17.0290	157.707

3.7.4 Belgium Phone Call data

The Belgium phone call data was originally published by the Belgium Statistical Survey and was utilized in the book of P. J. Rousseeuw & Leroy (2005). This dataset includes the annual number of international calls made from Belgium between 1950 and 1973. The dataset consists of two variables: the year ($\mathbf{x}$) and the number of calls received (y). The data represents simple linear regression data. But when a person without critical thinking looked into the residual plot of OLS, they couldn't find anything abnormal. If we consider the residual plots (figure: 3.11) from robust estimators and fitted vs. residual plot (figure: 3.12), we can readily identify 6 outliers in the y direction, and we can observe how these observations influence other observations by comparing the residual plots of the classical vs. robust estimator. RSh-S_n estimator based plot detects observations 15, 16, 17, 18, 19, 20 correctly as outliers along with other robust estimators . Due to the influence of these six outliers, observation 21 too, was detected as an outlier. The regression coefficients and other model metrics values are tabulated and given below.

The tabulated metrics of different real - life datasets given in the table 3.9 show that RSh-S_n estimator is capable of withstand outlier occupying datasets and gave metric values less than LTS, LMS, and S. Also, the regression coefficients similar to those of other robust estimators. The plots show the capability of RSh-S_n estimator to detect the outliers. All this gives assurance regarding the good performance of RSh-S_n estimator.

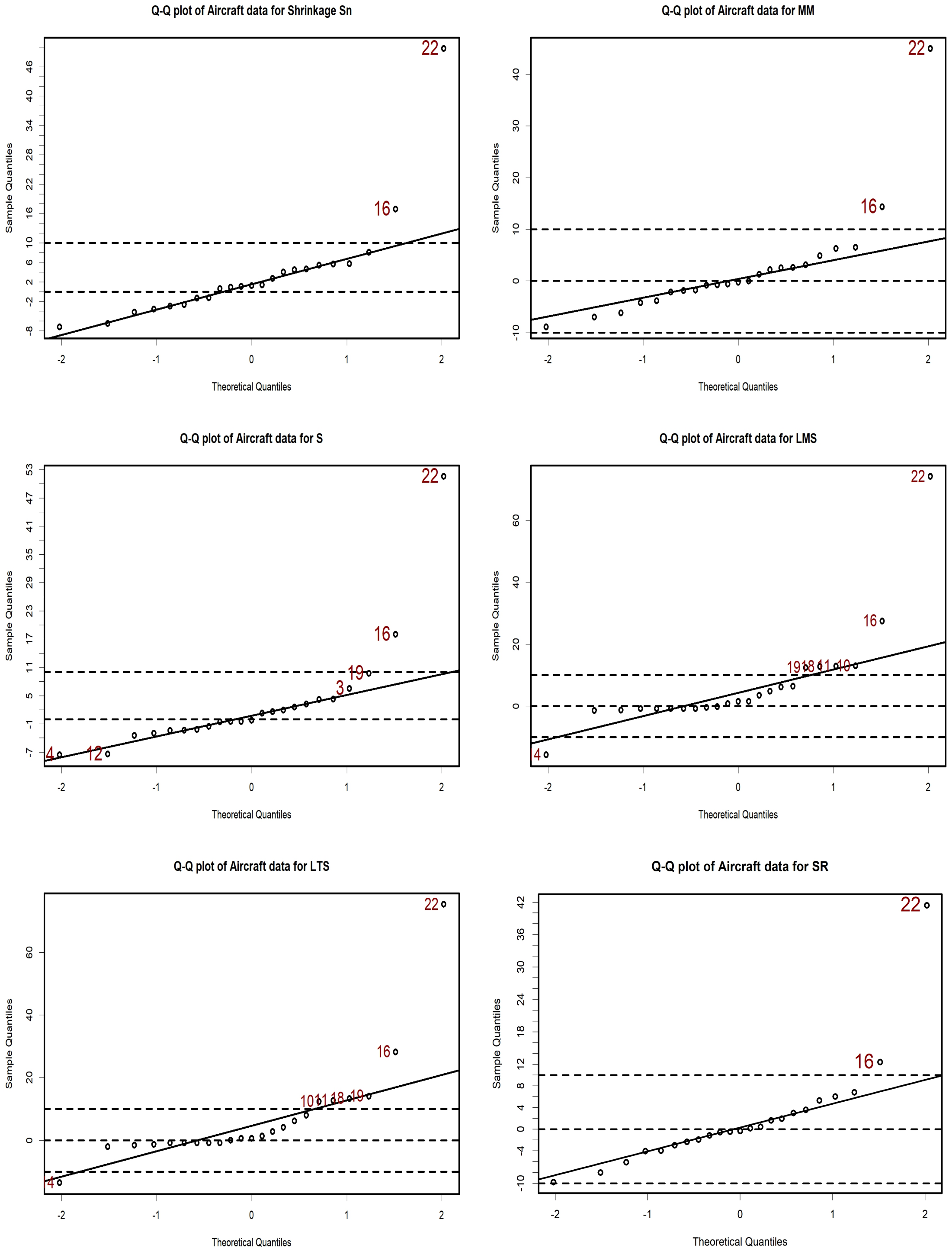


Figure 3.10: Aircraft data Q-Q plot of Residuals of $RSh - S_n$, MM, S, LMS, LTS, SR

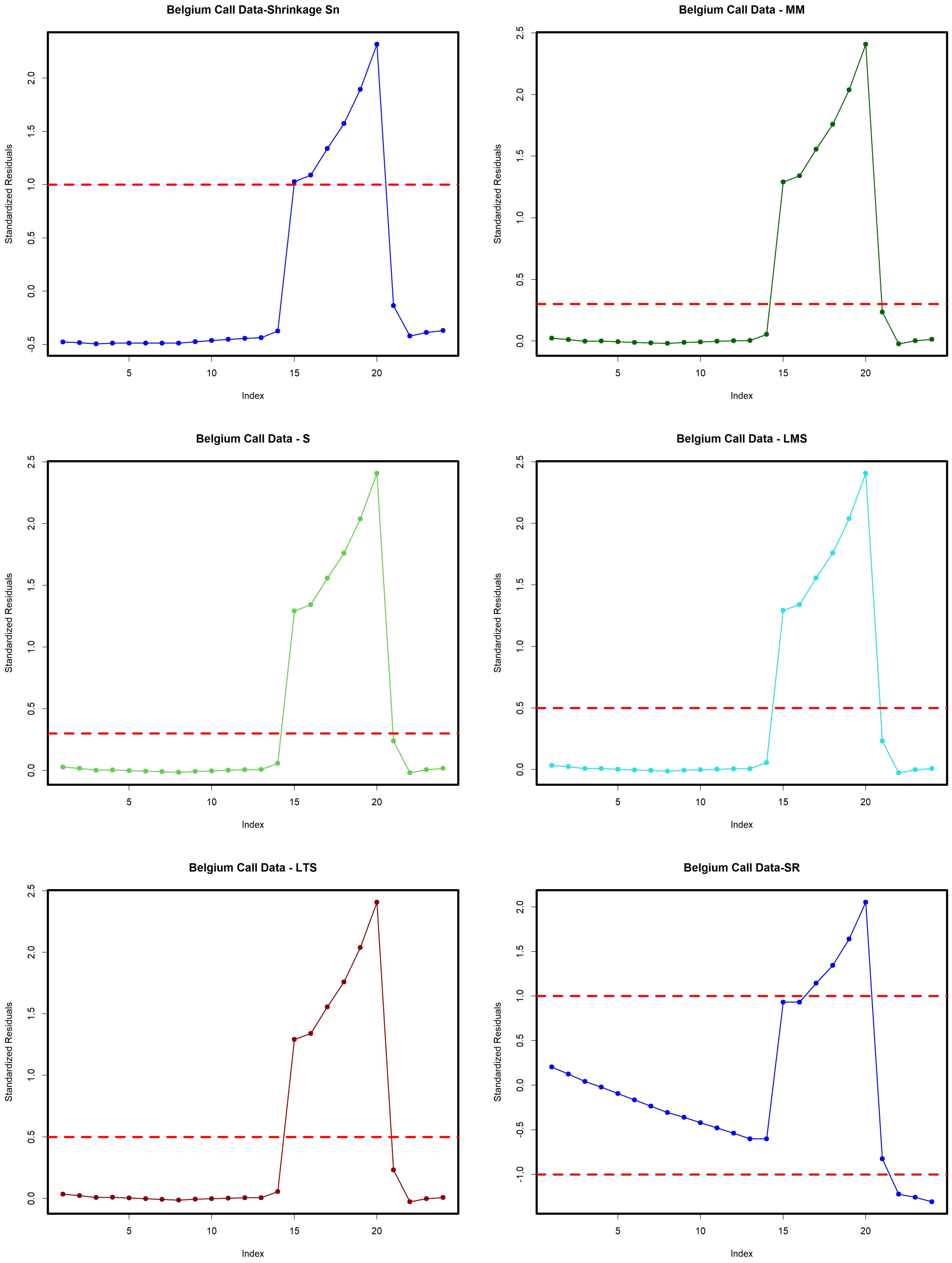
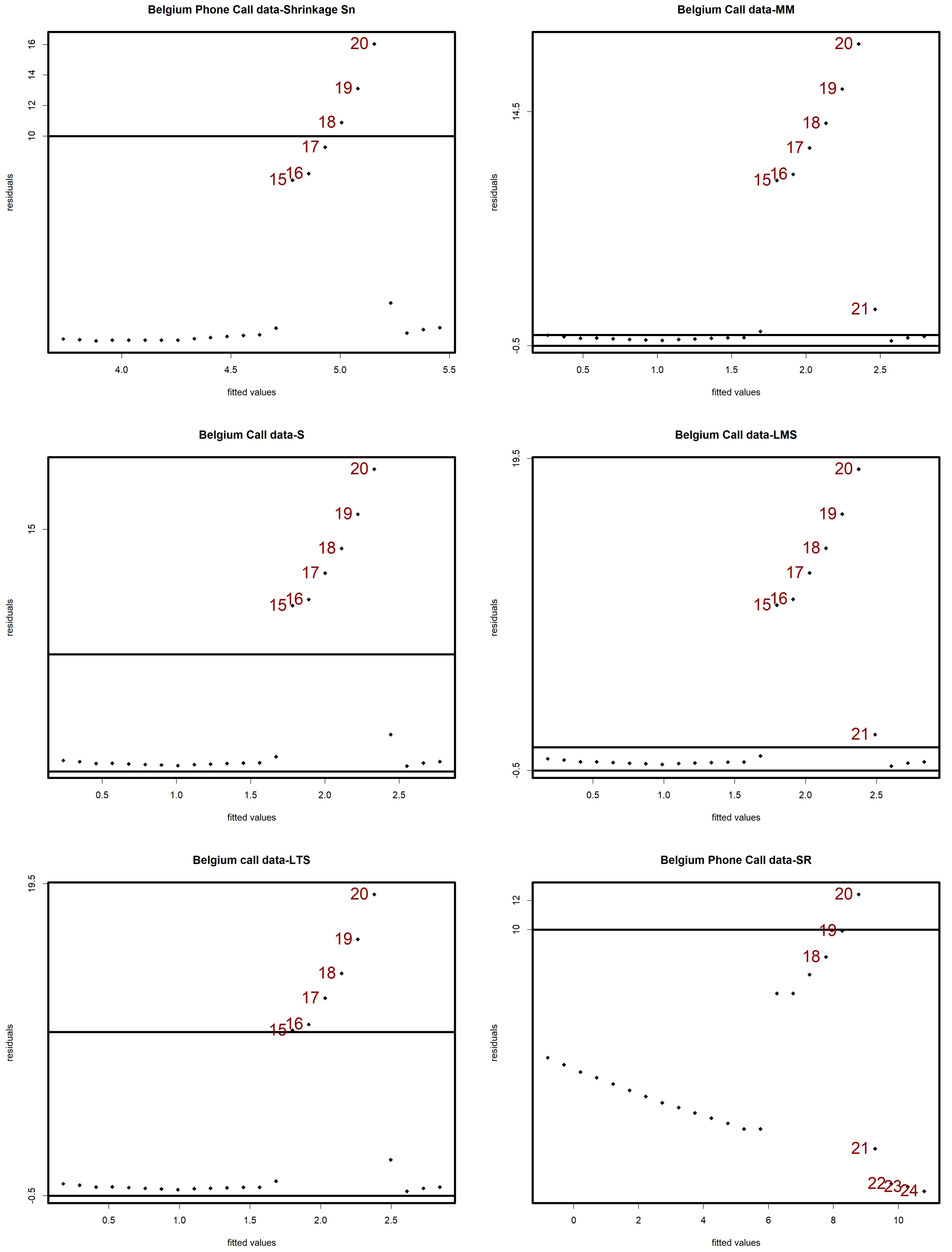


Figure 3.11: Belgium Phone Call data Standardized Residuals Index Plot for outlier detection of $RSh - S_n$, MM, S, LMS, LTS, SR



Belgium Phone Call data-Shrinkage Sn

Belgium Call data-MM

Belgium Call data-S

Belgium Call data-LMS

Belgium call data-LTS

Belgium Phone Call data-SR

Figure 3.12: Belgium Phone Call data fitted values versus residuals plot of $RSh{-}S_n$, MM, S, LMS, LTS, SR

3.8 Summary

In multiple regression, the response variable is related to p explanatory variables. Commonly, OLS regression is used for the estimation of regression coefficients. The outlier presence in the dataset makes OLS produce distorted regression coefficients. Also, outliers can increase the standard errors of the regression coefficients, making them appear less statistically significant and misleading goodness of fit. So while doing regression, it falls on the researcher to carefully look into each aspect rather than simply looking into the regression coefficient estimates. One should do a thorough examination of the model assumptions, the quality of the data, potential outliers, and the meaningfulness of the results. Careful observation, diagnostic testing, and domain expertise are essential to building a reliable and valid regression model. In this chapter we have proposed a robust regression estimator RSh-S_n based on Sh-S_n covariance estimator. Our extensive comparisons demonstrate that $RSh - S_n$ outperforms other existing robust regression methods. It is crucial to recognize that many available methods fail to deliver satisfactory performance with high-dimensional data. Not all available methods perform well with large datasets or high-dimensional data, and many are not proven to be sufficiently robust against the presence of outliers. We propose to make use of robust shrinkage estimators of location and scatter to estimate the regression parameters by utilizing the concept of shrinkage. The method produces the $RSh - S_n$ regression estimator. The advantages of the proposed estimator is shown through the simulation study. Simulation study demonstrates that our proposed estimator consistently outperforms other methods in terms of robustness, regardless of higher dimensions, larger contamination, or transformed data. In terms of efficiency, our estimator shows an advantage over existing approaches like LTS, LMS, S, and MM. A key feature of $RSh - S_n$ is that it utilizes all observations, unlike the sub-sample iterations employed in other methods, which adds to its stability. We also applied the new estimator to several real-life datasets and evaluated its performance. The results favor the effectiveness of our proposed method, in line with the findings from the simulation study. Comparisons of real-life models highlight the capabilities of our estimator and reveal how classical OLS can mislead researchers when dealing with datasets containing outliers.

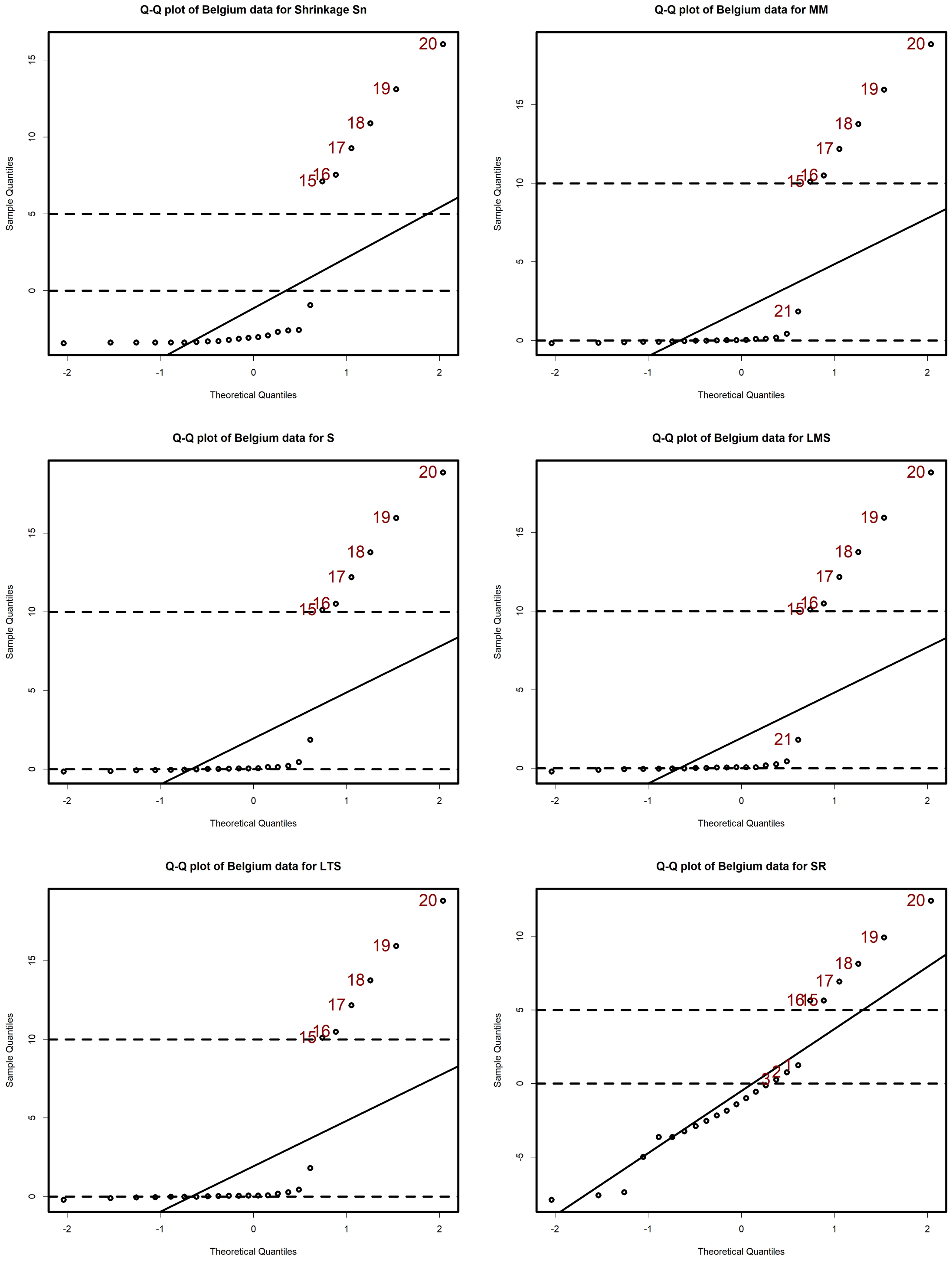


Figure 3.13: Belgium Phone Call data Q-Q plot of Residuals of $RSh - S_n$, MM, S, LMS, LTS, SR

Chapter 4

Robust multivariate RSh-S_η regression estimator

4.1 Introduction

In statistical modeling, regression analysis is a method used to estimate the relationships between variables. It involves a range of techniques for modeling and analyzing multiple variables. As discussed, the multivariate regression model allows us to examine the influence of several variables on one or more dependent variables within the same model.

Let $\mathbf{x} = (x_1, x_2, \ldots, x_p)^t$ be a p-dimensional predictor and a q-dimensional response $\mathbf{y} = (y_1, y_2, \ldots, y_p)^t$. Consider a multivariate regression model $\mathbf{y} = \mathbf{B}^t\mathbf{x} + \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$, where $\mathbf{B}$ is $p \times q$ slope matrix, $\boldsymbol{\alpha}$ is q-dimensional intercept vector, and $\boldsymbol{\varepsilon} = (\varepsilon_1, .., \varepsilon_q)^t$ denotes the $i.i.d.$ error term with mean zero and cov($\boldsymbol{\varepsilon}$)=$\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}$, a $q \times q$ positive definite matrix. Let $\boldsymbol{\mu}$ denote the location of joint variables $(\mathbf{x}, \mathbf{y})$ and $\boldsymbol{\Sigma}$ denotes their scatter matrix. Partition matrix of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ be as follows: $\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_{\mathbf{x}} \\ \boldsymbol{\mu}_{\mathbf{y}} \end{pmatrix}$ and $\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{\mathbf{xx}} & \boldsymbol{\Sigma}_{\mathbf{xy}} \\ \boldsymbol{\Sigma}_{\mathbf{yx}} & \boldsymbol{\Sigma}_{\mathbf{yy}} \end{pmatrix}$. Conventional maximum likelihood estimates of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, empirical mean $\hat{\boldsymbol{\mu}}$ and empirical covariance matrix $\hat{\boldsymbol{\Sigma}}$ are often used as estimates. The components of the resulting estimators $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ are utilized in the least squares equations (Johnson and Wichern (1988, p.301)) as follows:

$$\hat{\mathbf{B}} = \hat{\mathbf{\Sigma}}_{\mathbf{xx}}^{-1}\hat{\mathbf{\Sigma}}_{\mathbf{xy}}, \tag{4.1.1}$$

$$\hat{\boldsymbol{\alpha}} = \hat{\boldsymbol{\mu}}_{\mathbf{y}} - \hat{\mathbf{B}}^t\hat{\boldsymbol{\mu}}_{\mathbf{x}}, \tag{4.1.2}$$

$$\text{and} \quad \hat{\mathbf{\Sigma}}_{\varepsilon} = \hat{\mathbf{\Sigma}}_{\mathbf{yy}} - \hat{\mathbf{B}}^t\hat{\mathbf{\Sigma}}_{\mathbf{xx}}\hat{\mathbf{B}}. \tag{4.1.3}$$

It is widely recognized that, the classical multiple regression is highly sensitive to outliers in the data. This issue is equally prevalent in the context of multivariate regression, since classical $\hat{\mu}$ and $\hat{\Sigma}$ are sensitive in the presence of anomalies. To address this issue, one can substitute the classical estimates of location and scatter with highly robust estimates that are less sensitive to outliers, enabling a more robust analysis. R. A. Maronna & Yohai (1997) provide an overview of robust multivariate regression algorithms in the context of simultaneous equations models. Although, Koenker & Portnoy (1990) presented a M - type approach, their estimator lacked affine equivariance. P. J. Rousseeuw et al. (2004) proposed estimator based on the robust estimate of the location and dispersion of the joint distribution of the $(\mathbf{x}, \mathbf{y})$ variables using Minimum Covariance Determinant, which is computationally time - consuming and expensive. Ollila et al. (2002) and Ollila et al. (2003) introduced multivariate regression based on rank covariance matrices and evaluated their properties in their paper. A multivariate regression extension of the least trimmed squares estimator (MLTS) was investigated by Agulló et al. (2002), in which the slope matrix is obtained as the minimum of the determinant of the robust MCD scatter matrix of the residuals. Van Aelst & Willems (2005) introduced S estimators in multivariate regression similar to that of MLTS. Both LTS and S based multivariate regression estimators perform with OLS estimates as the initial fit. Sajana & Sajesh (2018) introduced a multivariate regression estimator based on Comedian covariance and empirically developed the properties of the proposed estimator; also checked the performance through a simulation study. The estimator performs better in terms of robustness and efficiency with respect to other estimators compared in their paper.

In this chapter, we propose to make use of reweighted, Shrinkage based S_n covariance estimator $(Sh - S_n)$ and Shrinkage - based L_1 median as robust alternatives to the classical estimates of $\mathbf{\Sigma}$ and $\boldsymbol{\mu}$ respectively, which we developed in the second chapter, multivariate regression estimator. We make use of the same estimator here and compare it with classical maximum likelihood estimates based multivariate regression, P. J. Rousseeuw et al. (2004), orthogonalized Gnanadesikan - Kettenring (OGK) estimates of R. A. Maronna & Zamar (2002) based multivariate regression estimates, Sajana & Sajesh (2018) and Kunjunni & Abraham (2022) based multivariate regression estimates. The next section of this chapter describes our proposed regression estimator, and the following section shows the simulation study on robustness and efficiency properties. We also prove the equivariance property empirically. We have presented two real-world examples and concluded the chapter with a summary.

4.2 $RSh - S_n$ multivariate regression estimator

The principle behind shrinkage estimation lies in the notion of "shrinking" an estimator $\hat{E}$ towards a target estimator $\hat{T}$, which serves to effectively diminish estimation errors. In chapter 2, we have introduced Sh-S_n covariance estimator and Shrinkage L_1 median location. We have shown some properties of our proposed estimators there. Here we make use of those estimators to develop multivariate regression estimators.

Consider $\mathbf{z} = (\mathbf{x},\, \mathbf{y})$, the joint variable with location and covariance matrix μ, Σ respectively. The associated squared Mahalanobis distance for each observation $\mathbf{z}_i, i = 1, 2, 3, \ldots, n$, based on initial estimates of Shrinkage L_1 median $\hat{\mu}_{Sh}$ and $Sh - S_n$ covariance estimate $\hat{\Sigma}_{Sh}$ be:

$$\mathrm{RD}^2(\mathbf{z}_i) = (\mathbf{z}_i - \hat{\mu}_{Sh})^t \hat{\Sigma}_{Sh}^{-1} (\mathbf{z}_i - \hat{\mu}_{Sh}). \tag{4.2.1}$$

The weight function based on above defined robust Mahalanobis distance be $w_i = w(\mathrm{RD}^2(\mathbf{z}_i))$, where a weight of 1 be assigned to the observations $(\mathbf{z}_i)$ with a Mahalanobis distance less than $\frac{\chi^2_{0.95,p} \times \mathrm{median}(\mathrm{RD}^2(\mathbf{z}_i))}{\chi^2_{0.5,p}}$. Thus the reweighted shrinkage location and S_n covariance matrix be defined as:

$$\hat{\boldsymbol{\mu}}^{\mathbf{1}} = \frac{\sum_{i=1}^n w_i \mathbf{z}_i}{\sum_{i=1}^n w_i}, \quad \hat{\boldsymbol{\Sigma}}^{\mathbf{1}} = \frac{\sum_{i=1}^n w_i (\mathbf{z}_i - \hat{\boldsymbol{\mu}}^{\mathbf{1}})(\mathbf{z}_i - \hat{\boldsymbol{\mu}}^{\mathbf{1}})^t}{\sum_{i=1}^n w_i}. \tag{4.2.2}$$

We now identify the regression estimates (WR) from 4.1.1 based on the above defined reweighted location and scatter estimators. It makes sense to employ weights in regression analyses that are based on the residuals that correspond to the original fit P. J. Rousseeuw & Leroy (2005). Let the residuals of the weighted regression estimates (WR) based on above defined reweighted estimators be:

$$\boldsymbol{r}_i = \mathbf{y_i} - \hat{\mathbf{B}}^t \mathbf{x_i} - \hat{\boldsymbol{\alpha}}. \tag{4.2.3}$$

Now we again reweight the WR estimator based on the residuals defined above. Consider the weights $w\mathbf{r_i}$=$w(\mathrm{RD}^2(\mathbf{r_i}))$ which assign a weight 1 to the residuals($\mathbf{r_i}$) with Mahalanobis distance less than $\chi^2_{q,0.99}$, where $\mathrm{RD}^2(\mathbf{r_i})$ be the Mahalanobis distance of WR residuals, defined as $\mathrm{RD}^2(\mathbf{r_i}) = \mathbf{r_i}^t(\hat{\boldsymbol{\Sigma}}_\varepsilon)^{-1}\mathbf{r_i}$. The final reweighted regression estimators be:

$$\mathbf{T}^R = (\sum_{i=1}^{n} w\mathbf{r_i}\mathbf{u}_i\mathbf{u}_i^t)^{-1}\sum_{i=1}^{n} w\mathbf{r_i}\mathbf{y_i}\mathbf{u}_i, \tag{4.2.4}$$

and

$$\hat{\boldsymbol{\Sigma}}_\varepsilon{}^R = \frac{\sum_{i=1}^n w\mathbf{r_i}(\mathbf{r}^R)_i(\mathbf{r}^R)_i^t}{\sum_{i=1}^n w\mathbf{r_i}}, \tag{4.2.5}$$

where $\mathbf{T}^R = ((\hat{\mathbf{B}}^R)^t, \hat{\boldsymbol{\alpha}}^R)^t$, $\mathbf{u}_i$=$(\mathbf{x}_i^t, 1)^t$ and $(\boldsymbol{r}^R)_i = \mathbf{y}_i - (\hat{\mathbf{B}}^R)^t\mathbf{x_i} - \hat{\boldsymbol{\alpha}}^R$. $\mathbf{T}^R$ be the two step reweighted shrinkage regression estimator (RSh-S_n). The superscript R implies the weights were based on initial regression. The robustness of these reweighted regression estimators is derived from the properties of the initial regression estimators. It is important to note that the weights now depend solely on the magnitude of the residual distance $w\mathbf{r_i}$. Unlike the initial estimates, good leverage points are no longer down weighted.

4.2.1 Pulpfibre data

Consider dataset, Lee et al. (1993) that includes measurements of various properties of pulp fibers as well as the characteristics of the paper produced from these fibers. The objective is to explore the relationships between the properties of pulp fibers and the characteristics of the resulting paper. The dataset comprises $n = 62$ measurements of four specific pulp fiber attributes: arithmetic fiber length, long fiber fraction, fine fiber fraction, and zero span tensile. The dataset includes measurements of four paper properties: breaking length, elastic modulus, stress at failure, and burst strength. Our objective is to predict the four paper properties using the four fiber characteristics. To achieve this, we initially applied classical multivariate regression to the dataset. Figure 4.1 shows diagnostic plot of classical estimator.

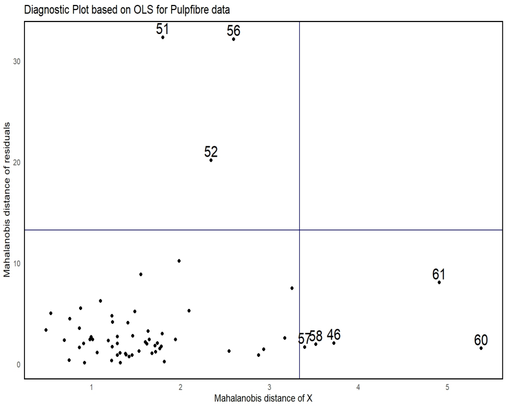


Figure 4.1: OLS Diagnostic Plot for Pulpfibre data

This diagnostic plot integrates information on regression outliers and leverage points, since it is based on Mahalanobis distance of the residuals versus Mahalanaobis distance of independent variables, providing a more comprehensive view than analyzing each distance individually. The horizontal and vertical line cut off values are $\sqrt{\chi^2_{p,0.975}}$, $\sqrt{\chi^2_{q,0.99}}$ respectively. For classical estimator, we can see that observations 51, 52, and 56 are identified as vertical outliers. Conversely, while some observations are recognized as leverage points (with observations 60 and 61 being the most prominent), they are not classified as regression outliers. To validate the results from classical multivariate regression, Rousseeuw applied univariate robust LTS regression to each response individually, using the same regressors. The results

Table 4.1: Outliers detected in Pulpfibre data using LTS regression as in P. J. Rousseeuw et al. (2004)

Response	Outliers
y_1	51,52,56,61
y_2	61
y_3	52,56,61
y_4	51,52,56

in 4.1 indicate that univariate LTS regression identify observations 51, 52, 56, and 61 as outliers. The results of LTS regression shows the necessity of finding outliers and substantiate the results using robust estimators. However, using univariate robust estimators on each response variable individually does detects outliers only

along the coordinate directions of the responses and fails to identify outliers that may be masked within these directions. In general, it is preferable to use a robust multivariate regression estimator that can detect all outliers and is efficient both statistically and computationally. Diagnostic plot allows us to categorize data points as regular observations, vertical outliers, good leverage points, or bad leverage points. Additionally, it helps us determine whether a point is an extreme outlier or just a borderline case. This analysis gave insight on the need and usefulness of a robust estimator and as an inspiration we develop two step reweighted regression estimator. We study the efficiency, robustness of our robust estimator in upcoming sections. Also, we evaluated the affine equivariance property too.

4.3 Efficiency

To assess the efficiency of the proposed regression estimator, we conducted the following simulation study similar to that in P. J. Rousseeuw et al. (2004). For different sample sizes n and various choices of p and q, we generated r datasets of size n from the multivariate standard Gaussian distribution $N(\mathbf{0}, \mathbf{I}_{p+q})$, where $\mathbf{B} = 0$ and $\boldsymbol{\alpha} = 0$. For each dataset ($l = 1, 2, \ldots, r$), we conduct the reweighted regression estimation as defined in section 2 based on shrinkage estimators, yielding $p \times q$ slope estimate $\hat{\mathbf{B}}^{(l)}$, intercept estimate $\hat{\boldsymbol{\alpha}}^{(l)}$ and the covariance matrix estimate $\hat{\boldsymbol{\Sigma}}_{\varepsilon}^{(l)}$ of the errors.

The variance estimate of slope coefficient is obtained as:

$$\operatorname{var}(\hat{\mathbf{B}}_{jk}) = n\operatorname{var}(\hat{\mathbf{B}}_{jk}^{(l)}), \text{where} \quad j = 1, 2, \ldots, p \quad \text{and} \quad k = 1, 2, \ldots, q. \tag{4.3.1}$$

Then the corresponding finite sample efficiency of the slope estimate is defined as $1/\text{ave}_{j,k}(\operatorname{var}(\hat{\mathbf{B}}_{jk}))$. Similarly, we calculate the finite-sample efficiency of the intercept vector. To assess the accuracy of the error scatter matrix, we use the standardized variance (Bickel & Lehmann (2011)) of the elements of the error covariance matrix, defined as follows:

$$\text{Stvar}\left(\hat{\boldsymbol{\Sigma}}_{\varepsilon jk}\right) = \frac{n\operatorname{var}_l((\hat{\boldsymbol{\Sigma}}_{\varepsilon}^{(l)})_{jk})}{[\text{ave}_l \text{ave}_j((\hat{\boldsymbol{\Sigma}}_{\varepsilon}^{(l)})_{jj})]^2} \text{for} \quad j = 1, \ldots, q \quad \text{and} \quad k = 1, \ldots, q.$$

The overall finite - sample efficiency of the off-diagonal elements is then defined as follows $1/\text{ave}_{j\neq k}(\text{Stvar}((\hat{\boldsymbol{\Sigma}}_\varepsilon)_{jk}))$. And the finite sample efficiency of diagonal elements is given by $2/\text{ave}_j(\text{Stvar}((\hat{\boldsymbol{\Sigma}}_\varepsilon)_{jj}))$.

Table 4.5 and tables in appendix B gives the result of our study. The results of RSh-S_n estimator along with other estimators shown in the tables. We have considered sample sizes n = 50, 100, 200, 500. All the simulated values are obtained after 1000 replications. The (p, q) dimensions considered throughout the study are (4,4), (4,8), (8,4), (10,10). From the tables it is clear that proposed regression estimator shows better performance in uncontaminated dataset. The table value shows that irrespective of the dimension, two step reweighted RSh-S_n estimator possess better efficiency than other robust estimators. Undoubtedly traditional least squares based on classical MLE will be the consistent and asymptotically efficient in an uncontaminated data. Thus, RSh-S_n is better than other robust estimators and not possessing over estimation issue.

4.4 Robustness

We performed simulations similar in P. J. Rousseeuw et al. (2004), to investigate the finite-sample robustness with dataset containing outliers. A vertical outlier is a point $(\mathbf{x}_i, \mathbf{y}_i)$ whose $\mathbf{x}_i$ is not outlying but does not follow the linear trend of the majority of the data. A point $(\mathbf{x}_i, \mathbf{y}_i)$ where $\mathbf{x}_i$ is an outlier is referred to as a leverage point. If this $(\mathbf{x}_i, \mathbf{y}_i)$ deviates from the pattern of the majority, it is termed a bad leverage point. Conversely, if it aligns with the majority pattern and does not negatively impact the fit, it is considered a good leverage point. Regression estimators often fail when confronted with vertical outliers or bad leverage points. In this study, we created datasets that include both types of outliers for the evaluation of estimators. For a given sample size n, we generate $r = 1000$ datasets from multivariate standard Normal distribution with mean $\mathbf{0}$ and identity matrix $\mathbf{I}_{p+q}$ as variance covariance matrix. For the evaluation purpose, we consider 10%, 20% and 40% contamination in datasets. Keeping $\mathbf{x}_i$, the q response variables are taken from multivariate normal distribution $\text{N}(\mathbf{2}\sqrt{\boldsymbol{\chi}^2_{\boldsymbol{p+q,0.99}}}, \mathbf{I}_q)$. This produces vertical outliers, because here only response variables are outlying. We also replaced the data with bad leverage points for which the p independent variables are generated according to multivariate normal distribution $\text{N}(\mathbf{2}\sqrt{\boldsymbol{\chi}^2_{\boldsymbol{p,0.99}}}, \mathbf{I}_p)$ and the q dependent variables are generated from multivariate normal N$(\mathbf{2}\sqrt{\boldsymbol{\chi}^{\mathbf{2}}_{\boldsymbol{q},\mathbf{0.99}}}, \mathbf{I}_q)$. Apart from the mentioned distribution, we also considered observations from multivariate normal N$(\mathbf{2}\sqrt{\boldsymbol{\chi}^{\mathbf{2}}_{\boldsymbol{p+q},\mathbf{0.99}}}, 0.1 \times \mathbf{I}_q)$, N$(\mathbf{2}\sqrt{\boldsymbol{\chi}^{\mathbf{2}}_{\boldsymbol{q},\mathbf{0.99}}}, 0.1 \times \mathbf{I}_q)$, N$(\mathbf{2}\sqrt{\boldsymbol{\chi}^{\mathbf{2}}_{\boldsymbol{p},\mathbf{0.99}}}, 0.1 \times \mathbf{I}_p)$ respectively for vertical outliers and bad leverage points. Rather than using observations from different distributions, which would make it easier for the estimators to detect outliers and perform, we opted to use the same multivariate normal distribution with varying parameters. As in efficiency study, we generate each dataset ($l = 1, 2, \ldots, r$), compute the slope matrix estimate $\hat{\mathbf{B}}^{(l)}$, the intercept estimate $\hat{\boldsymbol{\alpha}}^{(l)}$ and the covariance matrix estimate $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\varepsilon}}^{(l)}$ of the errors. In order to measure robustness, we make use of bias and MSE. As in the case of univariate component, the bias and MSE of the slope are defined as:

$$\text{bias}(\hat{\mathbf{B}}) = \sqrt{\text{ave}_{j,k}(\text{bias}(\hat{\mathbf{B}}_{jk})^2)}$$

and

$$\text{MSE}(\hat{\mathbf{B}}) = \text{ave}_{j,k}(\text{MSE}(\hat{\mathbf{B}}_{jk})).$$

Similarly bias, MSE for the intercept $\hat{\boldsymbol{\alpha}}$, for diagonal and off diagonal elements of $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\varepsilon}}$ are calculated. For study, we have considered n = 50, 100, 200 and 300 and (p, q) = (4,4), (4,8), (8,4), (10,10). Tables B.1 - B.34 show the results of robustness. For comparison purpose we have considered the metrics MSE and Bias. The results shows that RSh-S_n possesses smaller MSE and bias than other robust estimators across all contamination considered. We observed other estimators fails to perform as level of contamination increases with a decline in MSE and bias. Proposed RSh-S_n show smaller MSE and bias irrespective of the contamination in data. The contaminants have multivariate normal with 0.1 deviation in covariance matrix and the results of such small deviated multivariate normal are in favorable to RSh-S_n estimator. MSE, bias of RSh-S_n is least compared to other estimators in deviated contaminants too, irrespective of higher contamination and dimension.

Generalized approaches to regression, scale, affine equivariance, and the robustness of multiple regression estimators were introduced by P. J. Rousseeuw & Leroy (2005). Regression equivariance implies that if a linear function of the explanatory variables is added to the responses, the coefficients of that linear function are similarly added to the estimator. The $\mathbf{y}$ - equivariance of the estimator means that a linear transformation of the response variables results in the estimator being transformed in the same way. **x** - equivariance indicates that if the predictor variables undergo a linear transformation, the estimator will transform correspondingly. The three equivariance properties are proven for our proposed estimator empirically by considering contaminated scenarios same as for checking robustness, additionally 0% contamination too considered. MSE values of slope and intercept from the simulated samples are considered for the evaluation. Table B.36 - B.35 shows the results of affine equivariance. Sample size and level of contamination considered are same as that in robustness and efficiency above. The results shows that RSh-S_n possess least, almost negligible MSE for affine equivariance. Irrespective of the contamination level, dimension, our proposed estimator exhibit least MSE, gave us assurance on equivariance property of RSh-S_n empirically.

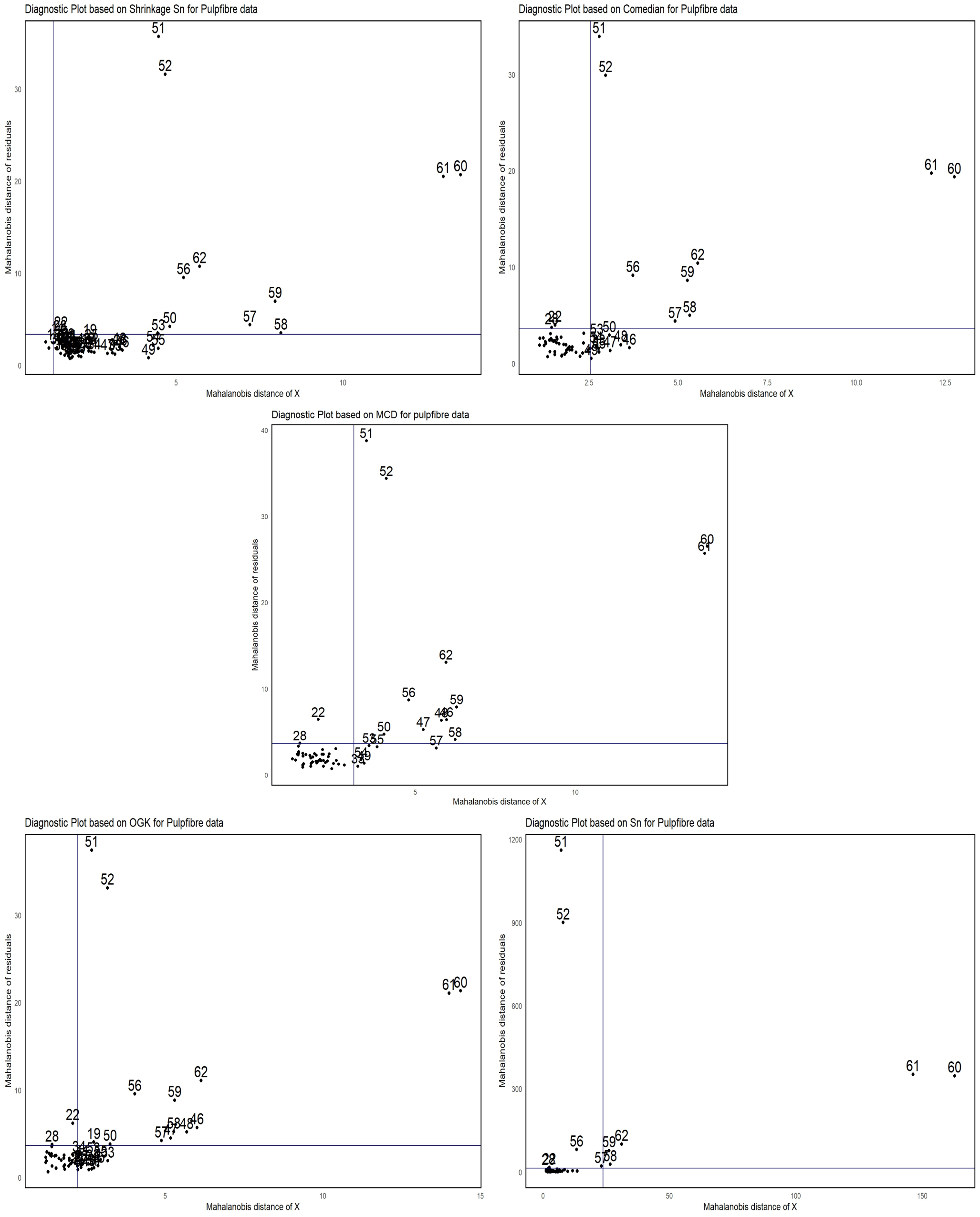
4.5 Real - Life examples

4.5.1 Pulpfibre data continuation

Figure 4.2 shows the diagnostic plot of robust RSh-S_n estimator along with other compared estimators. For plotting we consider the horizontal cutoff line as the cutoff of respective used Mahalanobis distance and vertical cut off as $\sqrt{\chi^2_{q,0.99}}$. The choice of the vertical line cutoff is usually between $\sqrt{\chi^2_{q,0.975}}$ and $\sqrt{\chi^2_{q,0.99}}$. Here we selected the latter as our cutoff point. Rousseeuw in his paper explored the sample and found the observation (59 - 62) were produced from fir wood. Most of the outlying observations were obtained from different pulping process. Observation 62 was obtained from the chemi - thermo mechanical process, while observations 22, 46 - 48, 58 - 61 were obtained from different pulping processes. Evidently from the diagnostic plot, we can affirm that proposed regression estimator detect all these observations as outliers which validate the efficacy of our estimator.

Table 4.2: Pulpfibre Data Table

Estimators	Slope MSE	Intercept MSE	Slope Bias	Intercept Bias
MCD	561.7581	2158.558	9.37032	-40.2621
OGK	522.1857	2003.755	9.020945	-38.7037
Comedian	634.0527	2230.64	10.72183	-40.7999
S_n	634.0527	2230.64	10.72183	-40.7999
$RSh - S_n$	534.7054	2076.603	10.58657	-43.9223
MLE	847.0833	2786.09	12.30624	-45.6509


Diagnostic Plot based on Shrinkage Sn for Pulpfibre data

Diagnostic Plot based on Comedian for Pulpfibre data

Diagnostic Plot based on MCD for pulpfibre data

Diagnostic Plot based on OGK for Pulpfibre data

Diagnostic Plot based on Sn for Pulpfibre data

Figure 4.2: Diagnostic plot based on different estimators for Pulpfibre data

4.5.2 School data

This data consists of $n = 70$ observations on school sites in US (Charnes et al. (1981); Roelant et al. (2009)). Data contains three response variable and 5 explanatory variables. The response variables are: total reading score measured by the Metropolitan Achievement test, total mathematics score measured by the Metropolitan Achievement test and the Coopersmith self - esteem inventory. The independent variables are: education level of mother, highest occupation of a family member, number of parental visits to the school, parent counseling concerning school - related topics and the number of teachers at the school. The diagnostic plots are given in figure 4.3. Cutoff lines are considered same as above mentioned. Classical estimator detects only one observation as vertical outliers and six observations as good leverage points. But robust estimators detect the observations 33, 35, 44, 59 as outliers, more than two moderate large good leverage points (1, 10, 50, 54, 66, 67. . .). Along with other robust estimators, our proposed estimator detects the aforementioned observations as susceptible which shows efficiency of our estimator in real - life datasets. Thus, it is always better to make use of robust estimator and RSh-S_n performs well in detecting the vertical outliers, leverage points.

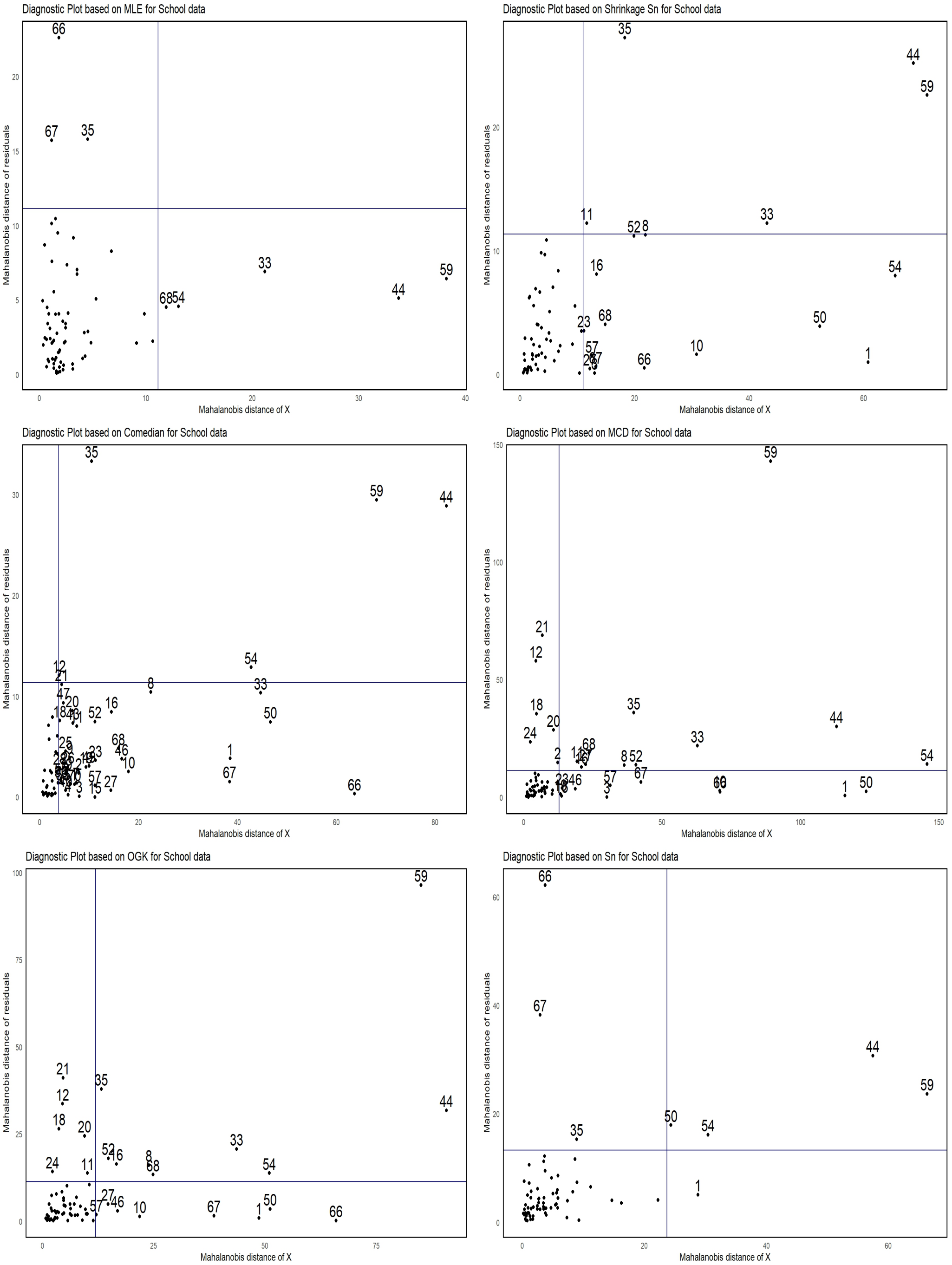
Also, we have tabulated the intercept, slope MSE, bias of these two examples and are given below. The table values in 4.3, shows the inconsistent performance of classical estimator based on MLE in these datasets. And RSh-S_n estimator exhibit outstanding performance with minimum MSE than other estimators.

Table 4.3: School Data

Estimator	Slope MSE	Intercept MSE	Slope Bias	Intercept Bias
MCD	3.8891	6.7543	0.7086	2.7644
OGK	3.8172	6.6208	0.7015	1.3937
Comedian	3.8073	6.4095	0.6528	2.2576
S_n	3.0669	6.7108	0.6427	2.2301
$RSh - S_n$	3.2752	6.6013	0.6532	2.2106
MLE	3.222	0.0487	0.6837	-0.181

4.6 Summary

Classical regression estimation estimators are highly sensitive to the presence of outliers in the dataset, which can significantly affect their robustness. It is better to make use of alternative estimators capable of detecting and withstanding outliers are necessary to ensure reliable results, even when outliers are present. Several works like Singer & Sen (1985), Koenker & Portnoy (1990) introduced robust regression estimators using M estimators. Ollila et al. (2002), Ollila et al. (2003) proposed estimators based on affine equivariant signs and ranks. P. J. Rousseeuw et al. (2004) proposed regression based on reweighted MCD estimator. Roelant et al. (2009) introduced non - reweighted multivariate regression based on generalised S estimators. Sajana & Sajesh (2018) introduced non reweighted multivari-



Diagnostic Plot based on MLE for School data

Diagnostic Plot based on Shrinkage Sn for School data

Diagnostic Plot based on Comedian for School data

Diagnostic Plot based on MCD for School data

Diagnostic Plot based on OGK for School data

Diagnostic Plot based on Sn for School data

Figure 4.3: Diagnostic plot based on different estimators for School data

Table 4.4: Robustness table-Contaminants from N($\mathbf{2}\sqrt{\boldsymbol{\chi}^{\mathbf{2}}_{\boldsymbol{p+q,0.99}}}, 0.1 \times \mathbf{I}_q$) and (p, q) = (4,4)

	Slope MSE	Intercept MSE	diagonal element mse	non-diagonal elements mse	slope bias	intercept bias	diagonal element bias	non diagonal element bias
		$n = 50$	δ = 10%	$p = 4$	$q = 4$	diag=0.1		
S_n	0.027062	0.024811	0.869399	0.02140	0.003809	0.001816	0.908377	-0.00498
MCD	0.056792	0.029701	0.605927	0.02885	-0.00122	-7.10e-5	0.740357	-0.00533
Comedian	0.027862	0.024672	0.852523	0.02171	0.002757	0.000244	0.898504	-0.00392
OGK	0.032135	0.025488	0.75953	0.02285	0.000946	0.00393	0.842867	-0.00075
RSh-S_n	0.027537	0.02368	0.80369	0.01998	0.00056	5.10e-3	0.87283	-0.00295
		$n = 100$	δ=10%	$p = 4$	$q = 4$	diag=0.1		
S_n	0.012804	0.011781	0.903699	0.011783	-0.00017	-0.00153	0.938678	0.006895
MCD	0.015236	0.01145	0.750362	0.01302	-0.00177	0.000362	0.850676	0.008832
Comedian	0.012505	0.011895	0.902279	0.011021	-0.00057	-0.00269	0.93753	0.00686
OGK	0.014642	0.013542	0.847704	0.012663	0.000848	0.000761	0.906497	0.008697
RSh-S_n	0.01248	0.0113	0.864418	0.010806	0.000346	0.002428	0.917793	0.006029
		$n = 200$	δ=10%	$p = 4$	$q = 4$	diag=0.1		
S_n	0.005973	0.005682	0.926016	0.005729	-8.00e-5	0.000141	0.956202	0.003477
MCD	0.006437	0.00573	0.846856	0.006063	0.000298	2.51e-6	0.913761	0.003293
Comedian	0.005955	0.005523	0.923031	0.005722	-5.90e-5	-9.10e-6	0.954904	0.00261
OGK	0.013524	0.011409	0.81934	0.011265	0.001151	0.003127	0.891392	0.00972
RSh-S_n	0.00591	0.005506	0.898654	0.005729	0.000355	-0.00044	0.941665	0.002543
		$n = 500$	δ=10%	$p = 4$	$q = 4$	diag=0.1		
S_n	0.002309	0.002166	0.93471	0.002198	4.65e-5	-0.00025	0.964442	0.001258
MCD	0.002404	0.00217	0.905404	0.002333	-1.90e-5	0.000597	0.949063	0.002317
Comedian	0.00231	0.002195	0.933463	0.002302	-0.00019	0.001192	0.963769	0.001277
OGK	0.002426	0.002143	0.893785	0.002358	-4.90e-5	0.000273	0.94283	0.000895
RSh-S_n	0.00266	0.00242	0.92209	0.00258	-0.00025	0.00093	0.95751	0.00021

Table 4.5: Efficiency table

Method	Slope	Intercept	Σ_{diag}	$\Sigma_{offdiag}$
	$n = 100$	$p = 4$	$q = 4$	
RSh-S_n	1.133559	0.987338	2.26897	1.166734
MLE	1.098142	1.062179	2.179757	1.118315
Comedian	1.208254	1.036421	2.626631	1.28434
MCD	1.855456	1.611511	4.231406	2.018595
OGK	1.251197	1.134508	2.785136	1.343533
S_n	1.055214	1.006972	2.244428	1.14198
	$n = 100$	$p = 4$	$q = 8$	
RSh-S_n	1.069516	1.05058	2.171636	1.082605
MLE	1.079249	1.052644	2.194896	1.087536
Comedian	1.216045	1.096154	2.454006	1.268528
MCD	1.781371	1.614493	3.718948	1.850268
OGK	1.253904	1.080404	2.717749	1.311943
S_n	1.074581	1.0278	2.206211	1.093251
	$n = 100$	$p = 8$	$q = 4$	
RSh-S_n	1.187912	1.027722	2.450559	1.203182
MLE	1.131539	1.114378	2.46635	1.156902
Comedian	1.27453	1.158034	2.739073	1.334702
MCD	2.173524	1.856234	4.767762	2.382701
OGK	1.324617	1.135834	2.935003	1.443301
S_n	1.132202	1.067909	2.397812	1.122268
	$n = 100$	$p = 10$	$q = 10$	
RSh-S_n	1.197286	1.11769	2.353268	1.186694
MLE	1.146476	1.141881	2.296645	1.135159
Comedian	1.318311	1.204311	2.674198	1.319547
MCD	2.350712	2.063462	4.494091	2.282524
OGK	1.414083	1.240724	2.915869	1.421498
S_n	1.261643	1.165052	2.627408	1.261015

ate regression based Comedian estimator. In this chapter, we develop a two step reweighted multivariate regression $RSh\text{-}S_n$. We evaluated the performance of our proposed estimator and compared with other existing estimators. We evaluated robustness, efficiency and affine equivariance through the metrics Slope, Intercept bias and MSE. To enhance efficiency, we made use of reweighting schemes and found that the best results are achieved by using $Sh\text{-}S_n$ - based robust distances. These distances are used to create a reweighted estimator of location and scatter, which then serves as the foundation for the initial regression. The robust residuals from this initial regression are then used to determine the weights for the final regression. This two step reweightedness gave us finely performing regression estimator with empirically high asymptotic efficiency. For comparing the robustness, we used contaminated simulated datasets. Contamination schemes are finely chosen to explore the performance of all the estimators. The results shows that our proposed estimator outperforms other estimators in terms of MSE and bias validating high robustness empirically. For higher contamination irrespective of the dimension in data, our proposed estimator works best in terms of MSE, bias and other estimators fails to perform in such situation. We empirically verified the affine equivariance, and the results provided positive confirmation of our proposed estimator equivariance property. We checked the performance of the estimator in two real - life datasets and made use of Diagnostic plot for evaluation purpose. The plots enforce our proposed estimator is capable of detecting outliers in the dataset. Thus our study assures using our proposed estimator in datasets with multiple outliers for multivariate regression estimation.

Chapter 5

Conclusions

Outliers undeniably impact statistical analyses, particularly in regression and multivariate data. They skew estimates, inflate variances, and directly lead to misleading conclusions. Addressing them is essential for ensuring accurate analysis and interpretation. Outlier detection has evolved with methods ranging from simple statistical measures, like z-scores or Mahalanobis distance, robust statistical estimators like M estimator, R estimator, LTS estimator, graphical techniques, multivariate methods etc. These methods are vital in preprocessing, as effectively addressing outliers enhances the reliability and robust analyses, especially in regression models. It is well established that classical sample estimators, such as the sample mean and sample covariance matrix, are significantly sensitive to outliers. Furthermore, outlier detection becomes increasingly demanding as data dimensionality increases, often referred to as the curse of dimensionality. Relying solely on univariate outlier detection measures is insufficient for effectively identifying multivariate outliers. In addition, OLS estimator is highly susceptible to distortion from even a single outlier, which can undermine the integrity of the analysis. In this thesis, we address these issues due to the presence of outliers in multivariate data and regression model. The introduction of thesis provides an overview of outliers, their impact on multivariate datasets, atypical observations in regression data, and how their presence influences data modeling and forecasting. The introduction includes an extensive literature review on robust regression estimators. Literature review includes advantages, disadvantages and need for further developments in existing robust regression estimators. The first contribution of the thesis contribute a robust Shrinkage based covariance estimator and extended it to a robust Mahalanobis distance measure for outlier detection in multivariate datasets. Shrinkage estimation offers the benefits of reducing estimation error and achieving a balance between bias and variance. The efficiency of the proposed shrinkage covariance estimator has been evaluated on non-contaminated data, and the results confirm that it performs well in such settings as well. The sensitivity curve of the proposed shrinkage covariance estimator, compared to the classical MLE, has been plotted. The curve clearly demonstrates that the proposed robust shrinkage covariance estimator remains bounded, regardless of the level of contamination or sample size. The results indicate that the proposed shrinkage estimator possesses a bounded influence function and demonstrates efficiency in non-contaminated data, making it a viable alternative to the classical covariance estimator. Also, shrinkage has the added advantage of yielding a positive definite and well-conditioned estimate, which is essential since the inverse of the covariance estimator is required to define the Mahalanobis distance. Thus, we extend the proposed shrinkage covariance estimator to a Mahalanobis distance measure. Evaluated the performance of proposed robust Mahalanobis distance measure along with FAST-MCD, OGK, RMDS, RMD based on S_n through an extensive simulation study. For the evaluation purpose, specific measures were examined, namely the true positive rate(TPR) and the false positive rate(FPR). Results demonstrates that the proposed distance measure exhibits commendable performance across all levels of contamination, as indicated by its high TPR. In terms of FPR, proposed distance measure consistently maintains a lower rate compared to the other approaches. The study empirically evaluated the properties like robustness, breakdown, correlated data performance and affine equivariance of proposed robust distance measure extensively. The results shows a positive evidence regarding the performance of distance measure in outlier detection. The proposed distance measure is able to detect all the possible outliers presented in the dataset with minimum masking and swamping effect. Also, applied proposed distance measure to a number of several benchmark datasets, and it successfully detects outliers in all cases. The second contribution of this thesis is a robust multiple linear regression approach that utilizes proposed robust shrinkage covariance estimator in an alternative definition of the OLS estimator. The proposed regression estimator, $RSh\text{-}S_n$, is presented in chapter 3 of our thesis. In multiple regression, the estimator involves weighting the observations with robust Mahalanobis distance based on proposed shrinkage covariance estimator, resulting in a robust shrinkage reweighted $RSh\text{-}S_n$ multiple regression estimator. The performance of proposed regression estimator evaluated along with other existing regression estimators using MSE metric. We have analyzed the robustness, efficiency, breakdown and affine equivariance properties of $RSh\text{-}S_n$ estimator empirically. The results support the efficacy of proposed estimator in respective properties. Also, checked the performance of $RSh\text{-}S_n$ estimator in a number of real - life datasets, which promise the usefulness of the estimator.

The fourth chapter of the thesis constitutes the $RSh\text{-}S_n$ multivariate regression estimator. In this chapter, we extend the Mahalanobis distance based on robust shrinkage covariance estimator to propose a two step reweighted multivariate linear regression estimator $RSh\text{-}S_n$. We evaluated the robustness and efficiency property of proposed estimator along with other existing estimators using metrics such as MSE and bias. The results show that, even in multivariate regression estimation, our estimator possesses better, promising performance in terms of robustness and efficiency. Also, we evaluated the performance in two benchmark datasets too, where, the results favour the $RSh\text{-}S_n$ estimator.

In conclusion, a robust alternative to covariance estimator developed, extended the estimator to a distance measure and evaluated the performance of distance measure thoroughly. Then, we extended the robust covariance estimator to alternative methods of multiple and multivariate linear regression estimators. Extensive analysis done on proposed estimators in two different chapters. The results highlight the advantages of the proposed estimators across various scenarios and demonstrate their empirical properties.

Chapter 6

Recommendations

- To extend the proposed robust estimator to non - linear regression and evaluate the performance.
- To study the robustness of quantile regression estimators.
- To develop robust Bayesian regression estimator.

List of Publications

Lakshmi, R., Sajesh, T. A., “Comparative Study on Robust Estimators and Evaluating their Performance in Multiple Regression”, Aligarh Journal of Statistics 43 (2023).

Lakshmi, R., & Sajesh, T. A. (2023). Empirical Study on Robust Regression Estimators and Their Performance. Reliability: Theory & Applications, 18(2 (73)), 466-478.

Lakshmi, R., & Sajesh, T. A. (2024). A robust distance-based approach for detecting multidimensional outliers. Journal of Applied Statistics, 1–21.
https://doi.org/10.1080/02664763.2024.2422403.

Lakshmi, R., & Sajesh, T. A., “Robust multiple regression based on Shrinkage S_n estimator ”, Accepted in Hacettepe Journal of Mathematics and Statistics.

Lakshmi R., & Sajesh, T. A., “Robust multivariate regression based on Shrinkage S_n estimator”, Communicated in Computational Statistics and Data Analysis.

List of Presentations

Lakshmi R, presented “Multivariate Outlier detection methods comparison study” in International Conference on Recent Advances of Probability and Statistics in Interdisciplinary Research (RAPSIR - 2024) organized by the Department of Statistics, Faculty of Science, University of Allahabad, Prayagraj, India in conjunction with 43rd Annual Convention of Indian Society for Probability and Statistics (ISPS) during 06 - 08 February 2024.

Lakshmi R, presented “Empirical study on robust regression estimators and their performance” in Eighth International Conference on Statistics for Twenty - first Century - 2022 (ICSTC - 2022) organized by the International Statistics Fraternity, School of Physical and Mathematical Sciences and Department of Statistics, University of Kerala, Trivandrum during 16 - 19 December 2022.

Lakshmi R, presented “Robust Regression based on Sn Covariance Estimator” in International Conference on Knowledge Discoveries on Statistical Innovations & Recent Advances In Optimization (ICON - KSRAO) organized by the Department of Statistics and Population Research Centre, Andhra University, Visakhapatnam during 29 - 30 December 2022.

Lakshmi R, presented “Multidimensional Adaptive Sn estimator and its performance” in International Conference on Applications of Statistics organized by the Department of Statistics, Pachhunga University College campus, Mizoram University during 2 - 3 February 2023.

Lakshmi R, presented “Comparison of ridge regression estimators: Exploring old and new” in National Seminar on Recent Developments and Applications in Statistics and Probability organized by Department of Statistics, St. Thomas College(Autonomous), Thrissur,Kerala during 14 - 15 February 2024.

References

Agulló, J., Croux, C., & Van Aelst, S. (2002). The multivariate least trimmed squares estimator. Citeseer.

Agulló, J., Croux, C., & Van Aelst, S. (2008). The multivariate least-trimmed squares estimator. Journal of Multivariate Analysis, 99(3), 311–338.

Anscombe, F. J. (1960). Rejection of outliers. Technometrics, 2(2), 123–146.

Atkinson, A., & Mulira, H.-M. (1993). The stalactite plot for the detection of multivariate outliers. Statistics and Computing, 3, 27–35.

Atkinson, A. C. (1994). Fast very robust methods for the detection of multiple outliers. Journal of the American Statistical Association, 89(428), 1329–1339.

Barnett, V., & Lewis, T. (1994). Outliers in statistical data.

Bassett Jr, G., & Koenker, R. (1978). Asymptotic theory of least absolute error regression. Journal of the American Statistical Association, 73(363), 618–622.

Beckman, R. J., & Cook, R. D. (1983). Outlier.......... s. Technometrics, 25(2), 119–149.

Bickel, P. J. (1975). One-step huber estimates in the linear model. Journal of the American Statistical Association, 70(350), 428–434.

Bickel, P. J., & Lehmann, E. L. (2011). Descriptive statistics for nonparametric models. iii. dispersion. In Selected works of el lehmann (pp. 499–518). Springer.

Billor, N., Hadi, A. S., & Velleman, P. F. (2000). Bacon: blocked adaptive computationally efficient outlier nominators. Computational statistics & data analysis, 34(3), 279–298.

Bose, A. (1995). Estimating the asymptotic dispersion of the l1 median. Annals of the Institute of Statistical Mathematics, 47(2), 267–272.

Bose, A., & Chaudhuri, P. (1993). On the dispersion of multivariate median. Annals of the Institute of Statistical Mathematics, 45(3), 541–550.

Brownlee, K. A. (1965). Statistical theory and methodology in science and engineering. A Wiley Publication in Applied Statistics.

Cabana, E., Lillo, R. E., & Laniado, H. (2020). Robust regression based on shrinkage with application to living environment deprivation. Stochastic environmental research and risk assessment, 34(2), 293–310.

Cabana, E., Lillo, R. E., & Laniado, H. (2021). Multivariate outlier detection based on a robust mahalanobis distance with shrinkage estimators. Statistical papers, 62, 1583–1609.

Campbell, N. (1989). Bushfire mapping using noaa avhrr data. Technical report, CSIRO.

Carroll, R. J., & Ruppert, D. (1985). Transformations in regression: A robust analysis. Technometrics, 27(1), 1–12.

Cerioli, A., Riani, M., & Atkinson, A. C. (2009). Controlling the size of multivariate outlier tests with the mcd estimator of scatter. Statistics and Computing, 19, 341–353.

Charnes, A., Cooper, W. W., & Ferguson, R. O. (1955). Optimal estimation of executive compensation by linear programming. Management science, 1(2), 138–151.

Charnes, A., Cooper, W. W., & Rhodes, E. (1981). Evaluating program and managerial efficiency: an application of data envelopment analysis to program follow through. Management science, 27(6), 668–697.

Coakley, C. W., & Hettmansperger, T. P. (1993). A bounded influence, high breakdown, efficient regression estimator. Journal of the American Statistical Association, 88(423), 872–880.

Croux, C., Rousseeuw, P. J., & Hössjer, O. (1994). Generalized s-estimators. Journal of the American Statistical Association, 89(428), 1271–1281.

Croux, C., Van Aelst, S., & Dehon, C. (2003). Bounded influence regression using high breakdown scatter matrices. Annals of the Institute of Statistical Mathematics, 55, 265–285.

Daudin, J., Duby, C., & Trecourt, P. (1988). Stability of principal component analysis studied by the bootstrap method. Statistics: A journal of theoretical and applied statistics, 19(2), 241–258.

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

DeMiguel, V., Martin-Utrera, A., & Nogales, F. J. (2013). Size matters: Optimal calibration of shrinkage estimators for portfolio selection. Journal of Banking & Finance, 37(8), 3018–3034.

Devlin, S. J., Gnanadesikan, R., & Kettenring, J. R. (1981). Robust estimation of dispersion matrices and principal components. Journal of the American Statistical Association, 76(374), 354–362.

Donoho, D. L., & Huber, P. J. (1983). The notion of breakdown point. A festschrift for Erich L. Lehmann, 157184.

Edgeworth, F. Y. (1887). On observations relating to several quantities. Hermathena, 6(13), 279–285.

EStimator, D. (1999). A fast algorithm for the minimum covariance. Technometrics, 41(3), 212.

Falk, M. (1997). On mad and comedians. Annals of the Institute of Statistical Mathematics, 49, 615–644.

Gervini, D., & Yohai, V. J. (2002). A class of robust and fully efficient regression estimators. The Annals of Statistics, 30(2), 583–616.

Gnanadesikan, R., & Kettenring, J. R. (1972). Robust estimates, residuals, and outlier detection with multiresponse data. Biometrics, 81–124.

Gray, J. B. (1985). Graphics for regression diagnostics. In American statistical association proceedings of the statistical computing section (pp. 102–107).

Grubbs, F. E. (1969). Procedures for detecting outlying observations in samples. Technometrics, 11(1), 1–21.

Hadi, A. S. (1992). Identifying multiple outliers in multivariate data. Journal of the Royal Statistical Society Series B: Statistical Methodology, 54(3), 761–771.

Hampel, F. R. (1968). Contributions to the theory of robust estimation. University of California, Berkeley.

Hampel, F. R. (1971). A general qualitative definition of robustness. The annals of mathematical statistics, 42(6), 1887–1896.

Hampel, F. R. (1974). The influence curve and its role in robust estimation. Journal of the american statistical association, 69(346), 383–393.

Hampel, F. R. (1975). Beyond location parameters: Robust concepts and methods. Eidgenössische Technische Hochschule Zürich.

Hawkins, D. (1980). Identification of outliers. Chapman and Hall.

Hawkins, D. M., Bradu, D., & Kass, G. V. (1984). Location of several outliers in multiple-regression data using elemental sets. Technometrics, 26(3), 197–208.

Hawkins, D. M., & Olive, D. J. (2002). Inconsistency of resampling algorithms for high-breakdown regression estimators and a new algorithm. Journal of the American Statistical Association, 97(457), 136–159.

Hodges Jr, J. L. (1967). Efficiency in normal samples and tolerance of extreme values for some estimates of location. In Proceedings of the fifth berkeley symposium on mathematical statistics and probability (Vol. 1, pp. 163–186).

Holland, P. W., & Welsch, R. E. (1977). Robust regression using iteratively reweighted least-squares. Communications in Statistics-theory and Methods, 6(9), 813–827.

Huber, P. J. (1992). Robust estimation of a location parameter. In Breakthroughs in statistics: Methodology and distribution (pp. 492–518). Springer.

Huber, P. J., & Ronchetti, E. M. (2011). Robust statistics. John Wiley & Sons.

James, W., & Stein, C. (1992). Estimation with quadratic loss. In Breakthroughs in statistics: Foundations and basic theory (pp. 443–460). Springer.

Jana, P., Rosadi, D., & Supandi, E. D. (2023). Comparison of robust estimation on multiple regression model. BAREKENG: Jurnal Ilmu Matematika dan Terapan, 17(2), 0979–0988.

Koenker, R., & Portnoy, S. (1990). M estimation of multivariate regressions. Journal of the American Statistical Association, 85(412), 1060–1068.

Krasker, W. S. (1980). Estimation in linear regression models with disparate data points. Econometrica: Journal of the Econometric Society, 1333–1346.

Krasker, W. S., & Welsch, R. E. (1982). Efficient bounded-influence regression estimation. Journal of the American statistical Association, 77(379), 595–604.

Kunjunni, S. O., & Abraham, S. T. (2022). Multidimensional outlier detection and robust estimation using sn covariance. Communications in Statistics-Simulation and Computation, 51(7), 3912–3922.

Lakshmi, R., & Sajesh, T. (2023). Empirical study on robust regression estimators and their performance. Reliability: Theory & Applications, 18(2 (73)), 466–478.

Ledoit, O., & Wolf, M. (2003a). Honey, i shrunk the sample covariance matrix. UPF economics and business working paper(691).

Ledoit, O., & Wolf, M. (2003b). Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. Journal of empirical finance, 10(5), 603–621.

Ledoit, O., & Wolf, M. (2004). A well-conditioned estimator for large-dimensional covariance matrices. Journal of multivariate analysis, 88(2), 365–411.

Lee, J., Roy, D., Hong, M., & Whiting, P. (1993). Relationships between properties of pulp-fibre and paper. University of Toronto.

Lopuhaä, H. P. (1989). On the relation between s-estimators and m-estimators of multivariate location and covariance. The Annals of Statistics, 1662–1683.

Lopuhaä, H. P., & Rousseeuw, P. J. (1991). Breakdown points of affine equivariant estimators of multivariate location and covariance matrices. The Annals of Statistics, 229–248.

Mallows, C. L. (1975). On some topics in robustness. Unpublished memorandum, Bell Telephone Laboratories, Murray Hill, NJ, 37.

Marazzi, A. (1993). Algorithms, routines, and s-functions for robust statistics. CRC Press.

Maronna, R., & Morgenthaler, S. (1986). Robust regression through robust covariances. Communications in Statistics-Theory and Methods, 15(4), 1347–1365.

Maronna, R. A. (1976). Robust m-estimators of multivariate location and scatter. The annals of statistics, 51–67.

Maronna, R. A., & Yohai, V. J. (1995). The behavior of the stahel-donoho robust multivariate estimator. Journal of the American Statistical Association, 90(429), 330–341.

Maronna, R. A., & Yohai, V. J. (1997). Robust estimation in simultaneous equations models. Journal of Statistical planning and Inference, 57(2), 233–244.

Maronna, R. A., & Zamar, R. H. (2002). Robust estimates of location and dispersion for high-dimensional datasets. Technometrics, 44(4), 307–317.

Mosteller, F., & Tukey, J. W. (1977). Data analysis and regression. a second course in statistics. Addison-Wesley series in behavioral science: quantitative methods.

Möttönen, J., Nordhausen, K., & Oja, H. (2010). Asymptotic theory of the spatial median. In Nonparametrics and robustness in modern statistical inference and time series analysis: A festschrift in honor of professor jana jurečková (Vol. 7, pp. 182–194). Institute of Mathematical Statistics.

Oja, H. (2010). Multivariate nonparametric methods with r: an approach based on spatial signs and ranks. Springer Science & Business Media.

Oja, H., & Niinimaa, A. (1985). Asymptotic properties of the generalized median in the case of multivariate normality. Journal of the Royal Statistical Society. Series B (Methodological), 372–377.

Ollila, E., Oja, H., & Hettmansperger, T. P. (2002). Estimates of regression coefficients based on the sign covariance matrix. Journal of the Royal Statistical Society Series B: Statistical Methodology, 64(3), 447–466.

Ollila, E., Oja, H., & Koivunen, V. (2003). Estimates of regression coefficients based on lift rank covariance matrix. Journal of the American Statistical Association, 98(461), 90–98.

Peña, D., & Prieto, F. J. (2001). Multivariate outlier detection and robust covariance matrix estimation. Technometrics, 43(3), 286–310.

PJ, H. F. R. E. R., & Stahel, W. (1986). Robust statistics. the approach based on the influence function. Wiley & Sons, New York.

Pollard, D. (1991). Asymptotics for least absolute deviation regression estimators. Econometric Theory, 7(2), 186–199.

Reimann, C., Filzmoser, P., & Garrett, R. G. (2005). Background and threshold: critical comparison of methods of determination. Science of the total environment, 346(1-3), 1–16.

Riani, M., Cerioli, A., Atkinson, A. C., Perrotta, D., & Torti, F. (2008). Fitting mixtures of regression lines with the forward search. In Mining massive data sets for security (pp. 271–286). IOS Press.

Rocke, D. M. (1996). Robustness properties of s-estimators of multivariate location and shape in high dimension. The Annals of statistics, 1327–1345.

Rocke, D. M., & Woodruff, D. L. (1993). Computation of robust estimates of multivariate location and shape. Statistica Neerlandica, 47(1), 27–42.

Rocke, D. M., & Woodruff, D. L. (1996). Identification of outliers in multivariate data. Journal of the American Statistical Association, 91(435), 1047–1061.

Roelant, E., Van Aelst, S., & Croux, C. (2009). Multivariate generalized s-estimators. Journal of Multivariate Analysis, 100(5), 876–887.

Rousseeuw, P. (1985). Multivariate estimation with high breakdown point. Mathematical Statistics and Applications B.

Rousseeuw, P., & Yohai, V. (1984). Robust regression by means of s-estimators. In Robust and nonlinear time series analysis: Proceedings of a workshop organized by the sonderforschungsbereich 123 "stochastische mathematische modelle", heidelberg 1983 (pp. 256–272).

Rousseeuw, P. J. (1984). Least median of squares regression. Journal of the American statistical association, 79(388), 871–880.

Rousseeuw, P. J., & Croux, C. (1993a). Alternatives to the median absolute deviation. Journal of the American Statistical Association, 88(424), 1273–1283. Retrieved from `https://www.tandfonline.com/doi/abs/10.1080/01621459.1993.10476408` doi: 10.1080/01621459.1993.10476408

Rousseeuw, P. J., & Croux, C. (1993b). Alternatives to the median absolute deviation. Journal of the American Statistical association, 88(424), 1273–1283.

Rousseeuw, P. J., Hampel, F. R., Ronchetti, E. M., & Stahel, W. A. (1986). Robust statistics: the approach based on influence functions. Wiley, New York.

Rousseeuw, P. J., & Leroy, A. M. (2005). Robust regression and outlier detection. John wiley & sons.

Rousseeuw, P. J., Van Aelst, S., Van Driessen, K., & Gulló, J. A. (2004). Robust multivariate regression. Technometrics, 46(3), 293–305.

Rousseeuw, P. J., & Van Driessen, K. (2000). An algorithm for positive-breakdown regression based on concentration steps. In Data analysis: Scientific modeling and practical application (pp. 335–346). Springer.

Rousseeuw, P. J., & Van Zomeren, B. C. (1990). Unmasking multivariate outliers and leverage points. Journal of the American Statistical association, 85(411), 633–639.

Ruppert, D., & Carroll, R. J. (1980). Trimmed least squares estimation in the linear model. Journal of the American Statistical Association, 75(372), 828–838.

Sajana, O., & Sajesh, T. (2018). Empirical robust multivariate regression parameter estimation using median approach. Int. J. Sci. Res. in Mathematical and Statistical Sciences Vol, 5, 5.

Sajesh, T., & Srinivasan, M. (2012). Outlier detection for high dimensional data using the comedian approach. Journal of statistical computation and simulation, 82(5), 745–757.

Sajesh, T., & Srinivasan, M. (2013). An overview of multiple outliers in multidimensional data. Sri Lankan Journal of Applied Statistics, 14(2), 87–120.

Sen, P. K. (1968). Estimates of the regression coefficient based on kendall's tau. Journal of the American statistical association, 63(324), 1379–1389.

Shevlyakov, G., Morgenthaler, S., & Shurygin, A. (2008). Redescending m-estimators. Journal of Statistical Planning and Inference, 138(10), 2906–2917.

Siegel, A. F. (1982). Robust regression using repeated medians. Biometrika, 69(1), 242–244.

Singer, J. M., & Sen, P. K. (1985). M-methods in multivariate linear models. Journal of multivariate Analysis, 17(2), 168–184.

Smith, R. E., Campbell, N., & Litchfield, R. (1984). Multivariate statistical techniques applied to pisolitic laterite geochemistry at golden grove, western australia. Journal of Geochemical Exploration, 22(1-3), 193–216.

Stromberg, A. J., Hössjer, O., & Hawkins, D. M. (2000). The least trimmed differences regression estimator and alternatives. Journal of the American Statistical Association, 95(451), 853–864.

Susanti, Y., Pratiwi, H., Sulistijowati, S., Liana, T., et al. (2014). M estimation, s estimation, and mm estimation in robust regression. International Journal of Pure and Applied Mathematics, 91(3), 349–360.

Theil, H. (1950). A rank-invariant method of linear and polynomial regression analysis. Indagationes mathematicae, 12(85), 173.

Van Aelst, S., & Willems, G. (2005). Multivariatr regression s-estimators for robust estimation and inference. Statistica Sinica, 981–1001.

Víšek, J. Á. (2006). The least trimmed squares. part iii: Asymptotic normality. Kybernetika, 42(2), 203–224.

Yohai, V. J. (1987). High breakdown-point and high efficiency robust estimates for regression. The Annals of statistics, 642–656.

Yohai, V. J., & Zamar, R. H. (1988). High breakdown-point estimates of regression by means of the minimization of an efficient scale. Journal of the American statistical association, 83(402), 406–413.

Yu, C., & Yao, W. (2017). Robust linear regression: A review and comparison. Communications in Statistics-Simulation and Computation, 46(8), 6261–6282.

Zhou, W., & Serfling, R. (2008). Multivariate spatial u-quantiles: A bahadur–kiefer representation, a theil–sen estimator for multiple regression, and a robust dispersion estimator. Journal of Statistical Planning and Inference, 138(6), 1660–1678.

Appendix A

A.1 Additional tables and figures from chapter 2

Table S1: Zero contaminated Multivariate Normal distribution FPR table for different sample size

p	Sh-S_n	**RMDS**	S_n	**FAST MCD**	**OGK**
			$n = 500$		
5	0.0204	0.0267	0.1827	0.03959	0.1112
10	0.0173	0.0282	0.0007	0.04456	0.10341
15	0.0113	0.0294	0.0613	0.05298	0.1082
20	0.0117	0.0312	0.0581	0.06402	0.09989
30	0.0089	0.0389	0.0585	0.0961	0.10478
50	0.0083	0.0467	0.0652	0.15891	0.10912
			$n = 100$		
5	0.0239	0.0382	0.1981	0.1363	0.1208
10	0.0178	0.0459	0.002	0.1821	0.1208
15	0.0210	0.0609	0.0809	0.2395	0.1184
20	0.0108	0.0716	0.0634	0.2751	0.1234
30	0.0081	0.1090	0.0618	0.2916	0.1287
50	0.0045	0.3761	0.0689	0.2451	0.1323

Table S2: Multivariate Normal distribution TPR table for $n = 100$

p	α	**FAST MCD**	**OGK**	**RMDS**	S_n	**Shrinkage** S_n
δ=5,λ=0.01						
	0.1	1	1	1	1	1
5	0.2	0.4	0.9936	1	1	1
	0.3	0	0.0017	0.8887	1	1
	0.1	1	1	1	1	1
10	0.2	0	1	1	1	1
	0.3	0	0	0.5661	0.9959	1
	0.1	0.544	1	1	1	1
15	0.2	0	1	1	1	1
	0.3	0	0	0.476	0.7959	1
	0.1	0	1	1	1	1
20	0.2	0	1	0.992	1	1
	0.3	0	0	0.3533	0.8777	1
	0.1	0	1	1	1	1
30	0.2	0	1	0.913	1	1
	0.3	0	0	0.2019	0.952	1
	0.1	0	1	0.954	1	1
50	0.2	0	1	0.678	1	1
	0.3	0	0	0.1179	0.9894	1
δ=5,λ=0.1						
	0.1	1	1	1	1	1
5	0.2	0.9760	1	1	1	1
	0.3	0.0003	0.0521	0.8983	1	1
	0.1	1	1	1	1	1
10	0.2	0	1	1	1	1
	0.3	0.0001	0.03307	0.6209	0.9993	1
	0.1	0.968	1	1	1	1
15	0.2	0	1	1	1	1
	0.3	0	0.0481	0.4855	0.9896	1
	0.1	0.216	1	1	1	1
20	0.2	0	1	0.9872	1	1
	0.3	0	0.0504	0.3503	1	1
	0.1	0	1	1	1	1
30	0.2	0	1	0.9162	1	1
	0.3	0	0.1056	0.218	1	1
	0.1	0	1	0.9648	1	1
50	0.2	0	1	0.703	1	1
	0.3	0	0.2035	0.1604	1	1
δ=5,λ=1						
	0.1	1	1	1	1	1
5	0.2	1	1	1	1	1
	0.3	1	0.9359	0.8943	1	0.9985
	0.1	1	1	1	1	1
10	0.2	1	1	1	1	1
	0.3	1	0.9928	0.858	1	1
	0.1	1	1	1	1	1
15	0.2	1	1	1	1	1
	0.3	0.7511	0.9987	0.8039	1	1
	0.1	1	1	1	1	1
20	0.2	0.9846	1	0.9942	1	1
	0.3	0.3816	0.9999	0.6417	1	1
	0.1	1	1	1	1	1
30	0.2	0.4572	1	0.897	1	1
	0.3	0.3228	1	0.4437	1	1
	0.1	0.4632	1	0.9436	1	1
50	0.2	0.2716	1	0.7082	1	1
	0.3	0.2596	1	0.3492	1	1

Table S3: Multivariate Normal distribution TPR table for $n = 100$

p	$\boldsymbol{\alpha}$	**FAST MCD**	**OGK**	**RMDS**	$\boldsymbol{S_n}$	Sh-S_n
			δ=10,λ=0.01			
5	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	0	0.8801	1	1	1
10	0.1	1	1	1	1	1
	0.2	0.004	1	1	1	1
	0.3	0	0.8533	0.996	1	1
15	0.1	1	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	0.7655	0.9579	1	1
20	0.1	0.8	1	1	1	1
	0.2	0	1	0.996	1	1
	0.3	0	0.7892	0.9161	1	1
30	0.1	0	1	1	1	1
	0.2	0	1	0.975	1	1
	0.3	0	0.84827	0.8703	1	1
50	0.1	0	1	0.9652	1	1
	0.2	0	1	0.8618	1	1
	0.3	0	0.9131	0.7487	1	1
			δ=10,λ=0.1			
5	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	0.5601	0.9988	1	1	1
10	0.1	1	1	1	1	1
	0.2	0.404	1	1	1	1
	0.3	0	1	0.992	1	1
15	0.1	1	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	1	0.9515	1	1
20	0.1	0.992	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	1	0.8885	1	1
30	0.1	0.0044	1	1	1	1
	0.2	0	1	0.9726	1	1
	0.3	0	1	0.85	1	1
50	0.1	0	1	0.972	1	1
	0.2	0	1	0.8756	1	1
	0.3	0	1	0.6861	1	1
			δ=10,λ=1			
5	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	1	1	1	1	1
10	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	1	1	0.9925	1	1
15	0.1	1	1	1	1	1
	0.2	1	1	0.9974	1	1
	0.3	0.8275	1	0.9444	1	1
20	0.1	1	1	1	1	1
	0.2	0.9874	1	0.9996	1	1
	0.3	0.39	1	0.88253	1	1
30	0.1	1	1	1	1	1
	0.2	0.4404	1	0.9444	1	1
	0.3	0.3163	1	0.8159	1	1
50	0.1	0.486	1	0.952	1	1
	0.2	0.275	1	0.8518	1	1
	0.3	0.2555	1	0.7103	1	1

Table S4: Multivariate Normal distribution TPR table for $n = 500$

p	α	**FAST MCD**	**OGK**	**RMDS**	S_n	Sh-S_n
			δ=5,λ=0.01			
5	0.1	1	1	1	1	1
	0.2	0.61	0.9916	1	1	1
	0.3	0	0.0011	0.7768	1	1
10	0.1	1	1	1	1	1
	0.2	0	0.994	1	1	1
	0.3	0	0	0	1	1
15	0.1	0.69	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	0	0	0.6473	1
20	0.1	0	1	1	1	1
	0.2	0	0.9952	1	1	1
	0.3	0	0	0	0.8719	1
30	0.1	0	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	0	0	0.9832	1
50	0.1	0	1	1	1	1
	0.2	0	1	0.0041	1	1
	0.3	0	0	0	1	1
			δ=5,λ=0.1			
5	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	0	0	0.7844	1	1
10	0.1	1	1	1	1	1
	0.2	0	0.9984	1	1	1
	0.3	0	0	0.0146	1	1
15	0.1	1	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	0	0	0.9873	1
20	0.1	0.1	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	0	0	0.98	1
30	0.1	0	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	0	0	0.9899	1
50	0.1	0	1	1	1	1
	0.2	0	1	0.1896	1	1
	0.3	0	0	0	1	1
			δ=5,λ=1			
5	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	1	0.9271	0.8365	1	1
10	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	1	0.9897	0.5839	1	1
15	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	0.943	0.9992	0.3278	1	1
20	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	0.4087	0.9997	0.1980	0.9988	1
30	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	0.0917	1	0.0946	0.9987	1
50	0.1	1	1	1	1	1
	0.2	0.2341	1	0.9999	1	1
	0.3	0.1789	1	0.1024	1	1

Table S5: Multivariate Normal distribution TPR table for $n = 500$

p	$\boldsymbol{\alpha}$	**FAST MCD**	**OGK**	**RMDS**	$\boldsymbol{S_n}$	Sh-S_n
δ=10,λ=0.01						
5	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	0	0.7047	1	1	1
10	0.1	1	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	0.5567	0.996	1	1
15	0.1	1	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	0.4867	0.9579	1	1
20	0.1	1	1	1	1	1
	0.2	0	1	0.996	1	1
	0.3	0	0.5871	0.9161	1	1
30	0.1	1	1	1	1	1
	0.2	0	1	0.975	1	1
	0.3	0	0.637	0.8703	1	1
50	0.1	0	1	0.9652	1	1
	0.2	0	1	0.8618	1	1
	0.3	0	0.84	0.7487	1	1
δ=10 λ=0.1						
5	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	0.68	1	1	1	1
10	0.1	1	1	1	1	1
	0.2	0.35	1	1	1	1
	0.3	0	1	0.992	1	1
15	0.1	1	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	1	0.9515	1	1
20	0.1	1	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	1	0.8885	1	1
30	0.1	1	1	1	1	1
	0.2	0	1	0.9726	1	1
	0.3	0	1	0.85	1	1
50	0.1	0	1	0.972	1	1
	0.2	0	1	0.8756	1	1
	0.3	0	1	0.6861	1	1
δ=10,λ=1						
5	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	1	1	1	1	1
10	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	1	1	0.9925	1	1
15	0.1	1	1	1	1	1
	0.2	1	1	0.9974	1	1
	0.3	0.9709	1	0.9444	1	1
20	0.1	1	1	1	1	1
	0.2	1	1	0.9996	1	1
	0.3	0.3693	1	0.8825	1	1
30	0.1	1	1	1	1	1
	0.2	1	1	0.9444	1	1
	0.3	0.0681	1	0.8159	1	1
50	0.1	1	1	0.952	1	1
	0.2	0.2582	1	0.8518	1	1
	0.3	0.1649	1	0.7103	1	1

Table S6: Multivariate Normal distribution TPR table for $n = 1000$

p	$\boldsymbol{\alpha}$	**FAST MCD**	**OGK**	**RMDS**	$\boldsymbol{S_n}$	$Sh\text{-}S_n$
δ=5,λ=0.01						
5	0.1	1	1	1	1	1
	0.2	0.7523	1	1	1	1
	0.3	0	0	1	1	1
10	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	0	0	0	1	1
15	0.1	0.7712	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	0	0	0.7875	1
20	0.1	0	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	0	0	0.8901	1
30	0.1	0	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	0	0	0.9409	0.9871
50	0.1	0	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	0	0	0.9562	0.9671
δ=5,λ=0.1						
5	0.1	1	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	0	0.9993	1	1
10	0.1	1	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	0	0.0867	1	1
15	0.1	1	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	0	0	1	1
20	0.1	0	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	0	0	1	1
30	0.1	0	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	0	0	0.9711	0.9887
50	0.1	0	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	0	0	0.9788	0.9998
δ=5,λ=1						
5	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	1	0.6415	0.9688	1	1
10	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	1	0.9182	0.8487	1	1
15	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	0.9683	0.9899	0.4229	1	1
20	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	0.4812	0.9988	0.2444	1	1
30	0.1	1	1	1	1	1
	0.2	0.9922	1	1	1	1
	0.3	0.0709	1	0.0999	1	1
50	0.1	1	1	1	1	1
	0.2	0.1935	1	1	1	1
	0.3	0.0992	1	0	1	1

Table S7: Multivariate Normal distribution TPR table for $n = 1000$

p	$\boldsymbol{\alpha}$	**FAST MCD**	**OGK**	**RMDS**	$\boldsymbol{S_n}$	$Sh\text{-}S_n$
δ=10,λ=0.01						
5	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	0	0.6318	1	1	1
10	0.1	1	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	0.4034	0.6965	1	1
15	0.1	1	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	0.4914	0.7771	0.9591	1
20	0.1	1	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	0.3199	0.7998	0.9679	1
30	0.1	0	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	0.5811	0.8541	0.9729	1
50	0.1	0	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	0.7618	0.8716	0.9913	1
δ=10,λ=0.1						
5	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	0.66	1	1	1	1
10	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	0	1	1	1	1
15	0.1	1	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	1	1	1	1
20	0.1	1	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	1	1	1	1
30	0.1	0	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	1	0.9983	1	1
50	0.1	0	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	1	0.9995	1	1
δ=10,λ=1						
5	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	1	1	1	1	1
10	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	1	1	1	1	1
15	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	0.9938	1	1	1	1
20	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	0.5911	1	1	1	1
30	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	0.1156	1	1	1	1
50	0.1	1	1	1	1	1
	0.2	0.2217	1	1	1	1
	0.3	0.1002	1	1	1	1

Figure A.1: Figure of FPR for multivariate normal distribution for $n = 1000$

Table S8: Multivariate Normal distribution FPR table for $n = 100$

p	$\boldsymbol{\alpha}$	**FAST MCD**	**OGK**	**RMDS**	$\boldsymbol{S_n}$	Sh-S_n
δ=5,λ=0.01						
5	0.1	0.0809	0.0866	0.0111	0.1428	0.0361
	0.2	0.2763	0.0711	0.0022	0.2723	0.0194
	0.3	0.6370	0.2887	0.0012	0.5303	0.0055
10	0.1	0.1360	0.0939	0.0106	0.0014	0.0344
	0.2	0.5217	0.0831	0.0036	0.0206	0.0197
	0.3	0.6389	0.3319	0.0135	0.2099	0.0069
15	0.1	0.2726	0.1008	0.0148	0.0507	0.0364
	0.2	0.5042	0.0867	0.0091	0.0355	0.0213
	0.3	0.5991	0.3414	0.0675	0.0189	0.0076
20	0.1	0.3901	0.1033	0.0188	0.0487	0.0348
	0.2	0.4928	0.0855	0.0327	0.0362	0.0194
	0.3	0.5713	0.3356	0.0713	0.02	0.0088
30	0.1	0.3727	0.1095	0.0409	0.0534	0.0367
	0.2	0.4368	0.0969	0.1279	0.0350	0.0232
	0.3	0.5	0.3345	0.0818	0.0201	0.0093
50	0.1	0.2776	0.1222	0.2339	0.0553	0.0329
	0.2	0.3125	0.0989	0.1654	0.0366	0.0214
	0.3	0.3571	0.34	0.0635	0.0182	0.0084
δ=5,λ=0.1						
5	0.1	0.0826	0.0816	0.0096	0.1414	0.0345
	0.2	0.0551	0.0659	0.0031	0.2022	0.0179
	0.3	0.4466	0.0738	0.0019	0.3306	0.0058
10	0.1	0.1309	0.0935	0.0123	0.0009	0.0323
	0.2	0.4363	0.0743	0.0038	0.0066	0.0209
	0.3	0.6034	0.0985	0.0155	0.0314	0.0067
15	0.1	0.1819	0.0952	0.0157	0.0504	0.0343
	0.2	0.4747	0.0739	0.0084	0.0326	0.021
	0.3	0.5862	0.1135	0.0472	0.0144	0.0089
20	0.1	0.3360	0.0976	0.0206	0.0508	0.0359
	0.2	0.4737	0.0787	0.0265	0.0324	0.0205
	0.3	0.5663	0.1169	0.0741	0.0163	0.0091
30	0.1	0.3625	0.1034	0.0501	0.0507	0.0357
	0.2	0.4311	0.0814	0.1314	0.0317	0.0215
	0.3	0.4989	0.1234	0.0775	0.0175	0.0079
50	0.1	0.2764	0.1183	0.2551	0.0561	0.0352
	0.2	0.3124	0.0904	0.1590	0.0344	0.0218
	0.3	0.3570	0.1141	0.0806	0.0156	0.0090
δ=5,λ=1						
5	0.1	0.0794	0.0847	0.0115	0.1150	0.0339
	0.2	0.0499	0.0555	0.0033	0.1004	0.0203
	0.3	0.0233	0.0334	0.0003	0.1027	0.0056
10	0.1	0.1368	0.0875	0.0109	0.0005	0.0346
	0.2	0.0889	0.0609	0.004	0.0005	0.0196
	0.3	0.0435	0.0379	0.0071	0.0005	0.0072
15	0.1	0.1784	0.0942	0.0135	0.0470	0.0344
	0.2	0.1165	0.0645	0.0076	0.0296	0.0204
	0.3	0.1055	0.0389	0.0499	0.0148	0.0077
20	0.1	0.2055	0.0902	0.0174	0.0462	0.0333
	0.2	0.1299	0.0682	0.0326	0.0332	0.0214
	0.3	0.2345	0.0426	0.0705	0.0162	0.0086
30	0.1	0.1990	0.0999	0.0449	0.0489	0.0369
	0.2	0.243	0.0709	0.1174	0.0294	0.0211
	0.3	0.2715	0.0430	0.0733	0.0155	0.0094
50	0.1	0.2182	0.1113	0.2300	0.0548	0.0350
	0.2	0.2366	0.0792	0.1369	0.0329	0.0233
	0.3	0.2373	0.0449	0.0925	0.0171	0.0087

Table S9: Multivariate Normal distribution FPR table for $n = 100$

p	$\boldsymbol{\alpha}$	**FAST MCD**	**OGK**	**RMDS**	$\boldsymbol{S_n}$	Sh-S_n
			δ=10,λ=0.01			
	0.1	0.0824	0.0837	0.0105	0.112	0.0336
5	0.2	0.0456	0.0690	0.0032	0.1527	0.0198
	0.3	0.6358	0.0586	0.0011	0.3487	0.006
	0.1	0.1326	0.0940	0.012	0.0008	0.0322
10	0.2	0.5142	0.0819	0.0040	0.0051	0.0191
	0.3	0.6399	0.0622	0.01	0.0845	0.0066
	0.1	0.1734	0.0937	0.0146	0.0509	0.0371
15	0.2	0.5069	0.0895	0.0069	0.0422	0.0209
	0.3	0.5995	0.0674	0.0396	0.0258	0.0074
	0.1	0.2449	0.1008	0.0173	0.0508	0.0366
20	0.2	0.4916	0.0936	0.0218	0.0387	0.0222
	0.3	0.5713	0.0652	0.0764	0.024	0.0084
	0.1	0.372	0.1047	0.0431	0.0537	0.0354
30	0.2	0.4368	0.1002	0.1249	0.0415	0.0227
	0.3	0.515	0.0655	0.1063	0.0244	0.0102
	0.1	0.2757	0.1156	0.2363	0.0583	0.0369
50	0.2	0.3135	0.1084	0.1283	0.0415	0.0218
	0.3	0.3718	0.0711	0.0687	0.0254	0.0101
			δ=10,λ=0.1			
	0.1	0.0809	0.0819	0.0115	0.1028	0.0329
5	0.2	0.0518	0.0608	0.0038	0.124	0.0165
	0.3	0.2235	0.0486	0.0011	0.2137	0.0045
	0.1	0.1330	0.0896	0.0118	0.0005	0.0362
10	0.2	0.3024	0.0702	0.0039	0.0019	0.0211
	0.3	0.6052	0.0521	0.0192	0.0102	0.0077
	0.1	0.1692	0.0982	0.0144	0.0509	0.0339
15	0.2	0.4745	0.0784	0.0123	0.0367	0.0217
	0.3	0.5869	0.0501	0.0431	0.0182	0.0081
	0.1	0.2020	0.1014	0.0199	0.0515	0.0347
20	0.2	0.4715	0.0847	0.0267	0.0354	0.0227
	0.3	0.5662	0.0509	0.0804	0.0198	0.008
	0.1	0.3604	0.1076	0.0449	0.0559	0.0344
30	0.2	0.4313	0.0843	0.1183	0.0351	0.0204
	0.3	0.4989	0.0547	0.097	0.0207	0.0089
	0.1	0.2768	0.1117	0.2313	0.0564	0.0355
50	0.2	0.3124	0.0934	0.1535	0.0369	0.0207
	0.3	0.3570	0.0545	0.0757	0.0182	0.0103
			δ=10,λ=1			
	0.1	0.0785	0.0793	0.0111	0.0925	0.0318
5	0.2	0.0471	0.0534	0.0029	0.0715	0.0161
	0.3	0.0219	0.0343	0.0010	0.0674	0.0058
	0.1	0.1325	0.0867	0.0124	0.0002	0.0341
10	0.2	0.0835	0.0634	0.0044	0.0002	0.0202
	0.3	0.0465	0.0362	0.016	0.0002	0.008
	0.1	0.1796	0.0927	0.0155	0.0466	0.0328
15	0.2	0.1151	0.0633	0.006	0.029	0.0198
	0.3	0.0927	0.0379	0.0458	0.014	0.0085
	0.1	0.2092	0.0949	0.0172	0.0484	0.0356
20	0.2	0.1324	0.0653	0.0254	0.0324	0.0196
	0.3	0.2253	0.0410	0.0719	0.0151	0.0087
	0.1	0.2041	0.1015	0.0449	0.0523	0.0356
30	0.2	0.2471	0.0758	0.1227	0.0292	0.0205
	0.3	0.2725	0.0442	0.0831	0.0154	0.0091
	0.1	0.2139	0.1107	0.2286	0.0524	0.0356
50	0.2	0.2366	0.0739	0.1588	0.0339	0.0219
	0.3	0.2389	0.0465	0.0711	0.0183	0.0087

Table S10: Multivariate Normal distribution FPR table for $n = 500$

p	$\boldsymbol{\alpha}$	**FAST MCD**	**OGK**	**RMDS**	$\boldsymbol{S_n}$	Sh-S_n
δ=5,λ=0.01						
	0.1	0.0134	0.0843	0.0076	0.1987	0.0334
5	0.2	0.048	0.0706	0.0014	0.3873	0.0185
	0.3	0.454	0.2799	0.0002	0.6447	0.0067
	0.1	0.0178	0.0920	0.0056	0.0023	0.0336
10	0.2	0.2102	0.0821	0.0007	0.0422	0.0213
	0.3	0.7099	0.3262	0.00003	0.3065	0.0079
	0.1	0.0373	0.0986	0.0045	0.0392	0.0349
15	0.2	0.3983	0.0879	0.0003	0.0250	0.0208
	0.3	0.8669	0.3447	0	0.0123	0.0096
	0.1	0.1109	0.1025	0.0041	0.0407	0.0364
20	0.2	0.5044	0.0869	0.0004	0.0255	0.0208
	0.3	0.981	0.336	0.00002	0.0128	0.0098
	0.1	0.2016	0.1100	0.0034	0.0403	0.0358
30	0.2	0.5928	0.0949	0.0002	0.0259	0.0218
	0.3	0.9799	0.3359	0.00001	0.0133	0.0113
	0.1	0.4267	0.1213	0.0025	0.0414	0.0349
50	0.2	0.5921	0.0981	0.00009	0.0268	0.0222
	0.3	0.9388	0.3436	0.0001	0.0141	0.0119
δ=5,λ=0.1						
	0.1	0.0159	0.0837	0.0072	0.1672	0.0335
5	0.2	0.0067	0.0664	0.0013	0.2562	0.0187
	0.3	0.1877	0.0714	0.0002	0.2795	0.0064
	0.1	0.0249	0.0875	0.0057	0.0024	0.0345
10	0.2	0.1371	0.0715	0.0009	0.0446	0.0203
	0.3	0.5025	0.0949	0.00003	0.3324	0.0086
	0.1	0.0293	0.0943	0.0051	0.0485	0.0342
15	0.2	0.2502	0.0764	0.0006	0.0292	0.0204
	0.3	0.7062	0.1072	0.00001	0.0183	0.0089
	0.1	0.0933	0.1034	0.0038	0.0375	0.0363
20	0.2	0.3975	0.0743	0.0003	0.0289	0.0211
	0.3	0.7959	0.1075	0.00001	0.0185	0.0102
	0.1	0.2123	0.1072	0.0034	0.0355	0.0354
30	0.2	0.5769	0.0818	0.0003	0.0299	0.0216
	0.3	0.8945	0.1131	0	0.0191	0.0107
	0.1	0.4904	0.114	0.0025	0.0465	0.0356
50	0.2	0.6604	0.0906	0.0002	0.0299	0.0225
	0.3	0.8084	0.1174	0.00006	0.0197	0.0115
δ=5,λ=1						
	0.1	0.0158	0.0809	0.0070	0.1404	0.0329
5	0.2	0.0108	0.0526	0.0013	0.3058	0.0182
	0.3	0.0102	0.032	0.0002	0.4008	0.0069
	0.1	0.0199	0.0872	0.0057	0	0.0339
10	0.2	0.0208	0.0596	0.0007	0	0.0198
	0.3	0.0105	0.0367	0.00001	0	0.0084
	0.1	0.0199	0.0895	0.005	0.0378	0.0339
15	0.2	0.0225	0.0656	0.0004	0.0299	0.0212
	0.3	0.0185	0.0387	0	0.0189	0.0096
	0.1	0.0299	0.0944	0.0043	0.0467	0.0345
20	0.2	0.0196	0.0673	0.0004	0.0389	0.0211
	0.3	0.0299	0.0405	0	0.0199	0.0101
	0.1	0.0449	0.0974	0.0034	0.0473	0.0360
30	0.2	0.0297	0.0737	0.0001	0.0387	0.0213
	0.3	0.0395	0.0424	0.00001	0.0298	0.0108
	0.1	0.1249	0.1090	0.0025	0.0444	0.0343
50	0.2	0.1319	0.0773	0.0002	0.0372	0.0225
	0.3	0.1961	0.0458	0.0001	0.0187	0.0117

Table S11: Multivariate Normal distribution FPR table for $n = 500$

p	$\boldsymbol{\alpha}$	**FAST MCD**	**OGK**	**RMDS**	$\boldsymbol{S_n}$	Sh-S_n
			δ=10,λ=0.01			
5	0.1	0.0232	0.0859	0.0105	0.1445	0.0318
	0.2	0.0083	0.0797	0.0032	0.2389	0.0188
	0.3	0.5494	0.0687	0.0011	0.4843	0.0062
10	0.1	0.0205	0.0912	0.012	0.0006	0.0332
	0.2	0.3909	0.0879	0.0040	0.0438	0.0201
	0.3	0.8096	0.0538	0.01	0.3619	0.008
15	0.1	0.0254	0.0896	0.0146	0.0464	0.0351
	0.2	0.4168	0.0926	0.0069	0.0336	0.0207
	0.3	0.9067	0.0788	0.0396	0.01	0.0089
20	0.1	0.0378	0.0888	0.0173	0.0464	0.0344
	0.2	0.5891	0.0879	0.0218	0.0323	0.0206
	0.3	0.9997	0.0469	0.0764	0.0108	0.0094
30	0.1	0.3137	0.0987	0.0431	0.0476	0.0363
	0.2	0.6996	0.0779	0.1249	0.0324	0.0223
	0.3	0.981	0.0431	0.1063	0.0129	0.0096
50	0.1	0.4593	0.0955	0.2363	0.0481	0.0367
	0.2	0.6796	0.7202	0.1283	0.0311	0.0217
	0.3	0.9593	0.4069	0.0687	0.0211	0.0099
			δ=10 λ=0.1			
5	0.1	0.0216	0.0773	0.0115	0.2506	0.0327
	0.2	0.0064	0.0699	0.0038	0.1721	0.0171
	0.3	0.0716	0.0511	0.0011	0.1187	0.006
10	0.1	0.0262	0.0911	0.0118	0.0337	0.0339
	0.2	0.1019	0.0798	0.00395	0.0109	0.0197
	0.3	0.4926	0.0491	0.0192	0.0067	0.0081
15	0.1	0.0381	0.0924	0.0144	0.0471	0.0347
	0.2	0.3332	0.0805	0.0123	0.0218	0.0206
	0.3	0.7035	0.0468	0.0431	0.0183	0.0084
20	0.1	0.0312	0.0927	0.01996	0.0389	0.0351
	0.2	0.3838	0.0699	0.0267	0.0295	0.0217
	0.3	0.7954	0.0466	0.0804	0.0198	0.0092
30	0.1	0.1948	0.0945	0.0449	0.0359	0.0354
	0.2	0.5765	0.0765	0.1183	0.0299	0.0224
	0.3	0.8956	0.0391	0.0967	0.0194	0.0096
50	0.1	0.393	0.0932	0.2313	0.0384	0.0365
	0.2	0.5708	0.0769	0.1535	0.0291	0.0218
	0.3	0.7988	0.0395	0.0757	0.0199	0.0105
			δ=10,λ=1			
5	0.1	0.0141	0.0843	0.0111	0.1362	0.0323
	0.2	0.0059	0.0618	0.0029	0.0897	0.0171
	0.3	0.0029	0.0391	0.001	0.0714	0.0057
10	0.1	0.0069	0.0721	0.0124	0.0004	0.0339
	0.2	0.0019	0.061	0.0044	0.0021	0.0202
	0.3	0.005	0.0355	0.016	0.0002	0.0076
15	0.1	0.0201	0.0802	0.0155	0.0382	0.0341
	0.2	0.0198	0.0671	0.006	0.0301	0.0203
	0.3	0.0151	0.0354	0.0458	0.0208	0.0085
20	0.1	0.0241	0.0798	0.0172	0.0365	0.0349
	0.2	0.0165	0.0614	0.0254	0.0278	0.0216
	0.3	0.0211	0.0387	0.0719	0.0198	0.0091
30	0.1	0.0429	0.0895	0.0449	0.0394	0.0357
	0.2	0.0381	0.0629	0.1227	0.0295	0.0216
	0.3	0.04	0.0339	0.0831	0.0179	0.0103
50	0.1	0.1457	0.0331	0.2286	0.0372	0.0353
	0.2	0.1411	0.0689	0.1588	0.0286	0.0228
	0.3	0.1433	0.0878	0.0711	0.0194	0.0101

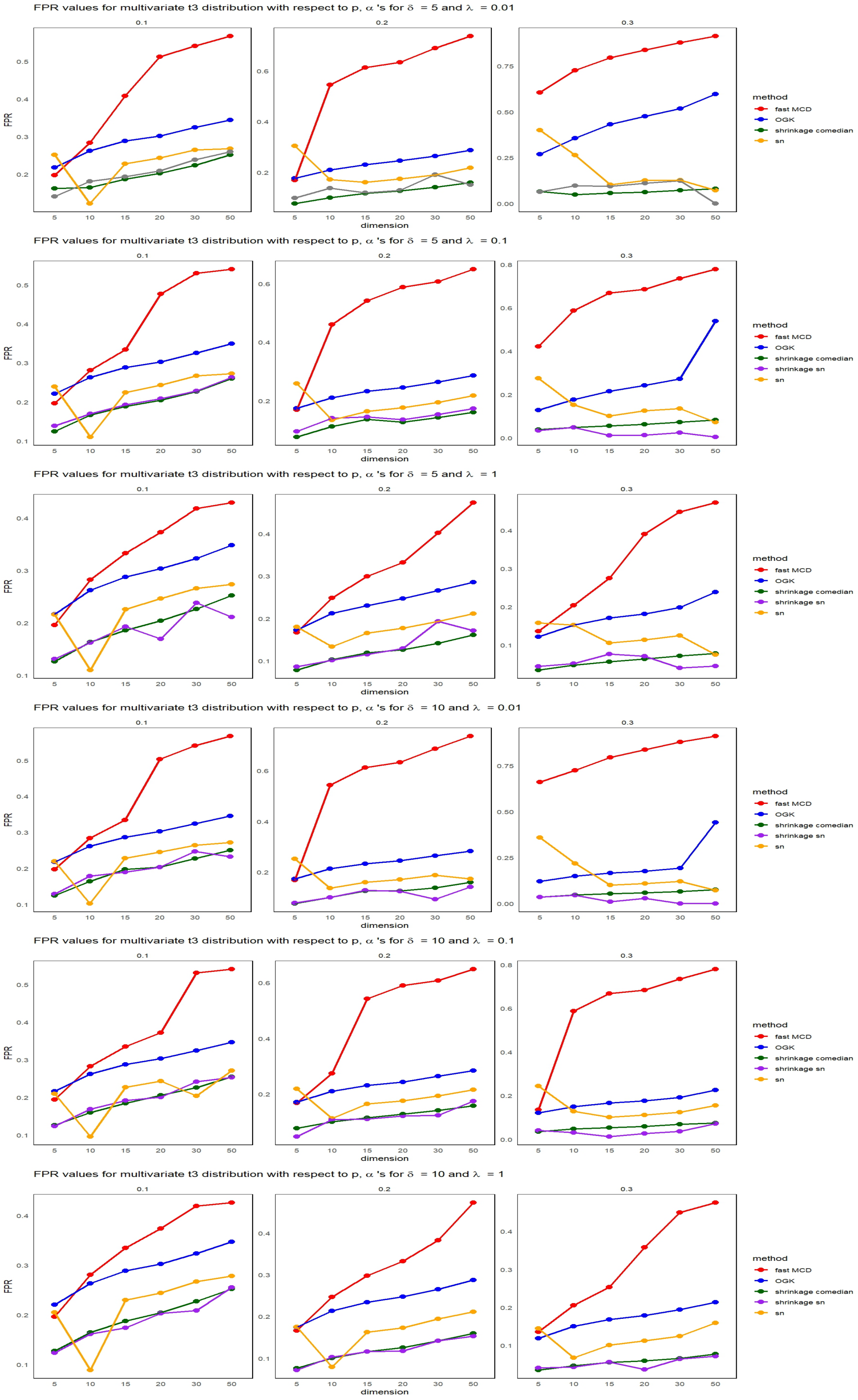


Figure A.2: (a) Figure of FPR for multivariate t_3 distribution for $n = 1000$

Table S12: Zero contamination FPR in multivariate t_3-distribution for different samples

$n = 1000$					
p	$Sh\text{-}S_n$	RMDS	S_n	MCD	OGK
5	0.1789	0.1802	0.2789	0.2313	0.25789
10	0.2262	0.2391	0.1376	0.3229	0.31578
15	0.2511	0.2718	0.2973	0.3761	0.33189
20	0.2710	0.2919	0.3205	0.4168	0.35644
30	0.2951	0.3226	0.3543	0.4819	0.37378
50	0.3168	0.3567	0.3830	0.4872	0.40729
$n = 500$					
5	0.1819	0.1888	0.2761	0.2382	0.26412
10	0.2128	0.2395	0.1280	0.3254	0.3103
15	0.2467	0.2689	0.2924	0.3796	0.34062
20	0.2724	0.2987	0.3263	0.4246	0.35864
30	0.2889	0.3231	0.3537	0.4678	0.38978
50	0.3107	0.3558	0.3828	0.4598	0.40662
$n = 100$					
5	0.1801	0.1923	0.2752	0.2528	0.2631
10	0.2288	0.2286	0.1278	0.3639	0.3018
15	0.2515	0.2667	0.3044	0.3891	0.3451
20	0.2718	0.3012	0.3168	0.3939	0.3578
30	0.2865	0.3387	0.3478	0.3538	0.3771
50	0.2978	0.3881	0.3632	0.2463	0.3981

Table S13: Multivariate t_3 distribution TPR table for $n = 100$

p	$\boldsymbol{\alpha}$	**FAST MCD**	**OGK**	**RMDS**	$\boldsymbol{S_n}$	Sh-S_n
δ=5,λ=0.01						
5	0.1	1	1	1	1	1
	0.2	0.0492	0.9874	0.996	1	1
	0.3	0.0045	0	0.2111	1	1
10	0.1	0.784	1	1	1	1
	0.2	0.0012	0.996	1	1	1
	0.3	0.0028	0	0.0641	0.7581	1
15	0.1	0.0172	1	1	1	1
	0.2	0.0026	1	1	1	1
	0.3	0.0017	0	0.0593	0.0967	1
20	0.1	0.0004	1	1	1	1
	0.2	0.0014	1	0.9916	1	1
	0.3	0.002	0.0013	0.0971	0.0937	0.9979
30	0.1	0.0008	1	1	1	1
	0.2	0.0014	1	0.9002	1	1
	0.3	0.0033	0	0.0975	0.1683	0.9585
50	0.1	0.0004	1	0.9988	1	1
	0.2	0.0014	1	0.503	1	1
	0.3	0.0012	0	0.0596	0.2849	0.8912
δ=5,λ=0.1						
5	0.1	1	1	1	1	1
	0.2	0.7826	1	0.9992	1	1
	0.3	0.0384	0.0587	0.2643	0.9996	0.9975
10	0.1	0.984	1	1	1	1
	0.2	0.0346	1	1	0.9978	1
	0.3	0.0483	0.0403	0.0915	0.8564	1
15	0.1	0.2992	1	1	1	1
	0.2	0.0388	1	1	1	1
	0.3	0.0469	0.0445	0.1108	0.5217	1
20	0.1	0.0472	1	1	1	1
	0.2	0.0316	1	0.9954	1	1
	0.3	0.0408	0.0545	0.1315	0.6395	1
30	0.1	0.0296	1	1	1	1
	0.2	0.0296	1	0.9394	1	1
	0.3	0.0361	0.1393	0.1428	0.7260	0.9684
50	0.1	0.0148	1	0.9948	1	1
	0.2	0.0226	1	0.6118	1	1
	0.3	0.0291	0.2209	0.1076	0.858	0.8833
δ=5,λ=1						
5	0.1	1	1	0.9996	1	1
	0.2	0.9560	1	0.9952	0.9996	1
	0.3	0.0025	0.9299	0.5903	0.9980	1
10	0.1	1	1	1	1	1
	0.2	0.0014	1	0.9996	0.9996	1
	0.3	0.0031	0.9953	0.5359	0.9300	1
15	0.1	0.6604	1	1	1	1
	0.2	0.0016	1	1	1	1
	0.3	0.0017	0.9997	0.5101	0.9971	1
20	0.1	0.0284	1	1	1	1
	0.2	0.0014	1	0.9996	1	1
	0.3	0.0023	0.9997	0.5367	0.9996	1
30	0.1	0.0004	1	1	1	1
	0.2	0.0018	1	0.9778	1	1
	0.3	0.0013	1	0.4260	0.9998	1
50	0.1	0.0016	1	0.9952	1	1
	0.2	0.0010	1	0.7516	1	1
	0.3	0.0019	1	0.3572	1	1

Table S14: Multivariate t_3 distribution TPR table for $n = 100$

p	$\boldsymbol{\alpha}$	**FAST MCD**	**OGK**	**RMDS**	$\boldsymbol{S_n}$	$Sh\text{-}S_n$
δ=10,λ=0.01						
5	0.1	1	1	1	1	1
	0.2	0.956	1	1	1	1
	0.3	0.0025	0.8753	1	1	1
10	0.1	1	1	1	1	1
	0.2	0.0014	1	1	1	1
	0.3	0.0031	0.792	1	1	1
15	0.1	0.6604	1	1	1	1
	0.2	0.0016	1	1	1	1
	0.3	0.0017	0.7701	0.9713	1	1
20	0.1	0.0284	1	1	1	1
	0.2	0.0014	1	0.9998	1	1
	0.3	0.0023	0.8376	0.9091	1	1
30	0.1	0.0004	1	1	1	1
	0.2	0.0018	1	0.985	1	1
	0.3	0.0013	0.8377	0.7876	1	1
50	0.1	0.0016	1	1	1	1
	0.2	0.001	1	0.8942	1	1
	0.3	0.0019	0.8569	0.5904	1	1
δ=10,λ=0.1						
5	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	0.2829	0.9881	1	1	1
10	0.1	1	1	1	1	1
	0.2	0.1148	1	1	1	1
	0.3	0.0508	1	0.9961	1	1
15	0.1	0.9488	1	1	1	1
	0.2	0.0376	1	1	1	1
	0.3	0.0413	1	0.9909	1	1
20	0.1	0.3872	1	1	1	1
	0.2	0.0338	1	1	1	1
	0.3	0.0411	1	0.942	1	1
30	0.1	0.0276	1	1	1	1
	0.2	0.0242	1	0.9974	1	1
	0.3	0.0401	1	0.7973	1	1
50	0.1	0.0216	1	0.992	1	1
	0.2	0.0216	1	0.9038	1	1
	0.3	0.0287	1	0.6307	1	1
δ=10,λ=1						
5	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	1	1	1	1	1
10	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	0.9261	1	1	1	1
15	0.1	1	1	1	1	1
	0.2	0.9936	1	1	1	1
	0.3	0.5189	1	0.9877	1	1
20	0.1	1	1	1	1	1
	0.2	0.7456	1	1	1	1
	0.3	0.3952	1	0.9028	1	1
30	0.1	0.9324	1	1	1	1
	0.2	0.3512	1	0.9672	1	1
	0.3	0.3335	1	0.8271	1	1
50	0.1	0.2872	1	0.9988	1	1
	0.2	0.256	1	0.8818	1	1
	0.3	0.25	1	0.6324	1	1

Table S15: Multivariate t_3 distribution TPR table for $n = 500$

p	$\boldsymbol{\alpha}$	**FAST MCD**	**OGK**	**RMDS**	$\boldsymbol{S_n}$	Sh-S_n
δ=5,λ=0.01						
	0.1	1	1	1	1	1
5	0.2	1	1	1	1	1
	0.3	0.0005	0	0.0089	1	1
	0.1	1	1	1	1	1
10	0.2	0.0003	1	1	1	1
	0.3	0.0004	0	0.0002	1	1
	0.1	0.67	1	1	1	1
15	0.2	0.0001	1	1	1	1
	0.3	0.0006	0	0.0003	0.7896	1
	0.1	0.0002	1	1	1	1
20	0.2	0.0003	1	1	1	1
	0.3	0.0006	0	0.0005	0.8976	1
	0.1	1	1	1	1	1
30	0.2	0.0001	1	0.3991	1	1
	0.3	0.0002	0	0.0004	0.9887	1
	0.1	0.0005	1	1	1	1
50	0.2	0.0016	1	0.0018	1	1
	0.3	0.0036	0	0.0002	1	1
δ=5,λ=0.1						
	0.1	1	1	1	1	1
5	0.2	1	1	1	1	1
	0.3	0.0142	0	0.0397	1	1
	0.1	1	1	1	1	1
10	0.2	0.0081	1	1	1	1
	0.3	0.0143	0	0.0097	1	1
	0.1	1	1	1	1	1
15	0.2	0.0082	1	1	1	1
	0.3	0.0118	0	0.0086	1	1
	0.1	0.1258	1	1	1	1
20	0.2	0.0067	1	1	1	1
	0.3	0.0109	0	0.0085	0.9876	1
	0.1	0.006	1	1	1	1
30	0.2	0.0059	1	0.6281	1	1
	0.3	0.0083	0	0.0094	0.9899	1
	0.1	0.0316	1	1	1	1
50	0.2	0.0378	1	0.0612	1	1
	0.3	0.0551	0.0009	0.01	1	1
δ=5,λ=1						
	0.1	1	1	1	1	1
5	0.2	1	1	1	1	1
	0.3	1	0.9393	0.4296	1	1
	0.1	1	1	1	1	1
10	0.2	1	1	1	1	1
	0.3	1	0.9787	0.2775	1	1
	0.1	1	1	1	1	1
15	0.2	1	1	1	1	1
	0.3	0.9255	0.9966	0.2368	1	1
	0.1	1	1	1	1	1
20	0.2	1	1	1	1	1
	0.3	0.4663	0.999	0.2225	1	1
	0.1	1	1	1	1	1
30	0.2	0.7471	1	1	1	1
	0.3	0.3301	0.99	0.2212	1	1
	0.1	0.8316	1	1	1	1
50	0.2	0.4702	1	0.7651	1	1
	0.3	0.4495	0.6456	0.2422	1	1

Table S16: Multivariate t_3 distribution TPR table for $n = 500$

p	$\boldsymbol{\alpha}$	**FAST MCD**	**OGK**	**RMDS**	$\boldsymbol{S_n}$	Sh-S_n
δ=10,λ=0.01						
	0.1	1	1	1	1	1
5	0.2	1	1	1	1	1
	0.3	0.0009	0.6517	1	1	1
	0.1	1	1	1	1	1
10	0.2	0.0003	1	1	1	1
	0.3	0.0007	0.5804	1	1	1
	0.1	1	1	1	1	1
15	0.2	0.0001	1	1	1	1
	0.3	0.0005	0.6023	1	1	1
	0.1	1	1	1	1	1
20	0.2	0.0003	1	1	1	1
	0.3	0.0004	0.6517	1	1	1
	0.1	0	1	1	1	1
30	0.2	0.0003	1	1	1	1
	0.3	0.0004	0.7242	1	1	1
	0.1	0.0016	1	1	1	1
50	0.2	0.002	1	1	1	1
	0.3	0.0033	0.8806	0.2622	1	1
δ=10,λ=0.1						
	0.1	1	1	1	1	1
5	0.2	1	1	1	1	1
	0.3	1	1	0.9999	1	1
	0.1	1	1	1	1	1
10	0.2	0.831	1	1	1	1
	0.3	0.0106	1	1	1	1
	0.1	1	1	1	1	1
15	0.2	0.0101	1	1	1	1
	0.3	0.0137	1	1	1	1
	0.1	1	1	1	1	1
20	0.2	0.0114	1	1	1	1
	0.3	0.0107	1	1	1	1
	0.1	0.0056	1	1	1	1
30	0.2	0.0053	1	1	1	1
	0.3	0.008	1	1	1	1
	0.1	0.0308	1	1	1	1
50	0.2	0.0442	1	1	1	1
	0.3	0.0538	0.4108	0.9605	1	1
δ=10,λ=1						
	0.1	1	1	1	1	1
5	0.2	1	1	1	1	1
	0.3	1	1	0.9998	1	1
	0.1	1	1	1	1	1
10	0.2	1	1	1	1	1
	0.3	1	1	1	1	1
	0.1	1	1	1	1	1
15	0.2	1	1	1	1	1
	0.3	0.9162	1	1	1	1
	0.1	1	1	1	1	1
20	0.2	1	1	1	1	1
	0.3	0.4907	1	1	1	1
	0.1	1	1	1	1	1
30	0.2	0.7832	1	1	1	1
	0.3	0.3356	1	1	1	1
	0.1	1	1	1	1	1
50	0.2	0.4752	1	1	1	1
	0.3	0.4544	1	0.9605	1	1

Table S17: Multivariate t_3 distribution TPR table for $n = 1000$

p	$\boldsymbol{\alpha}$	**FAST MCD**	**OGK**	**RMDS**	$\boldsymbol{S_n}$	Sh-S_n
δ=5,λ=0.01						
	0.1	1	1	1	1	1
5	0.2	1	1	1	1	1
	0.3	0.0034	0.0010	1	1	1
	0.1	1	1	1	1	1
10	0.2	0.0018	1	0.8891	1	1
	0.3	0.0029	0.0031	0.7789	1	1
	0.1	0.4410	1	1	1	1
15	0.2	0.0017	1	0.9111	1	1
	0.3	0.0041	0.0045	0.7991	0.8947	1
	0.1	0.0019	1	1	1	1
20	0.2	0.0019	1	0.9417	1	1
	0.3	0.0042	0.0039	0.8021	0.9105	0.9555
	0.1	0.0013	1	1	1	1
30	0.2	0.0019	1	0.9777	1	1
	0.3	0.0051	0.0056	0.8331	0.9399	0.9661
	0.1	0.0009	1	1	1	1
50	0.2	0.003	0.9109	0.9871	1	1
	0.3	0.0064	0.0033	0.8771	0.9541	0.9887
δ=5,λ=0.1						
	0.1	1	1	1	1	1
5	0.2	1	1	1	1	1
	0.3	0.0428	0.056	1	1	1
	0.1	1	1	1	1	1
10	0.2	0.0388	1	1	1	1
	0.3	0.0499	0.0402	1	1	1
	0.1	1	1	1	1	1
15	0.2	0.0412	1	1	1	1
	0.3	0.0598	0.0389	0.9522	1	1
	0.1	0.0891	1	1	1	1
20	0.2	0.0513	1	1	1	1
	0.3	0.0601	0.05	0.9451	0.9677	1
	0.1	0.0403	1	1	1	1
30	0.2	0.04489	1	1	1	1
	0.3	0.0715	0.0781	0.9789	0.9899	1
	0.1	0.0376	1	1	1	1
50	0.2	0.0519	1	1	1	1
	0.3	0.0810	0.0788	0.9881	0.9914	1
δ=5,λ=1						
	0.1	1	1	1	1	1
5	0.2	1	1	1	1	1
	0.3	1	0.998	1	1	1
	0.1	1	1	1	1	1
10	0.2	1	1	1	1	1
	0.3	1	1	1	1	1
	0.1	1	1	1	1	1
15	0.2	1	1	1	1	1
	0.3	0.8663	1	1	1	1
	0.1	1	1	1	1	1
20	0.2	1	1	1	1	1
	0.3	0.4178	1	0.9811	1	1
	0.1	1	1	1	1	1
30	0.2	1	1	1	1	1
	0.3	0.4862	1	0.9899	1	1
	0.1	0.9661	1	1	1	1
50	0.2	0.5028	1	1	1	1
	0.3	0.4961	0.6712	0.9912	1	1

Table S18: Multivariate t_3 distribution TPR table for $n = 1000$

p	$\boldsymbol{\alpha}$	**FAST MCD**	**OGK**	**RMDS**	$\boldsymbol{S_n}$	Sh-S_n
δ=10,λ=0.01						
5	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	0.0028	0.9891	1	1	1
10	0.1	1	1	1	1	1
	0.2	0.0021	1	1	1	1
	0.3	0.0033	1	1	1	1
15	0.1	1	1	1	1	1
	0.2	0.0020	1	1	1	1
	0.3	0.0044	1	1	1	1
20	0.1	0.0712	1	1	1	1
	0.2	0.0029	1	1	1	1
	0.3	0.0041	1	0.9663	1	1
30	0.1	0.0019	1	1	1	1
	0.2	0.0026	1	1	1	1
	0.3	0.0051	1	0.9799	1	1
50	0.1	0.0019	1	1	1	1
	0.2	0.0033	1	1	1	1
	0.3	0.0061	0.0077	0.9899	1	1
δ=10,λ=0.1						
5	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	1	1	1	1	1
10	0.1	1	1	1	1	1
	0.2	0.8991	1	1	1	1
	0.3	0.0546	1	1	1	1
15	0.1	1	1	1	1	1
	0.2	0.0582	1	1	1	1
	0.3	0.0674	1	1	1	1
20	0.1	1	1	1	1	1
	0.2	0.0439	1	1	1	1
	0.3	0.0668	1	1	1	1
30	0.1	0.0532	1	1	1	1
	0.2	0.0656	1	1	1	1
	0.3	0.0829	1	0.9988	1	1
50	0.1	0.0582	1	1	1	1
	0.2	0.0759	1	1	1	1
	0.3	0.0912	0.1708	0.9991	1	1
δ=10,λ=1						
5	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	1	1	1	1	1
10	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	1	1	1	1	1
15	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	0.9810	1	1	1	1
20	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	0.5912	1	1	1	1
30	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	0.4881	1	1	1	1
50	0.1	1	1	1	1	1
	0.2	0.5219	1	1	1	1
	0.3	0.5072	1	1	1	1

Table S19: Multivariate t_3 distribution FPR table for $n = 100$

p	$\boldsymbol{\alpha}$	**FAST MCD**	**OGK**	**RMDS**	$\boldsymbol{S_n}$	$Sh - S_n$
			δ=5,λ=0.01			
5	0.1	0.2108	0.1859	0.1304	0.2464	0.1737
	0.2	0.4374	0.1713	0.0849	0.3323	0.1306
	0.3	0.6323	0.2748	0.0517	0.5	0.0806
10	0.1	0.3398	0.1918	0.1736	0.1183	0.2261
	0.2	0.5456	0.2785	0.1229	0.1746	0.1715
	0.3	0.6401	0.3292	0.0846	0.3074	0.1145
15	0.1	0.443	0.0977	0.1978	0.2516	0.2522
	0.2	0.5188	0.1849	0.1526	0.1983	0.1942
	0.3	0.599	0.3322	0.1266	0.1642	0.1353
20	0.1	0.4328	0.1007	0.2239	0.2679	0.2702
	0.2	0.4958	0.2906	0.1758	0.2157	0.2144
	0.3	0.5699	0.333	0.1549	0.1718	0.1495
30	0.1	0.3786	0.1104	0.2615	0.2856	0.2944
	0.2	0.433	0.2946	0.2377	0.2349	0.234
	0.3	0.4985	0.3387	0.1466	0.1901	0.1643
50	0.1	0.276	0.1226	0.3272	0.3147	0.3205
	0.2	0.312	0.2985	0.2535	0.2579	0.2553
	0.3	0.3566	0.3387	0.1255	0.2177	0.1887
			δ=5,λ=0.1			
5	0.1	0.2117	0.1821	0.1272	0.245	0.1736
	0.2	0.2284	0.1677	0.0847	0.2812	0.131
	0.3	0.4514	0.1768	0.0543	0.3572	0.0841
10	0.1	0.3115	0.19	0.1689	0.222	0.1123
	0.2	0.4932	0.1712	0.1267	0.173	0.1355
	0.3	0.5904	0.1939	0.0867	0.1887	0.115
15	0.1	0.4079	0.1932	0.2042	0.2497	0.2471
	0.2	0.5	0.1745	0.1497	0.1991	0.197
	0.3	0.5722	0.2023	0.1222	0.1425	0.1314
20	0.1	0.4247	0.2962	0.2236	0.2666	0.2648
	0.2	0.4831	0.2773	0.1764	0.2154	0.2096
	0.3	0.5478	0.1742	0.1576	0.1581	0.1487
30	0.1	0.3738	0.2985	0.2576	0.2935	0.2927
	0.2	0.4203	0.2792	0.2304	0.238	0.2301
	0.3	0.4785	0.2171	0.1478	0.1714	0.1685
50	0.1	0.272	0.3197	0.3236	0.3204	0.3175
	0.2	0.3048	0.2923	0.2518	0.2616	0.2563
	0.3	0.3431	0.2138	0.1281	0.1896	0.188
			δ=5,λ=1			
5	0.1	0.2145	0.1804	0.1296	0.2278	0.1717
	0.2	0.1924	0.155	0.0888	0.2155	0.1313
	0.3	0.6319	0.1355	0.0581	0.2093	0.0818
10	0.1	0.3146	0.2897	0.1719	0.0991	0.2241
	0.2	0.5449	0.1632	0.1286	0.0965	0.1736
	0.3	0.639	0.1387	0.0905	0.0923	0.1188
15	0.1	0.3688	0.2925	0.2031	0.2509	0.2485
	0.2	0.5196	0.2674	0.1535	0.2016	0.2008
	0.3	0.5994	0.1401	0.1258	0.1405	0.1353
20	0.1	0.4293	0.294	0.2223	0.267	0.2665
	0.2	0.4957	0.2689	0.1787	0.2163	0.2138
	0.3	0.5698	0.1837	0.1611	0.1554	0.1488
30	0.1	0.3793	0.2978	0.2583	0.2943	0.2909
	0.2	0.4342	0.2695	0.2355	0.239	0.2378
	0.3	0.4993	0.1815	0.1657	0.1682	0.164
50	0.1	0.2765	0.3215	0.3289	0.32	0.3181
	0.2	0.3122	0.2792	0.2511	0.2593	0.2594
	0.3	0.3563	0.1857	0.1523	0.1909	0.1855

Table S20: Multivariate t_3 distribution FPR table for $n = 100$

p	$\boldsymbol{\alpha}$	**FAST MCD**	**OGK**	**RMDS**	$\boldsymbol{S_n}$	$Sh - S_n$
δ=10,λ=0.01						
5	0.1	0.2145	0.1811	0.1259	0.2188	0.1432
	0.2	0.1924	0.1689	0.0848	0.2458	0.1043
	0.3	0.6319	0.1058	0.0534	0.3683	0.0637
10	0.1	0.3146	0.1932	0.1712	0.0926	0.1789
	0.2	0.5449	0.1826	0.124	0.1182	0.1437
	0.3	0.639	0.1626	0.0823	0.2063	0.0909
15	0.1	0.3688	0.1965	0.2023	0.245	0.2019
	0.2	0.5196	0.1877	0.1507	0.1969	0.1547
	0.3	0.5994	0.1608	0.1219	0.1417	0.1019
20	0.1	0.4293	0.2135	0.224	0.2673	0.2174
	0.2	0.4957	0.1934	0.1773	0.2123	0.1685
	0.3	0.5698	0.1636	0.1425	0.1522	0.1162
30	0.1	0.3793	0.2047	0.2584	0.2888	0.2392
	0.2	0.4342	0.1907	0.2318	0.2334	0.1895
	0.3	0.4993	0.1675	0.1465	0.1657	0.1288
50	0.1	0.2765	0.2142	0.3185	0.314	0.2619
	0.2	0.3122	0.2048	0.2405	0.2544	0.2089
	0.3	0.3563	0.167	0.1208	0.182	0.1426
δ=10,λ=0.1						
5	0.1	0.2118	0.1792	0.1286	0.2165	0.1398
	0.2	0.1742	0.1618	0.0869	0.2241	0.1035
	0.3	0.3879	0.1463	0.0514	0.2705	0.0603
10	0.1	0.3067	0.192	0.1742	0.0981	0.1787
	0.2	0.4764	0.1743	0.1225	0.1039	0.1381
	0.3	0.5912	0.105	0.0882	0.1368	0.0893
15	0.1	0.3357	0.2027	0.2004	0.2487	0.2012
	0.2	0.5002	0.1789	0.1514	0.1956	0.1576
	0.3	0.5758	0.1502	0.1311	0.1385	0.103
20	0.1	0.386	0.2018	0.2212	0.2661	0.2191
	0.2	0.4829	0.2836	0.1793	0.2095	0.1725
	0.3	0.5482	0.253	0.1619	0.1589	0.1126
30	0.1	0.3728	0.2506	0.2572	0.2876	0.2373
	0.2	0.423	0.1897	0.2364	0.2349	0.1864
	0.3	0.4769	0.153	0.1499	0.1681	0.128
50	0.1	0.272	0.3134	0.3233	0.318	0.2598
	0.2	0.3054	0.2092	0.2422	0.2533	0.2055
	0.3	0.3434	0.1616	0.1321	0.1845	0.1447
δ=10,λ=1						
5	0.1	0.2088	0.1826	0.1273	0.2139	0.138
	0.2	0.1789	0.1551	0.0837	0.1885	0.1035
	0.3	0.1368	0.1331	0.0541	0.1754	0.0637
10	0.1	0.3123	0.186	0.1756	0.0935	0.1749
	0.2	0.2608	0.1633	0.1245	0.0821	0.138
	0.3	0.2154	0.1387	0.0808	0.0794	0.0874
15	0.1	0.3332	0.2944	0.1992	0.2467	0.2048
	0.2	0.2623	0.169	0.1447	0.1989	0.163
	0.3	0.3347	0.1393	0.1209	0.139	0.1029
20	0.1	0.3182	0.2953	0.2256	0.267	0.22
	0.2	0.2928	0.1786	0.1757	0.2134	0.1714
	0.3	0.3716	0.1419	0.1489	0.1542	0.1129
30	0.1	0.2682	0.2382	0.2558	0.2897	0.2362
	0.2	0.331	0.1841	0.2302	0.2329	0.1891
	0.3	0.3348	0.1419	0.163	0.1693	0.129
50	0.1	0.2386	0.2911	0.3171	0.3179	0.2623
	0.2	0.2403	0.2779	0.2397	0.2598	0.2092
	0.3	0.2415	0.1469	0.1269	0.1901	0.1474

Table S21: Multivariate t_3 distribution FPR table for $n = 500$

p	$\boldsymbol{\alpha}$	**FAST MCD**	**OGK**	**RMDS**	$\boldsymbol{S_n}$	$Sh\text{-}S_n$
δ=5,λ=0.01						
	0.1	0.1597	0.1784	0.1275	0.2345	0.1215
5	0.2	0.1378	0.1644	0.0861	0.2759	0.0601
	0.3	0.6089	0.3544	0.0495	0.3951	0.0359
	0.1	0.2001	0.1831	0.1701	0.1207	0.1688
10	0.2	0.4999	0.1682	0.1182	0.1494	0.1084
	0.3	0.7089	0.3185	0.0731	0.2459	0.0135
	0.1	0.1971	0.1844	0.196	0.1778	0.1723
15	0.2	0.5861	0.1683	0.1368	0.1626	0.0907
	0.3	0.7799	0.3009	0.0876	0.2389	0.0426
	0.1	0.4164	0.2845	0.2136	0.2469	0.2006
20	0.2	0.5851	0.1693	0.1515	0.1746	0.1229
	0.3	0.8189	0.2934	0.0981	0.1129	0.0328
	0.1	0.5883	0.2869	0.2398	0.2668	0.2296
30	0.2	0.6102	0.1706	0.1741	0.1929	0.1103
	0.3	0.8713	0.2855	0.1122	0.1235	0.0311
	0.1	0.5198	0.2897	0.2712	0.2925	0.2321
50	0.2	0.6687	0.1727	0.2005	0.2152	0.1589
	0.3	0.8669	0.2804	0.1369	0.1368	0.046
δ=5,λ=0.1						
	0.1	0.1543	0.1787	0.1278	0.2279	0.1288
5	0.2	0.1218	0.1599	0.0862	0.2433	0.0734
	0.3	0.4123	0.08	0.0516	0.2564	0.0354
	0.1	0.1907	0.1796	0.1697	0.1066	0.1689
10	0.2	0.4201	0.1613	0.1181	0.1289	0.1094
	0.3	0.5981	0.1098	0.0718	0.1449	0.032
	0.1	0.2201	0.1835	0.1952	0.2269	0.1842
15	0.2	0.4919	0.1627	0.1386	0.1645	0.1112
	0.3	0.6867	0.1249	0.0863	0.1035	0.041
	0.1	0.3656	0.2827	0.214	0.2462	0.2056
20	0.2	0.5337	0.1628	0.1525	0.1768	0.1125
	0.3	0.6819	0.1381	0.0987	0.1131	0.0326
	0.1	0.4449	0.2856	0.2391	0.1647	0.2293
30	0.2	0.5811	0.1636	0.1737	0.1989	0.1463
	0.3	0.7045	0.1519	0.1138	0.1246	0.0427
	0.1	0.5089	0.2887	0.2699	0.2932	0.2531
50	0.2	0.5899	0.1646	0.1999	0.2155	0.1403
	0.3	0.7069	0.164	0.139	0.1389	0.0374
δ=5,λ=1						
	0.1	0.148	0.1743	0.1304	0.2108	0.1426
5	0.2	0.1287	0.1531	0.0868	0.1789	0.0732
	0.3	0.0998	0.0739	0.0544	0.1521	0.0397
	0.1	0.1909	0.1776	0.17	0.2375	0.1767
10	0.2	0.1689	0.1538	0.1196	0.1508	0.1395
	0.3	0.1299	0.0705	0.0773	0.1167	0.041
	0.1	0.2105	0.1807	0.1971	0.2274	0.1737
15	0.2	0.1985	0.1557	0.1399	0.1676	0.1186
	0.3	0.1599	0.031	0.0926	0.1094	0.0117
	0.1	0.2221	0.2794	0.2155	0.2492	0.219
20	0.2	0.1885	0.1568	0.1563	0.1799	0.1276
	0.3	0.1597	0.0311	0.1024	0.1142	0.0122
	0.1	0.2305	0.281	0.2409	0.2649	0.2201
30	0.2	0.2995	0.1575	0.1765	0.1938	0.1372
	0.3	0.2922	0.0317	0.1185	0.1268	0.0141
	0.1	0.4095	0.2834	0.2699	0.2883	0.2422
50	0.2	0.4436	0.1576	0.2025	0.2133	0.1512
	0.3	0.4482	0.0342	0.1436	0.1399	0.0115

Table S22: Multivariate t_3 distribution FPR table for $n = 500$

p	$\boldsymbol{\alpha}$	**FAST MCD**	**OGK**	**RMDS**	$\boldsymbol{S_n}$	Sh-S_n
δ=10,λ=0.01						
	0.1	0.1489	0.0181	0.1283	0.3077	0.1241
5	0.2	0.1285	0.162	0.0848	0.2281	0.071
	0.3	0.6075	0.151	0.0513	0.3079	0.0359
	0.1	0.188	0.1814	0.1704	0.0989	0.1617
10	0.2	0.4981	0.1071	0.1162	0.1098	0.1017
	0.3	0.7098	0.147	0.0742	0.1819	0.03
	0.1	0.2167	0.1831	0.1968	0.2284	0.1861
15	0.2	0.5781	0.1773	0.1382	0.1629	0.1109
	0.3	0.7641	0.1452	0.0864	0.1013	0.0405
	0.1	0.2369	0.1854	0.2123	0.2487	0.2081
20	0.2	0.5957	0.1779	0.1521	0.1787	0.1267
	0.3	0.8188	0.1436	0.0977	0.1089	0.0427
	0.1	0.4767	0.186	0.2391	0.2689	0.2298
30	0.2	0.6402	0.1793	0.1744	0.1998	0.1309
	0.3	0.8693	0.1426	0.1119	0.1212	0.0477
	0.1	0.5234	0.1901	0.2693	0.289	0.2637
50	0.2	0.6589	0.1808	0.2017	0.2178	0.1588
	0.3	0.8638	0.1429	0.1363	0.1389	0.0539
δ=10,λ=0.1						
	0.1	0.1466	0.1771	0.129	0.2139	0.1249
5	0.2	0.1189	0.1563	0.0844	0.1976	0.1109
	0.3	0.0989	0.1412	0.0514	0.2139	0.0536
	0.1	0.1967	0.1784	0.1713	0.0845	0.1609
10	0.2	0.2067	0.1638	0.1171	0.0989	0.0939
	0.3	0.5861	0.1391	0.0725	0.1099	0.0416
	0.1	0.2299	0.1817	0.1954	0.1178	0.1813
15	0.2	0.4939	0.166	0.1385	0.161	0.1119
	0.3	0.6844	0.1363	0.0869	0.2256	0.0519
	0.1	0.2477	0.1822	0.215	0.2438	0.2155
20	0.2	0.5627	0.168	0.1544	0.1749	0.1379
	0.3	0.6878	0.1354	0.0985	0.1092	0.0421
	0.1	0.4459	0.1839	0.2386	0.2629	0.2253
30	0.2	0.5809	0.1718	0.1744	0.1922	0.1413
	0.3	0.7143	0.1365	0.1147	0.1235	0.0445
	0.1	0.5076	0.1875	0.2696	0.2907	0.2419
50	0.2	0.5859	0.1695	0.198	0.2112	0.1518
	0.3	0.7131	0.1375	0.1361	0.1398	0.055
δ=10,λ=1						
	0.1	0.1474	0.174	0.129	0.1978	0.1349
5	0.2	0.1256	0.1496	0.0844	0.1592	0.1105
	0.3	0.0999	0.1303	0.0514	0.1928	0.046
	0.1	0.1879	0.1707	0.1713	0.0898	0.1673
10	0.2	0.1688	0.1559	0.1171	0.0933	0.0941
	0.3	0.1319	0.1318	0.0725	0.0689	0.0478
	0.1	0.2189	0.1787	0.1954	0.2256	0.1895
15	0.2	0.1777	0.1567	0.1385	0.1634	0.1116
	0.3	0.1676	0.1322	0.0869	0.1014	0.0548
	0.1	0.2421	0.2802	0.215	0.2434	0.2058
20	0.2	0.2079	0.1587	0.1544	0.1769	0.1203
	0.3	0.2251	0.1325	0.0985	0.1106	0.0457
	0.1	0.2795	0.2281	0.2386	0.2637	0.2203
30	0.2	0.2589	0.1591	0.1744	0.1989	0.1408
	0.3	0.2914	0.1325	0.1147	0.1229	0.0498
	0.1	0.3917	0.2832	0.2696	0.2863	0.2416
50	0.2	0.4627	0.1609	0.198	0.2114	0.1597
	0.3	0.4473	0.1325	0.1361	0.1539	0.051

Table S23: Zero contaminated multivariate exponential distribution FPR table for different sample size

p	$Sh\text{-}S_n$	**RMDS**	S_n	**FAST MCD**	**OGK**
			$n = 1000$		
5	0.0004	0.0008	0.009	0.00278	0.02678
10	0.0003	0.0011	0	0.00284	0.02823
15	0.0003	0.0018	0.0011	0.00385	0.02943
20	0.0002	0.0019	0.0009	0.00416	0.0322
30	0.0002	0.0020	0.0017	0.00767	0.03378
50	0	0.0049	0.0018	0.01891	0.0411
			$n = 500$		
5	0.0004	0.0018	0.00196	0.0123	0.02687
10	0.0004	0.0026	0	0.0091	0.03108
15	0.0003	0.0037	0.0005	0.0138	0.0389
20	0.0003	0.003	0.0006	0.0189	0.03612
30	0	0.0089	0.0005	0.0376	0.03996
50	0	0.0047	0.0004	0.0818	0.0494
			$n = 100$		
5	0.0009	0.0062	0.0023	0.1233	0.0411
10	0.0003	0.0093	0	0.1775	0.0512
15	0.0003	0.0167	0.0003	0.2283	0.0615
20	0	0.0246	0.0006	0.2619	0.0714
30	0	0.0713	0.0005	0.381	0.0816
50	0	0.3390	0.002	0.2567	0.1117

Table S24: Multivariate Exponential distribution TPR table for $n = 100$

p	$\boldsymbol{\alpha}$	**FAST MCD**	**OGK**	**RMDS**	$\boldsymbol{S_n}$	Sh-S_n
δ=5,λ=0.01						
5	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	0	0.996	1	1	1
10	0.1	1	1	1	1	1
	0.2	0.24	1	1	1	1
	0.3	0	1	1	1	1
15	0.1	1	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	1	0.976	1	1
20	0.1	1	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	1	0.914	1	1
30	0.1	0.064	1	1	1	1
	0.2	0	1	0.9742	1	1
	0.3	0	1	0.9159	1	1
50	0.1	0	1	0.976	1	1
	0.2	0	1	0.9152	1	1
	0.3	0	1	0.8184	1	1
δ=5,λ=0.1						
5	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	0.948	1	1	1	1
10	0.1	1	1	1	1	1
	0.2	0.992	1	1	1	1
	0.3	0	1	0.9999	1	1
15	0.1	1	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	1	0.9384	1	1
20	0.1	1	1	1	1	1
	0.2	0	1	0.9958	1	1
	0.3	0.0001	1	0.9125	1	1
30	0.1	0.776	1	1	1	1
	0.2	0	1	0.9722	1	1
	0.3	0.0001	1	0.8901	1	1
50	0.1	0	1	0.9744	1	1
	0.2	0	1	0.9414	1	1
	0.3	0.0001	1	0.8075	1	1
δ=5,λ=1						
5	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	1	1	0.996	1	1
10	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	1	1	0.9988	1	1
15	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	0.8456	1	0.9677	1	1
20	0.1	1	1	1	1	1
	0.2	0.9968	1	0.9888	1	1
	0.3	0.3811	1	0.9207	1	1
30	0.1	1	1	0.9976	1	1
	0.2	0.4804	1	0.9772	1	1
	0.3	0.3367	1	0.8776	1	1
50	0.1	0.5316	1	0.9664	1	1
	0.2	0.2748	1	0.918	1	1
	0.3	0.2637	1	0.7952	1	1

Table S25: Multivariate Exponential distribution TPR table for $n = 100$

p	$\boldsymbol{\alpha}$	**FAST MCD**	**OGK**	**RMDS**	$\boldsymbol{S_n}$	Sh-S_n
			δ=10,λ=0.01			
	0.1	1	1	1	1	1
5	0.2	1	1	1	1	1
	0.3	0.88	1	0.996	1	1
	0.1	1	1	1	1	1
10	0.2	1	1	1	1	1
	0.3	0	1	0.996	1	1
	0.1	1	1	1	1	1
15	0.2	0	1	1	1	1
	0.3	0	1	0.9707	1	1
	0.1	1	1	1	1	1
20	0.2	0	1	1	1	1
	0.3	0	1	0.9356	1	1
	0.1	0.876	1	1	1	1
30	0.2	0	1	0.98	1	1
	0.3	0	1	0.828	1	1
	0.1	0	1	1	1	1
50	0.2	0	1	0.9136	1	1
	0.3	0	1	0.5979	1	1
			δ=10,λ=0.1			
	0.1	1	1	1	1	1
5	0.2	1	1	1	1	1
	0.3	1	1	1	1	1
	0.1	1	1	1	1	1
10	0.2	1	1	0.996	1	1
	0.3	0	1	0.9983	1	1
	0.1	1	1	1	1	1
15	0.2	0.872	1	1	1	1
	0.3	0	1	0.972	1	1
	0.1	1	1	1	1	1
20	0.2	0	1	0.996	1	1
	0.3	0	1	0.9181	1	1
	0.1	1	1	1	1	1
30	0.2	0	1	0.9718	1	1
	0.3	0	1	0.8599	1	1
	0.1	0.0004	1	0.9668	1	1
50	0.2	0	1	0.917	1	1
	0.3	0.0001	1	0.8609	1	1
			δ=10,λ=1			
	0.1	1	1	1	1	1
5	0.2	1	1	1	1	1
	0.3	1	1	0.9981	1	1
	0.1	1	1	1	1	1
10	0.2	1	1	1	1	1
	0.3	1	1	0.9915	1	1
	0.1	1	1	1	1	1
15	0.2	1	1	0.996	1	1
	0.3	0.798	1	0.9577	1	1
	0.1	1	1	1	1	1
20	0.2	0.9964	1	0.9978	1	1
	0.3	0.3948	1	0.9332	1	1
	0.1	1	1	1	1	1
30	0.2	0.4894	1	0.9526	1	1
	0.3	0.3141	1	0.858	1	1
	0.1	0.54	1	0.9792	1	1
50	0.2	0.282	1	0.9306	1	1
	0.3	0.2611	1	0.8748	1	1

Table S26: Multivariate Exponential distribution TPR table for $n = 500$

p	$\boldsymbol{\alpha}$	**FAST MCD**	**OGK**	**RMDS**	$\boldsymbol{S_n}$	Sh-S_n
δ=5,λ=0.01						
5	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	0	1	0.9999	0.9997	1
10	0.1	1	1	1	1	1
	0.2	0.7698	1	1	1	1
	0.3	0	1	0.9899	1	1
15	0.1	1	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	1	0.98	1	1
20	0.1	1	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	1	0.9739	0.9992	1
30	0.1	0.72	1	1	1	1
	0.2	0	1	0.9998	1	1
	0.3	0	1	0.9799	1	1
50	0.1	0	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	1	0.9718	1	1
δ=5,λ=0.1						
5	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	1	1	1	1	1
10	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	0	1	0.9991	1	1
15	0.1	1	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	1	0.9999	1	1
20	0.1	0.0114	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	1	0.9889	1	1
30	0.1	1	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	1	0.9891	1	1
50	0.1	0	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	1	0.9779	1	1
δ=5,λ=1						
5	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	1	1	1	1	1
10	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	1	1	1	1	1
15	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	1	1	1	1	1
20	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	0.6918	1	1	1	1
30	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	0.05	1	1	1	1
50	0.1	1	1	1	1	1
	0.2	0.2442	1	1	1	1
	0.3	0.1214	1	1	1	1

Table S27: Multivariate Exponential distribution TPR table for $n = 500$

p	$\boldsymbol{\alpha}$	**FAST MCD**	**OGK**	**RMDS**	$\boldsymbol{S_n}$	Sh-S_n
δ=10,λ=0.01						
	0.1	1	1	1	1	1
5	0.2	1	1	1	1	1
	0.3	1	1	1	1	1
	0.1	1	1	1	1	1
10	0.2	1	1	1	1	1
	0.3	0	1	1	1	1
	0.1	1	1	1	1	1
15	0.2	0	1	1	1	1
	0.3	0	1	1	1	1
	0.1	1	1	1	1	1
20	0.2	0	1	1	1	1
	0.3	0	1	0.9999	1	1
	0.1	1	1	1	1	1
30	0.2	0	1	1	1	1
	0.3	0	1	0.9991	1	1
	0.1	0	1	1	1	1
50	0.2	0	1	1	1	1
	0.3	0	1	0.9910	1	1
δ=10,λ=0.1						
	0.1	1	1	1	1	1
5	0.2	1	1	1	1	1
	0.3	1	1	1	1	1
	0.1	1	1	1	1	1
10	0.2	1	1	1	1	1
	0.3	0	1	1	1	1
	0.1	1	1	1	1	1
15	0.2	1	1	1	1	1
	0.3	0	1	1	1	1
	0.1	1	1	1	1	1
20	0.2	0	1	1	1	1
	0.3	0	1	1	1	1
	0.1	1	1	1	1	1
30	0.2	0	1	1	1	1
	0.3	0	1	1	1	1
	0.1	0.39	1	1	1	1
50	0.2	0	1	1	1	1
	0.3	0	1	1	1	1
δ=10,λ=1						
	0.1	1	1	1	1	1
5	0.2	1	1	1	1	1
	0.3	1	1	1	1	1
	0.1	1	1	1	1	1
10	0.2	1	1	1	1	1
	0.3	1	1	1	1	1
	0.1	1	1	1	1	1
15	0.2	1	1	1	1	1
	0.3	1	1	1	1	1
	0.1	1	1	1	1	1
20	0.2	1	1	1	1	1
	0.3	0.7519	1	1	1	1
	0.1	1	1	1	1	1
30	0.2	1	1	1	1	1
	0.3	0.0678	1	1	1	1
	0.1	1	1	1	1	1
50	0.2	0.2101	1	1	1	1
	0.3	0.09912	1	1	1	1

Table S28: Multivariate Exponential distribution TPR table for $n = 1000$

p	$\boldsymbol{\alpha}$	**FAST MCD**	**OGK**	**RMDS**	$\boldsymbol{S_n}$	Sh-S_n
			δ=5,λ=0.01			
5	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	0	1	1	1	1
10	0.1	1	1	1	1	1
	0.2	0.92	1	1	1	1
	0.3	0	1	1	1	1
15	0.1	1	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	1	1	1	1
20	0.1	1	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	1	1	1	1
30	0.1	0.44	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	1	1	1	1
50	0.1	0	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	1	1	1	1
			δ=5,λ=0.1			
5	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	1	1	1	1	1
10	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	1	1	1	1	1
15	0.1	1	1	1	1	1
	0.2	0	1	1	1	1
	0.3	1	1	1	1	1
20	0.1	1	1	1	1	1
	0.2	0	1	1	1	1
	0.3	1	1	1	1	1
30	0.1	1	1	1	1	1
	0.2	0	1	1	1	1
	0.3	1	1	1	1	1
50	0.1	0	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	1	1	1	1
			δ=5,λ=1			
5	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	1	1	1	1	1
10	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	1	1	1	1	1
15	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	1	1	1	1	1
20	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	0.9451	1	1	1	1
30	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	0.0987	1	1	1	1
50	0.1	1	1	1	1	1
	0.2	0.7109	1	1	1	1
	0.3	0.0518	1	1	1	1

Table S29: Multivariate Exponential distribution TPR table for $n = 1000$

p	$\boldsymbol{\alpha}$	**FAST MCD**	**OGK**	**RMDS**	$\boldsymbol{S_n}$	Sh-S_n
			δ=10,λ=0.01			
5	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	1	1	1	1	1
10	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	0	1	1	1	1
15	0.1	1	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	1	1	1	1
20	0.1	1	1	1	1	1
	0.2	0	1	1	1	1
	0.3	1	1	1	1	1
30	0.1	1	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	1	1	1	1
50	0.1	0	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	1	1	1	1
			δ=10,λ=0.1			
5	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	1	1	1	1	1
10	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	0	1	1	1	1
15	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	0	1	1	1	1
20	0.1	1	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	1	1	1	1
30	0.1	1	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	1	1	1	1
50	0.1	0	1	1	1	1
	0.2	0	1	1	1	1
	0.3	0	1	1	1	1
			δ=10,λ=1			
5	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	1	1	1	1	1
10	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	1	1	1	1	1
15	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	1	1	1	1	1
20	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	0.8567	1	1	1	1
30	0.1	1	1	1	1	1
	0.2	1	1	1	1	1
	0.3	0.0989	1	1	1	1
50	0.1	1	1	1	1	1
	0.2	0.7192	1	1	1	1
	0.3	0.1189	1	1	1	1

Table S30: Multivariate exponential distribution FPR table for $n = 100$

p	$\boldsymbol{\alpha}$	**FAST MCD**	**OGK**	**RMDS**	$\boldsymbol{S_n}$	$Sh\text{-}S_n$
δ=5,λ=0.01						
5	0.1	0.0712	0.0345	0.0008	0.0421	0.0036
	0.2	0.0375	0.037	0.0001	0.0952	0.0021
	0.3	0.6319	0.04	0.0011	0.2836	0.0003
10	0.1	0.116	0.0459	0.0017	0	0.0032
	0.2	0.4113	0.0577	0.0002	0.0005	0.0028
	0.3	0.6393	0.0557	0.0071	0.0282	0.0009
15	0.1	0.1516	0.0525	0.0023	0.0175	0.0049
	0.2	0.5049	0.062	0.0033	0.0245	0.0029
	0.3	0.5986	0.0681	0.0677	0.0221	0.001
20	0.1	0.1813	0.0577	0.0046	0.0204	0.0043
	0.2	0.4914	0.0663	0.0256	0.0269	0.0032
	0.3	0.5711	0.0722	0.0703	0.0245	0.0014
30	0.1	0.3636	0.0677	0.0287	0.0234	0.0039
	0.2	0.4369	0.0759	0.1171	0.0273	0.0031
	0.3	0.5	0.0841	0.0695	0.028	0.0025
50	0.1	0.2775	0.0836	0.2094	0.0268	0.0051
	0.2	0.3125	0.084	0.164	0.0286	0.0042
	0.3	0.3571	0.0958	0.0647	0.0284	0.0023
δ=5,λ=0.1						
5	0.1	0.0743	0.0353	0.0004	0.0384	0.0031
	0.2	0.0334	0.0321	0.0005	0.0665	0.0015
	0.3	0.0365	0.0317	0.0027	0.1455	0.0003
10	0.1	0.1216	0.0406	0.0012	0	0.004
	0.2	0.0682	0.0442	0.0003	0.0001	0.0024
	0.3	0.5965	0.0409	0.013	0.0011	0.0008
15	0.1	0.1493	0.0508	0.002	0.0176	0.004
	0.2	0.4682	0.0487	0.0048	0.0176	0.0035
	0.3	0.5853	0.0454	0.0421	0.0119	0.0009
20	0.1	0.1858	0.0557	0.0044	0.0194	0.0042
	0.2	0.4718	0.0525	0.0203	0.019	0.0029
	0.3	0.5659	0.054	0.0801	0.0133	0.0017
30	0.1	0.2303	0.065	0.0264	0.0232	0.0054
	0.2	0.4308	0.0614	0.1096	0.0202	0.0032
	0.3	0.4995	0.0595	0.0905	0.0124	0.0021
50	0.1	0.2768	0.0842	0.2218	0.0273	0.0049
	0.2	0.3125	0.0703	0.145	0.0236	0.0034
	0.3	0.3568	0.0637	0.0745	0.0162	0.0023
δ=5,λ=1						
5	0.1	0.0733	0.0315	0.0007	0.0351	0.0026
	0.2	0.0347	0.0235	0.0002	0.0241	0.0012
	0.3	0.0091	0.0162	0.0006	0.0237	0.0005
10	0.1	0.1202	0.0403	0.0008	0	0.0036
	0.2	0.0691	0.0314	0.0006	0	0.0023
	0.3	0.0243	0.0226	0.0026	0	0.0005
15	0.1	0.1496	0.0468	0.0022	0.0158	0.0034
	0.2	0.0891	0.0363	0.0042	0.0107	0.0029
	0.3	0.0723	0.0256	0.0478	0.0063	0.001
20	0.1	0.1869	0.0543	0.0048	0.0168	0.0036
	0.2	0.1045	0.0417	0.0254	0.0116	0.0028
	0.3	0.2146	0.0319	0.0708	0.0064	0.0016
30	0.1	0.192	0.0644	0.0249	0.0192	0.0048
	0.2	0.2307	0.0517	0.117	0.0139	0.0031
	0.3	0.2631	0.0356	0.0745	0.0079	0.0018
50	0.1	0.2077	0.0817	0.228	0.0254	0.005
	0.2	0.2374	0.0599	0.1431	0.0165	0.0039
	0.3	0.2358	0.0415	0.0718	0.0087	0.0019

Table S31: Multivariate Exponential distribution FPR table for $n = 100$

p	$\boldsymbol{\alpha}$	**FAST MCD**	**OGK**	**RMDS**	$\boldsymbol{S_n}$	Sh-S_n
δ=10,λ=0.01						
5	0.1	0.0724	0.0336	0.1298	0.0295	0.003
	0.2	0.035	0.0295	0.0856	0.0427	0.0011
	0.3	0.0854	0.0295	0.0531	0.1219	0.0002
10	0.1	0.1183	0.0397	0.1742	0	0.0028
	0.2	0.064	0.0391	0.1247	0.0002	0.0026
	0.3	0.6401	0.0465	0.0849	0.0082	0.0007
15	0.1	0.1471	0.0466	0.1988	0.0175	0.0032
	0.2	0.5041	0.0488	0.1485	0.0249	0.0024
	0.3	0.5991	0.0549	0.1262	0.0323	0.0015
20	0.1	0.1809	0.054	0.2243	0.0196	0.0036
	0.2	0.4923	0.0518	0.1732	0.0298	0.0031
	0.3	0.5714	0.0623	0.1531	0.0327	0.001
30	0.1	0.2134	0.0636	0.2576	0.0246	0.004
	0.2	0.4369	0.0578	0.2317	0.0314	0.0029
	0.3	0.5111	0.0701	0.1589	0.0399	0.0015
50	0.1	0.2775	0.0846	0.3162	0.0264	0.0051
	0.2	0.3325	0.0678	0.2718	0.0324	0.0038
	0.3	0.3971	0.0732	0.1349	0.042	0.0016
δ=10,λ=0.1						
5	0.1	0.0696	0.033	0.0007	0.027	0.0026
	0.2	0.036	0.0259	0.0003	0.0336	0.0012
	0.3	0.0091	0.0234	0.0006	0.0601	0.0005
10	0.1	0.1151	0.0388	0.0008	0	0.0034
	0.2	0.0671	0.0371	0.001	0.0001	0.0022
	0.3	0.6018	0.0339	0.0074	0.0001	0.0007
15	0.1	0.1505	0.0435	0.0024	0.016	0.0034
	0.2	0.1337	0.0448	0.0043	0.0185	0.0022
	0.3	0.5861	0.0407	0.046	0.02	0.0009
20	0.1	0.1822	0.0541	0.0068	0.0182	0.0041
	0.2	0.469	0.0482	0.024	0.0221	0.0027
	0.3	0.5666	0.0472	0.0679	0.0197	0.0014
30	0.1	0.191	0.0629	0.0208	0.024	0.005
	0.2	0.433	0.055	0.1275	0.0244	0.0032
	0.3	0.4993	0.0527	0.0579	0.0227	0.002
50	0.1	0.2766	0.0793	0.2206	0.0264	0.0054
	0.2	0.3124	0.0645	0.1298	0.0254	0.0031
	0.3	0.3569	0.0529	0.0901	0.0242	0.0021
δ=10,λ=1						
5	0.1	0.074	0.0282	0.0005	0.0246	0.0031
	0.2	0.0362	0.0223	0.0004	0.0142	0.0016
	0.3	0.011	0.0132	0.0002	0.0111	0.0002
10	0.1	0.1168	0.039	0.0012	0	0.004
	0.2	0.0724	0.0333	0.0006	0	0.0025
	0.3	0.023	0.0223	0.0109	0	0.0007
15	0.1	0.1564	0.0467	0.003	0.0162	0.0037
	0.2	0.0829	0.035	0.0029	0.012	0.0021
	0.3	0.0771	0.0255	0.0409	0.0077	0.0007
20	0.1	0.1864	0.0535	0.0051	0.0174	0.0044
	0.2	0.1063	0.0427	0.019	0.0153	0.0031
	0.3	0.2134	0.0307	0.0658	0.0112	0.0011
30	0.1	0.1948	0.0651	0.0241	0.0203	0.005
	0.2	0.2322	0.0466	0.1222	0.0142	0.0032
	0.3	0.2671	0.0337	0.0887	0.0103	0.0018
50	0.1	0.2072	0.0753	0.2246	0.0251	0.0054
	0.2	0.2363	0.0611	0.1342	0.0192	0.0034
	0.3	0.2369	0.0377	0.0607	0.0113	0.0023

Table S32: Multivariate Exponential distribution FPR table for $n = 500$

p	$\boldsymbol{\alpha}$	**FAST MCD**	**OGK**	**RMDS**	$\boldsymbol{S_n}$	Sh-S_n
δ=5,λ=0.01						
	0.1	0.0041	0.0228	0.0004	0.0401	0.0036
5	0.2	0.0018	0.0285	0	0.1078	0.0007
	0.3	0.5954	0.0366	0	0.2678	0
	0.1	0.0058	0.0274	0.0006	0	0.0059
10	0.2	0.0832	0.0417	0	0.009	0.0028
	0.3	0.6888	0.043	0	0.0289	0
	0.1	0.0089	0.0308	0.0005	0.0098	0.0062
15	0.2	0.4679	0.0467	0	0.0176	0.0009
	0.3	0.7598	0.0444	0	0.0091	0
	0.1	0.0189	0.0332	0	0.0089	0.0089
20	0.2	0.5743	0.0508	0	0.014	0.001
	0.3	0.8156	0.0467	0	0.0012	0
	0.1	0.1139	0.0368	0	0.0098	0.0079
30	0.2	0.581	0.0565	0	0.0195	0.0005
	0.3	0.8899	0.0521	0	0.0091	0
	0.1	0.3878	0.0417	0.0004	0.0161	0.0089
50	0.2	0.5798	0.0629	0	0.0101	0.0008
	0.3	0.9567	0.057	0	0.0098	0
δ=5,λ=0.1						
	0.1	0.0071	0.0221	0.129	0.0402	0.0056
5	0.2	0.0019	0.0226	0.0518	0.0698	0.0009
	0.3	0.0008	0.0242	0.0145	0.1765	0
	0.1	0.0087	0.0262	0.2618	0	0.0089
10	0.2	0.0037	0.032	0.0998	0	0.0005
	0.3	0.5671	0.0249	0.0217	0.0078	0
	0.1	0.0089	0.0292	0.4156	0.0049	0.0072
15	0.2	0.3897	0.0356	0.1689	0.0076	0.0008
	0.3	0.6896	0.0253	0.0323	0.0089	0
	0.1	0.9756	0.0307	0.5389	0.0072	0.0069
20	0.2	0.4892	0.0379	0.2399	0.0089	0.0005
	0.3	0.6678	0.0273	0.0487	0.0089	0
	0.1	0.0286	0.0343	0.7398	0.0071	0.0082
30	0.2	0.5891	0.0424	0.398	0.0091	0.0007
	0.3	0.7906	0.03	0.1238	0.0056	0
	0.1	0.3465	0.0404	0.9389	0.0099	0.0074
50	0.2	0.5781	0.0482	0.6771	0.0079	0.0009
	0.3	0.7912	0.0359	0.4589	0.0032	0
δ=5,λ=1						
	0.1	0.0096	0.0208	0.1361	0.0289	0.0079
5	0.2	0.0015	0.0145	0.061	0.0218	0.0009
	0.3	0.0008	0.0123	0.0189	0.023	0
	0.1	0.0058	0.0242	0.2682	0.0004	0.0039
10	0.2	0.0041	0.0194	0.099	0.0001	0.0021
	0.3	0.0008	0.0146	0.0145	0	0
	0.1	0.0088	0.0257	0.4109	0.0059	0.0081
15	0.2	0.0035	0.0226	0.1816	0.0043	0.0008
	0.3	0.002	0.0154	0.0498	0.0027	0
	0.1	0.0101	0.0281	0.5591	0.0034	0.0064
20	0.2	0.0079	0.0244	0.2309	0.0054	0.0009
	0.3	0.0067	0.0162	0.056	0.0071	0
	0.1	0.0189	0.0305	0.789	0.0084	0.0067
30	0.2	0.0126	0.0263	0.3904	0.0066	0.0009
	0.3	0.0254	0.0178	0.1291	0.0037	0
	0.1	0.0535	0.035	0.9307	0.0089	0.008
50	0.2	0.0689	0.0316	0.6751	0.007	0.0009
	0.3	0.0721	0.0198	0.4458	0.0027	0

Table S33: Multivariate Exponential distribution FPR table for $n = 500$

p	$\boldsymbol{\alpha}$	**FAST MCD**	**OGK**	**RMDS**	$\boldsymbol{S_n}$	Sh-S_n
			δ=10,λ=0.01			
	0.1	0.0078	0.0229	0.1298	0.0278	0.0039
5	0.2	0.0009	0.0287	0.0549	0.0497	0.0003
	0.3	0.0006	0.0336	0.0119	0.1267	0.0001
	0.1	0.0078	0.0275	0.2678	0	0.0089
10	0.2	0.0076	0.0363	0.0999	0.0006	0.0008
	0.3	0.7017	0.0523	0.0201	0.0078	0.0001
	0.1	0.0089	0.029	0.478	0.0089	0.0058
15	0.2	0.4645	0.0415	0.1689	0.0178	0.0008
	0.3	0.7478	0.062	0.0389	0.0178	0
	0.1	0.0101	0.0313	0.5378	0.0078	0.006
20	0.2	0.5478	0.0467	0.241	0.0167	0.0004
	0.3	0.819	0.0689	0.0518	0.0217	0
	0.1	0.0278	0.0349	0.7589	0.0108	0.0074
30	0.2	0.5978	0.0475	0.3934	0.0189	0.0005
	0.3	0.8989	0.0778	0.1189	0.0154	0
	0.1	0.3917	0.0478	0.9389	0.0239	0.0145
50	0.2	0.5856	0.0598	0.6689	0.0178	0.0001
	0.3	0.9589	0.0945	0.4481	0.0278	0
			δ=10,λ=0.1			
	0.1	0.0076	0.022	0.1289	0.0598	0.0038
5	0.2	0.0007	0.0235	0.0611	0.0378	0.0009
	0.3	0.0029	0.0221	0.0281	0.0318	0.0001
	0.1	0.0078	0.025	0.2798	0.0012	0.0083
10	0.2	0.0078	0.0306	0.198	0.0011	0.0018
	0.3	0.5856	0.0336	0.0276	0.0008	0.0001
	0.1	0.0084	0.0275	0.4189	0.0081	0.0056
15	0.2	0.0036	0.0339	0.1789	0.0129	0.0009
	0.3	0.6812	0.0425	0.0385	0.0198	0.0003
	0.1	0.0106	0.0296	0.5981	0.0089	0.007
20	0.2	0.4278	0.0361	0.2409	0.0189	0.0009
	0.3	0.6989	0.0449	0.0527	0.0178	0.0007
	0.1	0.0278	0.0336	0.7567	0.0091	0.0089
30	0.2	0.5891	0.0393	0.3981	0.0118	0.0008
	0.3	0.7879	0.0479	0.1265	0.0189	0.0001
	0.1	0.2589	0.0372	0.9679	0.0099	0.0088
50	0.2	0.5799	0.0423	0.6891	0.0106	0.0008
	0.3	0.8019	0.0547	0.4598	0.0081	0.0001
			δ=10,λ=1			
	0.1	0.0078	0.0199	0.1389	0.0189	0.0076
5	0.2	0.0067	0.0153	0.067	0.0091	0.0007
	0.3	0.0004	0.0099	0.0109	0.0079	0.0001
	0.1	0.0078	0.023	0.2717	0.0001	0.0067
10	0.2	0.0078	0.0209	0.1111	0.0001	0.0008
	0.3	0.0008	0.0172	0.0187	0.0001	0
	0.1	0.0078	0.0264	0.4089	0.0087	0.0061
15	0.2	0.0078	0.0239	0.1669	0.0049	0.0009
	0.3	0.0019	0.0197	0.0356	0.0078	0.0001
	0.1	0.05	0.0282	0.567	0.0089	0.0087
20	0.2	0.0087	0.0255	0.2378	0.0089	0.0006
	0.3	0.0067	0.0216	0.0589	0.0089	0.0001
	0.1	0.0278	0.0308	0.7679	0.0089	0.005
30	0.2	0.0189	0.0278	0.3817	0.0062	0.0006
	0.3	0.0267	0.0245	0.1127	0.0087	0
	0.1	0.0548	0.0355	0.9272	0.0093	0.0089
50	0.2	0.0692	0.0311	0.6678	0.0078	0.0005
	0.3	0.0715	0.0283	0.4567	0.0078	0

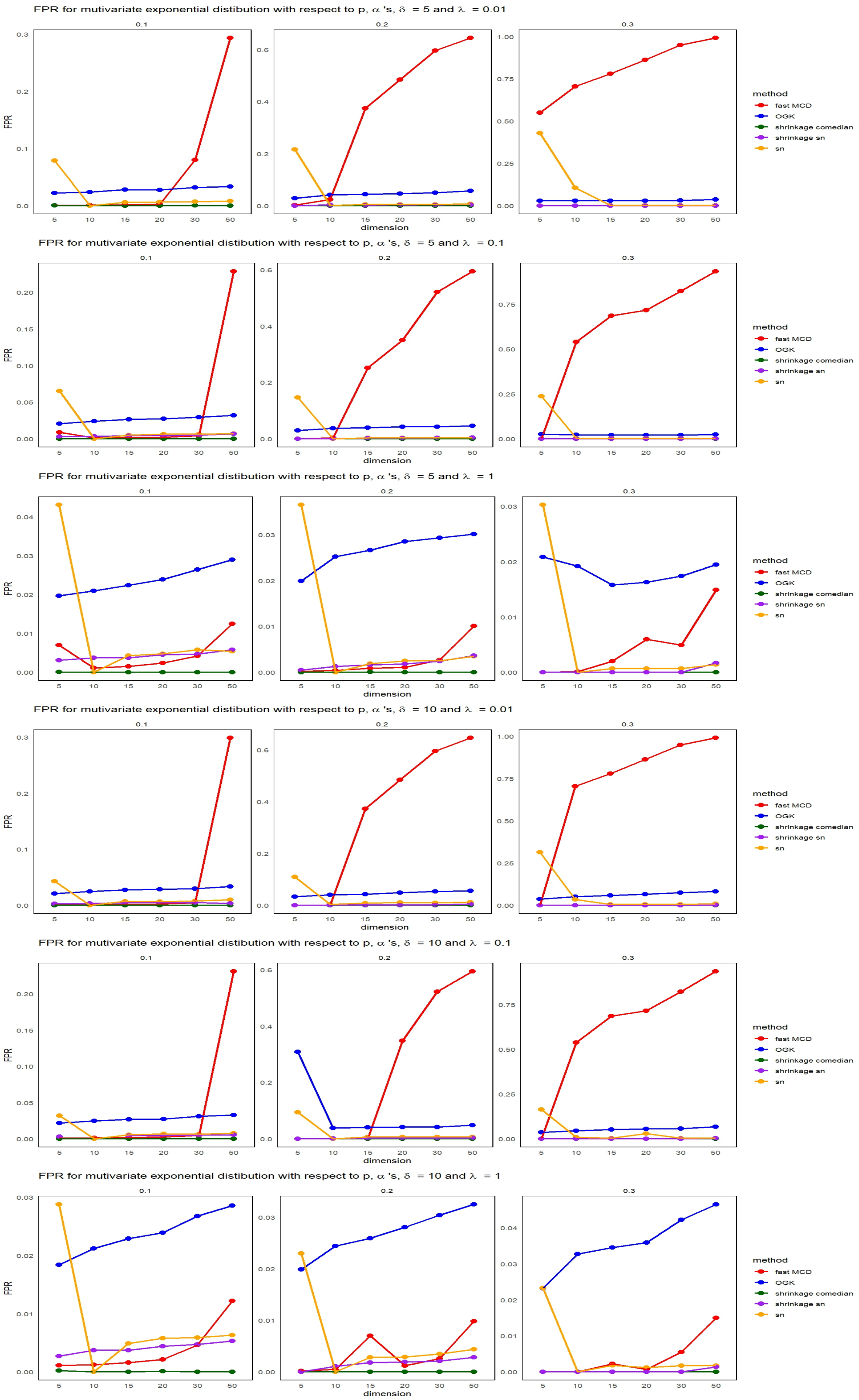


Figure A.3: (a) Figure of FPR for multivariate exponential distribution for $n = 1000$

Table S34: Affine Equivariance Property exhibiting table of RMD Sh-S_n across different sample size

		$n = 100$						$n = 500$						$n = 1000$					
	δ=5	**λ=0.01**		**λ=0.1**		**λ=1**		**λ=0.01**		**λ=0.1**		**λ=1**		**λ=0.01**		**λ=0.1**		**λ=1**	
p	α	**FPR**	**TPR**	**FPR**	**TPR**	**FPR**	**TPR**	**FPR**	**TPR**	**FPR**	**TPR**	**FPR**	**TPR**	**FPR**	**TPR**	**FPR**	**TPR**	**FPR**	**TPR**
	0.1	0.0844	1	0.0611	1	0.0842	1	0.0714	1	0.0604	1	0.0755	1	0.0421	1	0.0508	1	0.0574	1
5	0.2	0.0693	1	0.0666	0.954	0.0625	0.963	0.0675	0.9673	0.0615	1	0.0626	1	0.0489	0.9875	0.0485	0.9566	0.0599	0.98
	0.3	0.0393	0.6527	0.0224	0.735	0.0326	0.7817	0.0666	1	0.0657	1	0.0242	0.8039	0.0189	0.95	0.0156	0.9657	0.0094	0.968
	0.1	0.1088	1	0.1199	1	0.126	1	0.0931	1	0.1161	1	0.0896	1	0.0978	1	0.0633	1	0.0591	1
10	0.2	0.1255	1	0.1238	0.9899	0.1226	0.998	0.1239	1	0.1192	1	0.1136	1	0.0761	0.9899	0.0859	1	0.0824	1
	0.3	0.0942	0.8923	0.0634	0.9017	0.058	0.9033	0.0709	0.8606	0.1092	1	0.0631	0.9352	0.0566	0.7393	0.0237	0.77	0.0291	0.85
	0.1	0.1262	1	0.1146	1	0.1241	1	0.1136	1	0.1238	1	0.1255	1	0.0892	1	0.0969	1	0.0752	1
15	0.2	0.1301	1	0.1351	1	0.1226	1	0.1136	1	0.1298	1	0.1322	1	0.1209	1	0.0897	1	0.1126	1
	0.3	0.092	0.9899	0.0858	0.9493	0.0543	0.9277	0.0969	0.9245	0.0463	0.9	0.0587	0.9084	0.1261	0.9016	0.0232	0.8598	0.0564	0.8695
	0.1	0.1426	1	0.1504	1	0.1541	1	0.1588	1	0.1564	1	0.1541	1	0.0853	1	0.098	1	0.1327	1
20	0.2	0.2035	1	0.194	1	0.1973	1	0.1797	1	0.1897	1	0.1887	1	0.1231	1	0.1274	1	0.1522	1
	0.3	0.1207	0.9833	0.1384	0.986	0.1076	0.9817	0.1173	0.9838	0.113	0.9739	0.0948	0.9647	0.0739	0.9208	0.0761	0.9291	0.0414	0.8867
	0.1	0.1662	1	0.1661	1	0.1538	1	0.1548	1	0.1821	1	0.1798	1	0.1113	1	0.1108	1	0.1198	1
30	0.2	0.2045	1	0.1918	1	0.1921	1	0.1887	1	0.1849	1	0.1997	1	0.1585	1	0.1485	1	0.1289	1
	0.3	0.1223	1	0.1447	1	0.1329	0.99	0.1054	0.9887	0.1376	0.9961	0.0997	0.9921	0.0464	0.9026	0.1074	0.9811	0.0424	0.9075
	δ=10	**λ=0.01**		**λ=0.1**		**λ=1**		**λ=0.01**		**λ=0.1**		**λ=1**		**λ=0.01**		**λ=0.1**		**λ=1**	
	α	**FPR**	**TPR**	**FPR**	**TPR**	**FPR**	**TPR**	**FPR**	**TPR**	**FPR**	**TPR**	**FPR**	**TPR**	**FPR**	**TPR**	**FPR**	**TPR**	**FPR**	**TPR**
	0.1	0.0706	1	0.0798	1	0.0794	1	0.0541	1	0.1127	1	0.0796	1	0.0476	1	0.0391	1	0.0365	1
5	0.2	0.0631	1	0.07	0.99	0.0768	0.985	0.059	1	0.0585	0.98	0.0616	0.9768	0.0498	1	0.0497	1	0.0329	0.95
	0.3	0.0772	0.9463	0.0524	0.974	0.0701	0.945	0.0559	0.9486	0.0656	0.9546	0.053	0.9395	0.0493	1	0.0538	0.9661	0.0327	0.9457
	0.1	0.1261	1	0.1253	1	0.106	1	0.1204	1	0.1059	1	0.1082	1	0.0969	1	0.0956	0.9876	0.1074	1
10	0.2	0.1545	1	0.1316	1	0.1379	1	0.1229	1	0.1127	1	0.1279	1	0.0851	1	0.0953	1	0.0793	1
	0.3	0.1176	0.95	0.1048	0.965	0.102	0.9707	0.111	0.9783	0.0856	0.97	0.0948	0.9872	0.0798	0.9815	0.0615	0.9313	0.0772	0.9789
	0.1	0.1186	1	0.1167	1	0.1116	1	0.1268	1	0.1296	1	0.1452	1	0.1042	1	0.1089	1	0.0944	1
15	0.2	0.1328	1	0.1606	1	0.1531	1	0.1315	1	0.1341	1	0.1596	1	0.0837	1	0.0787	1	0.0757	1
	0.3	0.1298	1	0.1084	0.97	0.1016	0.983	0.1299	0.9783	0.1042	1	0.1044	0.9877	0.069	0.9639	0.0729	0.9536	0.075	0.9723
	0.1	0.1571	1	0.137	1	0.1648	1	0.1616	1	0.1466	1	0.1509	1	0.1572	1	0.1821	1	0.1549	1
20	0.2	0.2049	1	0.2038	1	0.1939	1	0.2001	1	0.1854	1	0.1809	1	0.0852	1	0.1384	1	0.0984	1
	0.3	0.1633	1	0.1741	0.9999	0.1511	0.9847	0.167	1	0.1473	1	0.1349	0.9896	0.1328	0.9885	0.0997	0.9662	0.1117	0.9854
	0.1	0.1487	1	0.1518	1	0.2146	1	0.1748	1	0.1849	1	0.1782	1	0.1253	1	0.1269	1	0.1257	1
30	0.2	0.2049	1	0.1955	1	0.157	1	0.2072	1	0.2014	1	0.1984	1	0.1742	1	0.1509	1	0.1468	1
	0.3	0.1597	1	0.1743	1	0.1811	1	0.13	1	0.1651	1	0.1394	1	0.0801	0.97	0.1461	0.993	0.0811	0.9792

Table S35: Breakdown Property table of RMD Sh-S_n for different sample sizes

		n = **100**				n = **500**				n = **1000**			
		Symmetric		**Asymmetric**		**Symmetric**		**Asymmetric**		**Symmetric**		**Asymmetric**	
p	α	**FPR**	**TPR**	**FPR**	**TPR**	**FPR**	**TPR**	**FPR**	**TPR**	**FPR**	**TPR**	**FPR**	**TPR**
10	0.1	0.0101	1	0.0112	1	0.0094	1	0.0104	1	0.0101	1	0.0101	1
	0.2	0.0056	1	0.0044	1	0.005	1	0.0043	1	0.0058	1	0.0038	1
	0.3	0.0014	1	0.0004	1	0.0021	1	0.0004	1	0.0018	1	0.0004	1
	0.4	0.0003	1	0	1	0	1	0	1	0.0002	1	0	1
	0.45	0.0002	1	0	1	0	1	0	1	0	1	0	1
30	0.1	0.0054	1	0.0065	1	0.0039	1	0.0048	1	0.0041	1	0.0044	1
	0.2	0.0003	1	0.0041	1	0.0018	1	0.0019	1	0.0027	1	0.0018	1
	0.3	0.0001	1	0.0001	1	0.0009	1	0	1	0.001	1	0.0005	1
	0.4	0.0002	1	0	1	0	1	0	1	0.0003	1	0	1
	0.45	0	1	0	1	0	1	0	1	0	1	0	1
50	0.1	0.0044	1	0.0038	1	0.0024	1	0.0038	1	0.0025	1	0.0021	1
	0.2	0.0016	1	0.0091	1	0.0012	1	0.0018	1	0.0011	1	0.0008	1
	0.3	0.0003	1	0.1127	1	0.0005	1	0.0003	1	0.0003	1	0	1
	0.4	0.0002	1	0.0059	1	0	1	0	1	0	1	0	1
	0.45	0.0002	1	0.0209	1	0	1	0	1	0	1	0	1
80	0.1	0.0008	1	0.0034	1	0.0012	1	0.0001	1	0.0011	1	0.0008	1
	0.2	0.0015	1	0.11	1	0.0006	1	0.0005	1	0	1	0.0006	1
	0.3	0.0005	1	0.1278	1	0.0005	1	0	1	0	1	0	1
	0.4	0	1	0.0752	1	0	1	0	1	0	1	0	1
	0.45	0	1	0.0387	1	0	1	0	1	0	1	0	1
100	0.1	0.0024	1	0.003	1	0.001	1	0.0008	1	0	1	0.0008	1
	0.2	0.0012	1	0.1471	1	0.0004	1	0.0003	1	0.0008	1	0.0002	1
	0.3	0.0005	1	0.1383	1	0.0001	1	0	1	0	1	0	1
	0.4	0	1	0.0705	1	0	1	0	1	0	1	0	1
	0.45	0	1	0.0427	1	0	1	0	1	0	1	0	1

Appendix B

B.1 Additional tables from chapter 4

Table B.1: Robustness table-Contaminants from N$(\mathbf{2}\sqrt{\boldsymbol{\chi}^{\mathbf{2}}_{\boldsymbol{p+q,0.99}}}, 0.1 \times \mathbf{I}_q)$ and (p, q) = (4,8)

	Slope MSE	Intercept MSE	diagonal element mse	non diagonal elements mse	slope bias	intercept bias	diagonal element bias	non diagonal element bias
		$n = 50$	$\delta = 10\%$	$p = 4$	$q = 8$	diag = 0.1		
S_n	0.026368	0.025535	0.898845	0.021827	-0.00054	-0.0024	0.924452	-0.00526
MCD	0.051203	0.036295	0.687359	0.029132	-0.00179	-1.34e-3	0.793941	-0.00536
Comedian	0.026599	0.02533	0.887046	0.02133	-0.00068	0.003809	0.917734	-0.00195
OGK	0.032623	0.029198	0.783143	0.023087	0.001405	0.005655	0.858097	-0.00088
RSh-S_n	0.02599	0.024524	0.849369	0.020729	-0.00123	3.60e-3	0.898265	-0.00183
		$n = 100$	$\delta = 10\%$	$p = 4$	$q = 8$	diag = 0.1		
S_n	0.01207	0.011584	0.935365	0.011002	0.00052	0.001621	0.955791	0.001446
MCD	0.017647	0.013201	0.750804	0.013175	0.000476	-0.00092	0.850689	0.002867
Comedian	0.012224	0.011975	0.928054	0.010863	0.000587	0.000461	0.951575	0.002234
OGK	0.013814	0.011857	0.845729	0.011493	-0.00083	0.00274	0.906208	0.004591
RSh-S_n	0.011325	0.011505	0.884337	0.01069	-0.00068	0.001695	0.928422	0.001768
		$n = 200$	$\delta = 10\%$	$p = 4$	$q = 8$	diag = 0.1		
S_n	0.005859	0.005726	0.943262	0.005523	6.59e-4	-0.00039	0.965365	0.004226
MCD	0.006482	0.005635	0.870071	0.005925	0.000157	3.95e-4	0.926094	0.004519
Comedian	0.005881	0.005699	0.944197	0.005567	-1.60e-4	-9.90e-5	0.965948	0.004217
OGK	0.006364	0.005658	0.892352	0.005751	-0.00016	0.000539	0.938439	0.004844
RSh-S_n	0.005851	0.005604	0.91682	0.00553	0.000377	0.000142	0.951738	0.003961
		$n = 500$	$\delta = 10\%$	$p = 4$	$q = 8$	diag = 0.1		
S_n	0.002333	0.002207	0.957668	0.002265	2.36e-4	0.000351	0.976269	0.00084
MCD	0.002394	0.002244	0.924019	0.002312	-4.50e-4	0.000588	0.958785	0.00059
Comedian	0.002303	0.002211	0.954446	0.002251	-0.0004	0.000591	0.974639	0.001185
OGK	0.002394	0.002237	0.922555	0.002337	6.94e-4	0.001017	0.958018	0.000802
RSh-S_n	0.00263	0.00254	0.94306	0.00254	5.1323e-06	0.00098	0.96847	0.00010

Table B.2: Robustness table-Contaminants from N($\mathbf{2}\sqrt{\boldsymbol{\chi}^{\mathbf{2}}_{\boldsymbol{p+q},\mathbf{0.99}}}$, $0.1\times\mathbf{I}_q$) and (p,q) = (8,4)

	Slope MSE	Intercept MSE	diagonal element mse	non diagonal elements mse	slope bias	intercept bias	diagonal element bias	non diagonal element bias
		$n = 50$	$\delta = 10\%$	$p = 8$	$q = 4$	diag = 0.1		
S_n	0.030071	0.029982	0.710073	0.020074	-0.00024	0.001393	0.818126	-0.00602
MCD	0.071754	0.040567	0.452425	0.025876	-0.00251	2.57e-3	0.632953	-0.00422
Comedian	0.03016	0.0283	0.702836	0.019552	0.000189	0.001694	0.814418	-0.00497
OGK	0.036603	0.029532	0.601802	0.019494	0.001023	0.007873	0.746853	0.002305
RSh-S_n	0.021185	0.02856	0.6768	0.019388	0.000149	5.77e-4	0.797973	-0.00343
		$n = 100$	$\delta = 10\%$	$p = 8$	$q = 4$	diag = 0.1		
S_n	0.01468	0.011625	0.822838	0.010568	0.000283	-0.00167	0.895149	0.005459
MCD	0.018414	0.012494	0.632756	0.012816	-6.80e-4	0.00015	0.778974	0.005181
Comedian	0.013838	0.011703	0.814563	0.010544	0.000571	0.001131	0.890538	0.005298
OGK	0.014344	0.012224	0.749691	0.011096	5.24e-6	0.0003	0.852167	0.009678
RSh-S_n	0.013151	0.011495	0.786627	0.010442	0.000276	0.001099	0.875251	0.007938
		$n = 200$	$\delta = 10\%$	$p = 8$	$q = 4$	diag = 0.1		
S_n	0.006087	0.005765	0.887469	0.005593	-9.50e-5	0.000937	0.936072	0.003351
MCD	0.006688	0.005708	0.854231	0.005953	-1.50e-4	7.88e-4	0.889898	0.002215
Comedian	0.006099	0.005683	0.87817	0.00559	1.93e-4	7.88e-4	0.930824	0.005072
OGK	0.006438	0.005796	0.833996	0.00575	-0.00015	-0.00057	0.90655	0.00307
RSh-S_n	0.006072	0.005668	0.859239	0.005493	-0.00018	-0.00071	0.920615	0.001929
		$n = 500$	$\delta = 10\%$	$p = 8$	$q = 4$	diag = 0.1		
S_n	0.002327	0.002167	0.919339	0.00227	4.89e-4	-4.70e-4	0.956496	0.000819
MCD	0.006688	0.005708	0.804231	0.005953	-1.50e-4	0.000788	0.889898	0.002215
Comedian	0.002319	0.002195	0.919471	0.002271	-0.00012	0.000267	0.956583	0.000463
OGK	0.002443	0.002253	0.885856	0.002292	4.43e-4	3.71e-5	0.938705	0.002677
RSh-S_n	0.00269	0.00242	0.90464	0.00249	1.1846e-05	0.00105	0.94839	0.00027

Table B.3: Robustness table-Contaminants from N($\mathbf{2}\sqrt{\boldsymbol{\chi}^{\mathbf{2}}_{\boldsymbol{p+q,0.99}}}, 0.1 \times \mathbf{I}_q$)

	Slope MSE	Intercept MSE	diagonal element mse	non diagonal elements mse	slope bias	intercept bias	diagonal element bias	non diagonal element bias
		$n = 50$	$\delta = 10\%$	$p = 10$	$q = 10$	diag = 0.1		
S_n	0.035379	0.034299	0.624847	0.019899	0.000192	0.001867	0.764586	0.027986
MCD	0.052203	0.044448	0.537251	0.023585	-0.00015	2.92e-4	0.701854	0.035432
Comedian	0.03154	0.029931	0.679281	0.019648	0.000477	0.000523	0.800382	0.02496
OGK	0.04065	0.037191	0.582545	0.02048	0.000208	0.006032	0.735208	0.032768
RSh-S_n	0.03277	0.030403	0.639593	0.018907	0.000251	1.97e-3	0.775733	0.023259
		$n = 100$	$\delta = 10\%$	$p = 10$	$q = 10$	diag = 0.1		
S_n	0.014004	0.013292	0.762375	0.010382	0.000578	0.000734	0.860593	0.011466
MCD	0.02302	0.017691	0.608272	0.013155	6.72e-5	0.001166	0.763183	0.01508
Comedian	0.012894	0.012416	0.806761	0.010174	-0.00035	-0.00078	0.886806	0.01022
OGK	0.01515	0.013676	0.721858	0.010639	-3.70e-4	0.001832	0.836498	0.012231
RSh-S_n	0.013372	0.012622	0.773067	0.010031	0.000441	0.001128	0.867663	0.010838
		$n = 200$	$\delta = 10\%$	$p = 10$	$q = 10$	diag = 0.1		
S_n	0.00634	0.005954	0.854939	0.005492	1.40e-5	0.000272	0.918432	0.00517
MCD	0.007377	0.006179	0.775626	0.006039	8.47e-5	-3.50e-4	0.873666	0.004764
Comedian	0.006077	0.005807	0.892463	0.005505	8.34e-4	-9.50e-4	0.938882	0.004963
OGK	0.006604	0.006088	0.830516	0.005617	-0.00017	0.002482	0.905209	0.005521
RSh-S_n	0.00613	0.005825	0.858965	0.005326	-0.00024	0.001276	0.920977	0.005581
		$n = 500$	$\delta = 10\%$	$p = 10$	$q = 10$	diag = 0.1		
S_n	0.002376	0.002208	0.919783	0.002225	-1.10e-4	-5.20e-5	0.9566	0.001779
MCD	0.002418	0.002309	0.900211	0.002273	1.25e-4	0.000241	0.94635	0.001675
Comedian	0.002348	0.002271	0.936253	0.002213	-0.00012	-0.00116	0.965302	0.002007
OGK	0.002436	0.002339	0.903601	0.002292	-1.50e-4	3.31e-4	0.948139	0.0024
RSh-S_n	0.03154	0.029931	0.679281	0.019648	0.000477	0.000523	0.800382	0.02496

Table B.4: Robustness table-Contaminants from N($\mathbf{2}\sqrt{\boldsymbol{\chi}^{\mathbf{2}}_{\boldsymbol{p+q},\mathbf{0.99}}}$, $\mathbf{I}_q$) and $(p, q) =$ (4,4)

	Slope MSE	Intercept MSE	diagonal element mse	non diagonal elements mse	slope bias	intercept bias	diagonal element bias	non diagonal element bias
		$n = 50$	$\delta = 10\%$	$p = 4$	$q = 4$			
S_n	0.027104	0.026141	0.871365	0.022099	-0.00017	-0.0017	0.909667	-0.00548
MCD	0.054961	0.030504	0.600277	0.028164	-2.90e-5	-2.78e-3	0.736767	-0.001
Comedian	0.032238	0.026643	0.761882	0.022298	0.002052	0.003696	0.845044	-0.00103
OGK	0.032483	0.025911	0.746421	0.022184	0.000958	0.004303	0.836388	0.001282
RSh-S_n	0.026918	0.024661	0.826234	0.021195	0.0004	1.17e-3	0.883227	-0.00357
		$n = 100$	$\delta = 10\%$	$p = 4$	$q = 4$			
S_n	0.01222	0.011744	0.909175	0.011397	0.000622	0.000536	0.941562	0.006297
MCD	0.015602	0.01156	0.766	0.01277	-2.80e-5	-0.00137	0.859509	0.005811
Comedian	0.013078	0.011632	0.821849	0.011563	0.000731	0.001822	0.893043	0.008888
OGK	0.013281	0.011338	0.825212	0.011239	1.36e-3	-0.00072	0.895182	0.007572
RSh-S_n	0.012423	0.011437	0.864166	0.010954	0.000505	0.00206	0.917553	0.005991
		$n = 200$	$\delta = 10\%$	$p = 4$	$q = 4$			
S_n	0.006075	0.00561	0.927595	0.005794	-2.20e-4	-0.00045	0.956908	0.002803
MCD	0.006341	0.005631	0.86313	0.005914	1.24e-4	1.23e-3	0.922112	0.003357
Comedian	0.006279	0.005663	0.87052	0.005857	-3.30e-4	8.01e-4	0.926773	0.004161
OGK	0.006323	0.00565	0.866304	0.006013	-1.70e-5	0.000297	0.924135	0.005033
RSh-S_n	0.006041	0.005542	0.893558	0.00544	-0.00019	0.00033	0.939649	0.004496
		$n = 500$	$\delta = 10\%$	$p = 4$	$q = 4$			
S_n	0.002324	0.002114	0.93284	0.002287	7.79e-5	-3.00e-5	0.963423	0.002273
MCD	0.002454	0.002223	0.904046	0.002375	-1.80e-5	-0.00058	0.948241	0.002253
Comedian	0.002425	0.002141	0.892591	0.002396	-0.00048	0.001091	0.942233	0.002711
OGK	0.002397	0.002234	0.922206	0.002297	-4.10e-5	6.97e-4	0.957776	0.00134
RSh-S_n	0.002367	0.00227	0.914538	0.002251	0.000522	0.000228	0.953882	0.00139

Table B.5: Robustness table-Contaminants from N($\mathbf{2}\sqrt{\boldsymbol{\chi}^{\mathbf{2}}_{\boldsymbol{p+q},\mathbf{0.99}}}, \mathbf{I}_q$) and $(p, q) =$ (4,8)

	Slope MSE	Intercept MSE	diagonal element mse	non diagonal elements mse	slope bias	intercept bias	diagonal element bias	non diagonal element bias
		$n = 50$	$\delta = 10\%$	$p = 4$	$q = 8$			
S_n	0.026036	0.024352	0.898834	0.021601	0.000235	6.76e-5	0.925397	-0.00177
MCD	0.052629	0.036653	0.67425	0.028786	3.13e-3	-4.60e-4	0.784536	-0.00185
Comedian	0.026419	0.02595	0.895646	0.021803	-7.60e-5	-0.00015	0.922931	-0.00296
OGK	0.032762	0.027617	0.781278	0.022477	0.001767	0.003449	0.855827	-0.00057
RSh-S_n	0.026652	0.02444	0.852299	0.020924	-0.00079	4.28e-3	0.898662	-0.00315
		$n = 100$	$\delta = 10\%$	$p = 4$	$q = 8$			
S_n	0.01207	0.011492	0.93025	0.010913	0.000931	-0.00204	0.953118	0.00125
MCD	0.017246	0.012694	0.755446	0.013573	2.71e-4	-0.00177	0.85355	0.00309
Comedian	0.013746	0.012192	0.837849	0.011448	0.000796	0.002605	0.902437	0.003119
OGK	0.013819	0.012194	0.840424	0.011384	5.36e-4	0.002834	0.90312	0.004237
RSh-S_n	0.012228	0.011432	0.885593	0.010725	4.20e-6	-0.00143	0.929856	0.001533
		$n = 200$	$\delta = 10\%$	$p = 4$	$q = 8$			
S_n	0.005922	0.005727	0.951432	0.00569	2.68e-5	0.000311	0.969582	0.004796
MCD	0.006504	0.005943	0.873553	0.005972	-2.10e-4	2.88e-4	0.927869	0.004187
Comedian	0.006353	0.005859	0.886311	0.005809	-3.90e-4	1.94e-3	0.935359	0.005867
OGK	0.006369	0.005742	0.835858	0.005715	-6.10e-4	0.001927	0.907761	0.004887
RSh-S_n	0.00598	0.005617	0.919595	0.005549	-0.00052	0.001082	0.953175	0.004802
		$n = 500$	$\delta = 10\%$	$p = 4$	$q = 8$			
S_n	0.002305	0.002237	0.955434	0.00224	-4.50e-4	6.80e-4	0.975091	0.000792
MCD	0.002346	0.002193	0.92329	0.002277	-9.80e-5	0.000793	0.958386	0.000411
Comedian	0.002286	0.002301	0.955494	0.002256	3.59e-5	-0.00134	0.975122	0.00014
OGK	0.002397	0.002234	0.922206	0.002297	-4.10e-5	6.97e-4	0.957776	0.00134
RSh-S_n	0.002327	0.002199	0.938934	0.002256	0.000349	0.001011	0.96662	0.000289

Table B.6: Robustness table-Contaminants from N($\mathbf{2}\sqrt{\boldsymbol{\chi}^{\mathbf{2}}_{\boldsymbol{p+q},\mathbf{0.99}}}, \mathbf{I}_q$) and $(p, q) =$ (8,4)

	Slope MSE	Intercept MSE	diagonal element mse	non diagonal elements mse	slope bias	intercept bias	diagonal element bias	non diagonal element bias
		$n = 50$	$\delta = 10\%$	$p = 8$	$q = 4$			
S_n	0.030228	0.027111	0.709027	0.019555	0.000525	2.45e-3	0.817944	-0.00341
MCD	0.071942	0.042172	0.466495	0.026127	2.23e-3	-1.87e-3	0.642774	-0.00385
Comedian	0.036868	0.029341	0.602844	0.018968	-2.20e-4	0.002407	0.74797	0.001453
OGK	0.037403	0.030711	0.606282	0.01967	0.000517	0.006056	0.751147	-0.00039
RSh-S_n	0.030355	0.027533	0.66581	0.019232	-0.00012	1.43e-3	0.791538	-0.00179
		$n = 100$	$\delta = 10\%$	$p = 8$	$q = 4$			
S_n	0.012977	0.012155	0.819943	0.010506	0.000567	-0.0014	0.894061	0.006268
MCD	0.018887	0.012857	0.642024	0.013133	1.03e-3	-0.0009	0.784671	0.008915
Comedian	0.014252	0.012338	0.749359	0.010878	-0.00105	0.000931	0.852171	0.009796
OGK	0.014383	0.012097	0.749174	0.010909	6.13e-4	-0.0002	0.85233	0.009278
RSh-S_n	0.013304	0.011938	0.791517	0.010294	2.01e-4	0.003988	0.877971	0.004615
		$n = 200$	$\delta = 10\%$	$p = 8$	$q = 4$			
S_n	0.006025	0.005856	0.878332	0.005491	-7.90e-5	0.000477	0.931262	0.002079
MCD	0.006612	0.005562	0.809249	0.006022	4.23e-4	-2.70e-3	0.892674	0.001315
Comedian	0.006405	0.005796	0.827204	0.005455	3.07e-4	1.76e-3	0.903246	0.003814
OGK	0.006369	0.005742	0.835858	0.005715	-6.10e-4	0.001927	0.907761	0.004887
RSh-S_n	0.006235	0.00583	0.857884	0.005459	-0.00031	0.000102	0.920022	0.003575
		$n = 500$	$\delta = 10\%$	$p = 8$	$q = 4$			
S_n	0.00234	0.00226	0.916931	0.002298	-4.70e-4	-5.70e-4	0.95514	0.001272
MCD	0.002443	0.002189	0.886638	0.002319	6.42e-5	-0.00048	0.939115	0.000909
Comedian	0.002436	0.002232	0.88465	0.0023	-2.40e-4	-0.00032	0.937966	0.001095
OGK	0.002441	0.002285	0.901099	0.002305	-1.20e-4	7.61e-4	0.946845	0.002174
RSh-S_n	0.002357	0.002254	0.897109	0.00224	2.94e-5	-0.00028	0.944794	0.000946

Table B.7: Robustness table-Contaminants from N($\mathbf{2}\sqrt{\boldsymbol{\chi}^{\mathbf{2}}_{\boldsymbol{p+q},\mathbf{0.99}}}, \mathbf{I}_q$) and $(p,q) =$ (10,10)

	Slope MSE	Intercept MSE	diagonal element mse	non diagonal elements mse	slope bias	intercept bias	diagonal element bias	non diagonal element bias
		$n = 50$	$\delta = 10\%$	$p = 10$	$q = 10$			
S_n	0.035186	0.033186	0.628093	0.019893	-0.00029	1.92e-3	0.766835	0.026861
MCD	0.05331	0.045228	0.548384	0.024382	-4.90e-4	9.81e-4	0.708176	0.036003
Comedian	0.031939	0.030156	0.674917	0.019562	-2.60e-4	-0.00049	0.797826	0.024018
OGK	0.041184	0.036058	0.581502	0.020469	0.000189	0.003186	0.734167	0.030117
RSh-S_n	0.032333	0.029723	0.637209	0.018516	-0.0005	1.36e-3	0.774981	0.022447
		$n = 100$	$\delta = 10\%$	$p = 10$	$q = 10$			
S_n	0.013988	0.013156	0.766434	0.010543	-2.20e-5	0.000967	0.863244	0.011695
MCD	0.023567	0.017868	0.602532	0.013419	8.36e-4	0.000879	0.758948	0.016278
Comedian	0.013097	0.012964	0.818361	0.010321	-0.00027	2.16e-5	0.893156	0.011145
OGK	0.015392	0.01365	0.720919	0.010768	-1.90e-4	0.004502	0.835836	0.01286
RSh-S_n	0.013361	0.012747	0.771019	0.010076	2.26e-4	0.002149	0.866525	0.009514
		$n = 200$	$\delta = 10\%$	$p = 10$	$q = 10$			
S_n	0.006289	0.005989	0.857448	0.005516	3.12e-4	0.002718	0.919868	0.004935
MCD	0.007359	0.00635	0.773152	0.006109	3.28e-4	2.85e-5	0.872256	0.005688
Comedian	0.006057	0.005785	0.886198	0.005354	-8.00e-5	1.56e-3	0.935541	0.004612
OGK	0.006611	0.006031	0.829817	0.005625	1.22e-4	0.001466	0.90463	0.005339
RSh-S_n	0.006175	0.005768	0.860194	0.005371	-0.0003	0.00143	0.921546	0.005657
		$n = 500$	$\delta = 10\%$	$p = 10$	$q = 10$			
S_n	0.002381	0.002312	0.91882	0.002253	-3.20e-4	4.00e-4	0.95622	0.001573
MCD	0.002433	0.002329	0.902059	0.002315	-4.00e-4	-0.00067	0.947311	0.002049
Comedian	0.00234	0.002267	0.935814	0.002231	-8.50e-5	0.000351	0.965056	0.001224
OGK	0.002441	0.002285	0.901099	0.002305	-1.20e-4	7.61e-4	0.946845	0.002174
RSh-S_n	0.002338	0.002314	0.920812	0.00223	-1.60e-5	0.001107	0.957193	0.001919

Table B.8: Robustness table-Contaminants from N($\mathbf{2}\sqrt{\boldsymbol{\chi}^{\mathbf{2}}_{\boldsymbol{p+q},\mathbf{0.99}}}, \mathbf{I}_q$), (p, q)=(4,4) and δ=20%

	Slope MSE	Intercept MSE	diagonal element mse	non-diagonal elements mse	slope bias	intercept bias	diagonal element bias	non diagonal element bias
		$n = 50$	$\delta = 20\%$	$p = 4$	$q = 4$			
S_n	0.0302	0.02727	0.85839	0.02415	0.00093	0.00073	0.89935	-0.00289
MCD	0.05076	0.0412	0.68838	0.02982	-0.00125	0.00157	0.79067	-0.00293
Comedian	0.02994	0.02718	0.85414	0.02346	-0.00014	-0.00159	0.8974891	-0.00478
OGK	0.03399	0.02742	0.77882	0.024663	-0.00078	0.00538	0.85393	-0.00113
RSh-S_n	0.02979	0.02737	0.83589	0.02353	0.00014	0.00471	0.88700	-0.0056
		$n = 100$	$\delta = 20\%$	$p = 4$	$q = 4$			
S_n	0.01373	0.01292	0.88807	0.01219	0.00061	-0.00023	0.92924	0.00928
MCD	0.01745	0.01693	0.79481	0.0147	-0.00118	-0.00137	0.87478	0.01117
Comedian	0.01382	0.01298	0.92580	0.01234	0.00022	-0.00112	0.94930	0.00698
OGK	0.014927	0.01339	0.94275	0.01302	0.00294	0.00091	0.95761	0.01195
RSh-S_n	0.01379	0.01297	0.87333	0.01210	0.00090	-0.00109	0.92097	0.00568
		$n = 200$	$\delta = 20\%$	$p = 4$	$q = 4$			
S_n	0.00681	0.00629	0.91992	0.00627	0.00026	0.00025	0.95231	0.00228
MCD	0.00701	0.00692	0.88238	0.00677	-0.00074	0.00042	0.93189	0.00254
Comedian	0.00639	0.00632	0.96182	0.00609	0.00015	0.00241	0.97431	0.00436
OGK	0.00704	0.00639	0.96717	0.00653	0.00082	0.002191	0.97723	0.00609
RSh-S_n	0.00683	0.00646	0.90559	0.00614	0.00047	7.08e-5	0.94482	0.00089
		$n = 500$	$\delta = 20\%$	$p = 4$	$q = 4$			
S_n	0.00262	0.00247	0.93383	0.00255	-0.00026	0.00099	0.96364	0.00039
MCD	0.00267	0.00269	0.93775	0.00261	0.00039	-0.00068	0.96565	0.00077
Comedian	0.00258	0.00256	0.98175	0.00255	-0.00049	0.00077	0.98826	0.00242
OGK	0.00263	0.00255	0.97934	0.00264	0.00036	0.00033	0.98715	0.00355
RSh-S_n	0.00266	0.00246	0.91953	0.00251	0.00089	6.15e-5	0.95616	0.00198

Table B.9: Robustness table-Contaminants from N($\mathbf{2}\sqrt{\boldsymbol{\chi}^{\mathbf{2}}_{\boldsymbol{p+q},\mathbf{0.99}}}, \mathbf{I}_q$),$(p, q)$ = (4,4) and $\delta = 40\%$

	Slope MSE	Intercept MSE	diagonal element mse	non-diagonal elements mse	slope bias	intercept bias	diagonal element bias	non diagonal element bias
		$n = 50$	$\delta = 40\%$	$p = 4$	$q = 4$			
S_n	0.35211	2.12064	88.3139	74.9682	0.02736	1.10210	7.57509	6.69784
MCD	0.06166	0.14459	5.10457	3.83927	0.00316	0.04959	1.16849	0.29616
Comedian	0.07304	0.20331	8.23891	6.28624	0.01107	0.09460	1.4986	0.61259
OGK	0.24763	2.06194	68.0633	55.9697	0.17555	1.12809	6.87226	5.96703
RSh-S_n	0.04199	0.03976	0.83585	0.03122	0.00137	-0.00386	0.87939	-0.00342
		$n = 100$	$\delta = 40\%$	$p = 4$	$q = 4$			
S_n	0.17121	2.47706	106.977	91.9338	0.05115	1.27435	8.73594	7.85035
MCD	0.02066	0.03643	1.70257	0.75216	0.00331	0.00620	0.97224	0.06375
Comedian	0.02559	0.08129	3.66991	2.36436	0.00931	0.03717	1.16523	0.24823
OGK	0.15370	2.61437	92.12617	76.86808	0.19075	1.38645	8.52452	7.60009
RSh-S_n	0.01897	0.01790	0.88176	0.01688	0.00095	0.00061	0.92145	0.00965
		$n = 200$	$\delta = 40\%$	$p = 4$	$q = 4$			
S_n	0.10238	2.98402	126.0541	108.3557	0.07587	1.51423	10.105	9.19703
MCD	0.00891	0.00863	0.93663	0.00853	-0.00113	0.00138	0.95895	0.00499
Comedian	0.010701	0.03056	1.90188	0.82885	0.00623	0.01251	1.04234	0.09413
OGK	0.10160	3.07438	111.7085	93.8952	0.18774	1.59264	9.79609	8.85033
RSh-S_n	0.00910	0.00820	0.90489	0.00829	-0.00114	-0.00255	0.94256	0.00521
		$n = 500$	$\delta = 40\%$	$p = 4$	$q = 4$			
S_n	0.05855	3.46906	142.8108	122.8374	0.11323	1.74201	11.3629	10.4506
MCD	0.00349	0.008536	1.08623	0.14533	-0.00039	0.00049	0.96910	0.01381
Comedian	0.00454	0.00557	1.0357	0.07907	0.00501	0.00256	0.97785	0.01777
OGK	0.06420	3.80079	137.7814	116.8007	0.1799	1.88859	11.4503	10.4991
RSh-S_n	0.00352	0.00320	0.92963	0.00334	0.00061	-0.00072	0.96053	0.00297

Table B.10: Robustness table-Contaminants from N$(\mathbf{2}\sqrt{\boldsymbol{\chi}^{\mathbf{2}}_{\boldsymbol{p+q},\mathbf{0.99}}}, \mathbf{I}_q)$,$(p,q) = (4,8)$ and $\delta = 20\%$

	Slope MSE	Intercept MSE	diagonal element mse	non-diagonal elements mse	slope bias	intercept bias	diagonal element bias	non diagonal element bias
		$n = 50$	$\delta = 20\%$	$p = 4$	$q = 8$			
S_n	0.02985	0.02757	0.89167	0.02372	0.00022	-0.00246	0.91829	-0.00171
MCD	0.04564	0.03866	0.74667	0.02948	-0.00042	-0.00168	0.82991	-0.00268
Comedian	0.03395	0.03111	0.89529	0.02369	-0.00155	-0.00172	0.920004	-0.00703
OGK	0.02945	0.02885	0.90123	0.02387	-0.0008	-0.00046	0.92355	-0.00212
RSh-S_n	0.03051	0.02776	0.86582	0.02344	-0.00076	0.00254	0.90531	-0.00123
		$n = 100$	$\delta = 20\%$	$p = 4$	$q = 8$			
S_n	0.01373	0.01317	0.92738	0.01242	0.00013	0.00077	0.95016	0.00211
MCD	0.01716	0.01613	0.80859	0.01427	0.00019	-0.00156	0.88370	0.00402
Comedian	0.01364	0.013508	0.94392	0.012340	0.00038	0.00130	0.95839	0.00257
OGK	0.01362	0.01348	0.93933	0.01223	0.00047	-0.00146	0.95648	0.00172
RSh-S_n	0.01373	0.01265	0.89824	0.01202	-2.50e-5	5.39e-5	0.93452	0.00214
		$n = 200$	$\delta = 20\%$	$p = 4$	$q = 8$			
S_n	0.00653	0.00625	0.95156	0.006172	-6.10e-5	0.00399	0.96932	0.00774
MCD	0.00707	0.00704	0.90170	0.00668	0.00083	-0.00075	0.94239	0.00538
Comedian	0.00668	0.00643	0.97214	0.006298	0.000133	0.000511	0.97970	0.00521
OGK	0.006512	0.00639	0.96979	0.00618	0.00032	0.00096	0.97843	0.00433
RSh-S_n	0.00671	0.00641	0.92568	0.00618	0.00026	0.00154	0.95560	0.00438
		$n = 500$	$\delta = 20\%$	$p = 4$	$q = 8$			
S_n	0.00272	0.00263	0.95723	0.00253	-0.00107	-0.00369	0.97559	-3.00e-5
MCD	0.00260	0.00263	0.95060	0.00256	9.74e-5	0.00022	0.97231	0.00084
Comedian	0.002603	0.00253	0.95310	0.00254	0.000723	0.00048	0.96880	0.00039
OGK	0.00252	0.00257	0.99024	0.00252	0.00024	0.00019	0.99262	0.00054
RSh-S_n	0.00259	0.00253	0.94393	0.00254	0.00035	0.00048	0.96880	0.00039

Table B.11: Robustness table-Contaminants from N($\mathbf{2}\sqrt{\boldsymbol{\chi}^{\mathbf{2}}_{\boldsymbol{p+q},\mathbf{0.99}}}$, $\mathbf{I}_q$),(p, q) = (4,8) and δ = 40%

	Slope MSE	Intercept MSE	diagonal element mse	non-diagonal elements mse	slope bias	intercept bias	diagonal element bias	non diagonal element bias
		$n = 50$	$\delta = 40\%$	$p = 4$	$q = 8$			
S_n	0.043008	0.03916	0.99702	0.13393	0.00179	0.00229	0.90948	0.00685
MCD	1.10588	4.99271	153.939	138.717	0.03637	1.90590	10.9517	10.2712
Comedian	0.04310	0.03926	0.8802	0.05615	0.00159	0.00301	0.90338	0.00150
OGK	0.13148	0.25842	11.0502	7.9792	0.2029	0.1882	2.07959	1.1726
RSh-S_n	0.04175	0.03909	0.8569	0.03099	-0.0019	-0.00078	0.89257	-0.00197
		$n = 100$	$\delta = 40\%$	$p = 4$	$q = 8$			
S_n	0.01849	0.01784	0.91315	0.01628	-0.00064	-0.0013	0.93828	0.00251
MCD	0.14467	1.90623	108.966	98.25095	0.09733	0.77056	6.46039	5.56713
Comedian	0.01890	0.01796	0.92561	0.01641	0.001394	-0.00103	0.94517	0.00326
OGK	0.12073	0.45787	24.4393	18.8932	0.22968	0.36320	3.39921	2.47033
RSh-S_n	0.01879	0.01789	0.91176	0.01632	-0.0015	0.00139	0.93776	0.00265
		$n = 200$	$\delta = 40\%$	$p = 4$	$q = 8$			
S_n	0.00877	0.00813	0.94621	0.00831	-0.00386	-0.00368	0.96353	0.00680
MCD	0.04710	1.13776	65.92517	58.908	0.06842	0.45254	4.20805	3.26446
Comedian	0.009016	0.00862	0.96128	0.008340	-0.000217	0.00154	0.97213	0.00646
OGK	0.11755	0.81882	48.09808	38.7176	0.24354	0.63746	5.47545	4.53889
RSh-S_n	0.00895	0.00865	0.93599	0.00836	-0.00131	0.00097	0.95889	0.00547
		$n = 500$	$\delta = 40\%$	$p = 4$	$q = 8$			
S_n	0.00331	0.00323	0.96203	0.00337	-0.00127	0.00141	0.97745	0.00038
MCD	0.015927	0.52698	25.93215	22.43054	0.03849	0.21017	2.33193	1.37015
Comedian	0.003529	0.00340	0.98661	0.00337	0.00018	0.00072	0.98999	0.00084
OGK	0.10927	1.22138	80.4669	66.3807	0.22964	0.92963	7.98539	7.04439
RSh-S_n	0.00346	0.00341	0.95473	0.00336	-0.0008	0.00065	0.97356	0.00087

Table B.12: Robustness table-Contaminants from N($\mathbf{2}\sqrt{\boldsymbol{\chi}^{\mathbf{2}}_{\boldsymbol{p+q,0.99}}}$, $\mathbf{I}_q$), $(p, q) = (8{,}4)$ and $\delta = 20\%$

	Slope MSE	Intercept MSE	diagonal element mse	non-diagonal elements mse	slope bias	intercept bias	diagonal element bias	non diagonal element bias
		$n = 50$	$\delta = 20\%$	$p = 8$	$q = 4$			
S_n	0.03357	0.03127	0.69526	0.02230	-0.00191	0.00164	0.80639	-0.0047
MCD	0.06377	0.05434	0.52516	0.02753	0.00354	-0.0036	0.68487	-0.00553
Comedian	0.03336	0.03258	0.71466	0.02132	0.00042	0.00444	0.81985	-0.00564
OGK	0.03488	0.03058	0.79689	0.02399	-0.00089	0.00871	0.86366	-0.00098
RSh-S_n	0.03463	0.03043	0.65955	0.02023	-0.00206	0.00436	0.78517	-0.00814
		$n = 100$	$\delta = 20\%$	$p = 8$	$q = 4$			
S_n	0.01464	0.01365	0.80411	0.01147	0.00133	-0.0013	0.88316	0.0069
MCD	0.02019	0.01851	0.65954	0.01404	0.00051	0.00120	0.79408	0.00859
Comedian	0.014704	0.01413	0.85518	0.01197	-0.0006	0.00132	0.91151	0.00620
OGK	0.01461	0.01391	0.84919	0.01187	0.00425	0.00471	0.91644	0.00988
RSh-S_n	0.01455	0.01312	0.78861	0.01169	7.36e-5	0.00088	0.87456	0.00883
		$n = 200$	$\delta = 20\%$	$p = 8$	$q = 4$			
S_n	0.00689	0.00686	0.87209	0.00605	-9.70e-5	-0.00217	0.92732	0.00274
MCD	0.00768	0.00738	0.80849	0.00679	0.00034	-0.00021	0.89146	0.00590
Comedian	0.00681	0.00655	0.87193	0.00623	0.00038	0.00057	0.92701	0.00206
OGK	0.00693	0.00681	0.91868	0.00613	0.00296	0.00237	0.95222	0.00413
RSh-S_n	0.00689	0.00686	0.87209	0.00605	-9.71e-5	-0.00217	0.92732	0.00274
		$n = 500$	$\delta = 20\%$	$p = 8$	$q = 4$			
S_n	0.00264	0.00264	0.91704	0.00259	-0.00021	0.00059	0.95489	0.00075
MCD	0.00270	0.00268	0.90326	0.00254	-0.00012	0.00014	0.94763	0.00186
Comedian	0.00267	0.00258	0.96727	0.00249	0.00024	0.00078	0.98095	0.00151
OGK	0.00282	0.00255	0.96945	0.00255	0.00316	-0.00185	0.9821	0.00333
RSh-S_n	0.00269	0.00252	0.90211	0.00253	-7.60e-5	0.00102	0.9472	0.00198

Table B.13: Robustness table-Contaminants from N($\mathbf{2}\sqrt{\boldsymbol{\chi}^{\mathbf{2}}_{\boldsymbol{p+q,0.99}}}$, $\mathbf{I}_q$),(p, q) = (8,4) and $\delta = 40\%$

	Slope MSE	Intercept MSE	diagonal element mse	non-diagonal elements mse	slope bias	intercept bias	diagonal element bias	non diagonal element bias
		$n = 50$	$\delta = 40\%$	$p = 8$	$q = 4$			
S_n	0.65477	4.49132	187.4866	166.8014	0.00732	1.91785	12.61918	11.8209
MCD	1.10588	4.99271	153.9388	138.7171	0.03637	1.90590	10.95174	10.27123
Comedian	0.08008	0.19894	6.76004	5.36965	0.00618	0.07377	1.24657	0.48109
OGK	0.26099	1.41466	51.98879	43.6377	0.07299	0.78048	5.23331	4.42547
RSh-S_n	0.05081	0.04479	0.62296	0.02713	0.00044	0.00141	0.75465	-0.01123
		$n = 100$	$\delta = 40\%$	$p = 8$	$q = 4$			
S_n	0.32181	5.07361	265.0977	238.2694	0.02099	2.16953	15.7887	14.9364
MCD	0.24450	3.08433	154.1151	139.5274	0.00349	1.27764	9.44718	8.64571
Comedian	0.02494	0.07119	3.67962	2.58406	0.00243	0.02834	1.05494	0.20349
OGK	0.15232	1.84470	84.32217	71.8368	0.07698	1.02136	7.36655	6.48206
RSh-S_n	0.02031	0.01889	0.76514	0.01541	-0.00141	0.00271	0.85653	0.01131
		$n = 200$	$\delta = 40\%$	$p = 8$	$q = 4$			
S_n	0.13881	5.63183	302.1547	271.9262	0.02784	2.34434	17.1804	16.2865
MCD	0.05585	1.92001	108.7991	98.1236	0.00615	0.77234	6.43164	5.52993
Comedian	0.00980	0.01175	1.05728	0.17460	0.00120	0.00140	0.94323	0.026271
OGK	0.09047	2.24832	110.871	94.70453	0.08036	1.24512	9.08741	8.15866
RSh-S_n	0.00927	0.00871	0.85027	0.00785	0.00032	0.001097	0.91304	0.00552
		$n = 500$	$\delta = 40\%$	$p = 8$	$q = 4$			
S_n	0.04544	6.1107	300.593	269.6979	0.0551	2.4702	17.291	16.376
MCD	0.01287	1.34594	74.024	66.1225	0.01029	0.53375	4.6888	3.73511
Comedian	0.00399	0.00338	0.96841	0.00372	0.00695	0.00278	0.98051	0.00804
OGK	0.04652	2.7300	133.465	114.091	0.08045	1.4778	10.582	9.6279
RSh-S_n	0.00360	0.00342	0.90273	0.00332	8.02e-6	0.00049	0.94657	0.00132

Table B.14: Robustness table-Contaminants from N($\mathbf{2}\sqrt{\boldsymbol{\chi}^2_{\boldsymbol{p+q,0.99}}}$, $\mathbf{I}_q$),(p, q)=(10,10) and δ=20%

	Slope MSE	Intercept MSE	diagonal element mse	non-diagonal elements mse	slope bias	intercept bias	diagonal element bias	non-diagonal element bias
		$n = 50$	δ = 20%	p = 10	q =10			
S_n	0.03946	0.03737	0.62071	0.02136	5.87e-5	0.00402	0.76008	0.02979
MCD	0.18668	0.32411	17.407	14.6341	0.00333	0.20601	2.32459	1.6307
Comedian	0.03551	0.03353	0.65291	0.02126	0.00038	0.00109	0.78266	0.02788
OGK	0.04362	0.03995	0.65303	0.02217	0.00077	0.00833	0.78249	0.03345
RSh-S_n	0.03619	0.03492	0.62575	0.02059	-0.00072	0.00279	0.76535	0.02836
		$n = 100$	δ = 20%	p = 10	q =10			
S_n	0.01591	0.01496	0.76640	0.01161	-0.00027	0.00164	0.86172	0.01279
MCD	0.02294	0.02076	0.95109	0.28122	0.00022	0.00114	0.80914	0.040083
Comedian	0.01493	0.01445	0.81212	0.01136	0.00030	0.00124	0.88852	0.01163
OGK	0.01685	0.01506	0.73256	0.01172	-0.00034	0.00468	0.84174	0.013794
RSh-S_n	0.01519	0.01442	0.77330	0.011306	-0.00016	0.00294	0.86661	0.01135
		$n = 200$	δ = 20%	p = 10	q =10			
S_n	0.01567	0.014839	0.76146	0.011589	0.00031	0.00276	0.85925	0.01267
MCD	0.00790	0.00772	0.87530	0.00656	0.00018	0.00203	0.92913	0.00656
Comedian	0.006871	0.00677	0.90201	0.00608	0.00041	0.00069	0.94349	0.00562
OGK	0.00727	0.00671	0.89834	0.00625	0.00015	0.00209	0.94151	0.00591
RSh-S_n	0.00698	0.00658	0.85951	0.00602	-0.00032	0.00209	0.92048	0.00481
		$n = 500$	δ = 20%	p = 10	q =10			
S_n	0.00268	0.00259	0.92724	0.00253	-0.00128	-0.00194	0.96014	0.00214
MCD	0.00271	0.00277	0.91255	0.00257	0.00017	-0.00036	0.95256	0.00228
Comedian	0.00262	0.00254	0.95852	0.00248	0.00015	0.00053	0.97657	0.00248
OGK	0.00258	0.00255	0.95926	0.00246	0.00254	0.00028	0.97687	0.00224
RSh-S_n	0.00266	0.00257	0.92344	0.00249	0.00029	0.00035	0.95829	0.00155

Table B.15: Robustness table-Contaminants from N($\mathbf{2}\sqrt{\boldsymbol{\chi}^{\mathbf{2}}_{\boldsymbol{p+q},\mathbf{0.99}}}$, $\mathbf{I}_q$),(p, q) = (10,10) and δ = 40%

	Slope MSE	Intercept MSE	diagonal element mse	non diagonal elements mse	slope bias	intercept bias	diagonal element bias	non diagonal element bias
		$n = 50$	δ = 40%	$p = 10$	$q = 10$			
S_n	0.05721	0.0522	0.56855	0.02759	0.00254	0.00557	0.71874	0.04188
MCD	0.89826	6.00411	230.46	213.62	0.05921	1.87974	12.04471	11.35095
Comedian	0.05617	0.05185	0.57761	0.02743	0.00093	-0.00464	0.72452	0.04005
OGK	0.06095	0.05666	0.66691	0.11091	0.00029	0.00393	0.72788	0.05581
RSh-S_n	0.05591	0.05162	0.57793	0.02738	-0.00076	-0.00063	0.72355	0.03901
		$n = 100$	δ = 40%	$p = 10$	$q = 10$			
S_n	0.02139	0.02036	0.74991	0.01495	-2.60e-6	0.00083	0.84851	0.01577
MCD	0.38991	5.31591	335.99	313.521	0.04517	1.7746	14.483	13.705
Comedian	0.02116	0.02048	0.85555	0.01489	0.00064	0.00112	0.91616	0.01659
OGK	0.02666	0.02482	1.2451	0.27052	0.07492	0.01901	0.98575	0.13838
RSh-S_n	0.02095	0.02005	0.74579	0.01474	0.00016	0.00081	0.84621	0.01597
		$n = 200$	δ = 40%	$p = 10$	$q = 10$			
S_n	0.00949	0.00935	0.84754	0.00793	3.41e-5	0.00128	0.91195	0.00718
MCD	0.16635	5.79316	459.79	429.114	0.02545	1.9671	17.822	16.948
Comedian	0.00933	0.00938	0.87223	0.00791	-3.00e-5	-0.00136	0.92555	0.00757
OGK	0.01663	0.01355	1.34789	0.22998	0.08342	0.01624	1.0632	0.14346
RSh-S_n	0.00939	0.00891	0.85125	0.0078	-0.00016	-0.00186	0.913901	0.00616
		$n = 500$	δ = 40%	$p = 10$	$q = 10$			
S_n	0.00351	0.00353	0.91729	0.00345	0.00044	0.00435	0.95469	0.00199
MCD	0.06606	7.41780	635.76	593.08	0.02723	2.4625	22.963	22.017
Comedian	0.00354	0.00348	0.94594	0.00333	0.00013	-0.00051	0.96919	0.00297
OGK	0.01203	0.00558	1.3406	0.14947	0.09007	0.01748	1.0991	0.14387
RSh-S_n	0.00353	0.00344	0.91675	0.00331	-0.00044	0.00113	0.95409	0.00257

Table B.16: Robustness table-Contaminants from N($\mathbf{2}\sqrt{\boldsymbol{\chi}^{\mathbf{2}}_{\boldsymbol{p+q,0.99}}}, 0.1 \times \mathbf{I}_q$), (p, q) = (4,4) and δ = 20%

	Slope MSE	Intercept MSE	diagonal element mse	non-diagonal elements mse	slope bias	intercept bias	diagonal element bias	non-diagonal element bias
		$n = 50$	$\delta = 20\%$	$p = 4$	$q = 4$	diag = 0.1		
S_n	0.03075	0.02774	0.85563	0.02364	0.00018	-0.00308	0.89928	-0.00583
MCD	0.05089	0.03241	0.65104	0.03019	-0.00204	0.00068	0.76749	-0.00694
Comedian	0.03048	0.02779	0.86394	0.02443	0.00011	0.00439	0.90155	-0.00505
OGK	0.03454	0.02974	0.78734	0.02376	-4.00e-5	0.00607	0.85762	-0.0025
RSh-S_n	0.03024	0.02771	0.82304	0.02354	-0.00019	0.00457	0.88019	-0.0057
		$n = 100$	$\delta = 20\%$	$p = 4$	$q = 4$	diag = 0.1		
S_n	0.014313	0.013099	0.901658	0.012564	0.000809	-0.00017	0.93652	0.00592
MCD	0.016517	0.013104	0.889987	0.013742	-0.00132	-0.00087	0.942827	0.005901
Comedian	0.014158	0.013117	0.902919	0.012439	-0.0012	-0.00032	0.936348	0.00573
OGK	0.014642	0.013542	0.897704	0.012663	0.000848	0.000761	0.936497	0.008697
RSh-S_n	0.014138	0.013011	0.882499	0.012114	0.000674	-5.80e-5	0.92597	0.005487
		$n = 200$	$\delta = 20\%$	$p = 4$	$q = 4$	diag = 0.1		
S_n	0.00669	0.00657	0.91924	0.00641	-0.00011	0.00284	0.95191	0.00348
MCD	0.00719	0.00644	0.87161	0.00672	-0.0013	0.00117	0.92621	0.00484
Comedian	0.00676	0.00635	0.91669	0.00634	0.00015	-0.00092	0.95086	0.00467
OGK	0.00699	0.00619	0.88655	0.00671	4.93e-5	0.00189	0.93447	0.00601
RSh-S_n	0.00661	0.00632	0.90445	0.00629	8.74e-5	0.00249	0.94411	0.00081
		$n = 500$	$\delta = 20\%$	$p = 4$	$q = 4$	diag = 0.1		
S_n	0.00257	0.00241	0.93676	0.00259	0.000414	-4.50e-5	0.96514	0.00081
MCD	0.00269	0.00251	0.93088	0.00256	-0.00019	-0.00012	0.96203	0.00057
Comedian	0.00263	0.00243	0.93226	0.00252	-0.00039	-9.76e-5	0.96284	1.97e-5
OGK	0.00269	0.00243	0.90249	0.00266	8.45e-5	5.13e-5	0.94727	0.00279
RSh-S_n	0.00244	0.00235	0.91845	0.00256	3.89e-4	0.00115	0.95566	0.00021

Table B.17: Robustness table-Contaminants from N($\mathbf{2}\sqrt{\boldsymbol{\chi}^{\mathbf{2}}_{\boldsymbol{p+q},\mathbf{0.99}}}$, $0.1 \times \mathbf{I}_q$), (p,q) = (4,8) and δ = 20%

	Slope MSE	Intercept MSE	diagonal element MSE	non-diagonal elements MSE	slope bias	intercept bias	diagonal element bias	non-diagonal element bias
		$n = 50$	δ = 20%	$p = 4$	$q = 8$	diag = 0.1		
S_n	0.034704	0.031897	0.764784	0.023529	-0.00084	0.000461	0.899752	0.026747
MCD	0.047819	0.036478	0.724979	0.028639	-0.00086	7.38e-6	0.816456	-0.00208
Comedian	0.02924	0.028465	0.895826	0.023991	0.000321	-0.00228	0.92042	-0.00292
OGK	0.034667	0.030594	0.802752	0.023804	0.000448	0.008045	0.866346	-0.00156
RSh-S_n	0.029632	0.027593	0.75648	0.023149	-8.30e-5	0.0004134	0.846933	-0.00268
		$n = 100$	δ = 20%	$p = 4$	$q = 8$	diag = 0.1		
S_n	0.01377	0.01327	0.93041	0.01244	0.00097	-0.00168	0.95109	0.00331
MCD	0.01817	0.01458	0.78737	0.01432	-0.00059	-0.001	0.87088	0.00275
Comedian	0.01356	0.01326	0.92725	0.01231	0.00092	0.0008	0.95011	0.00296
OGK	0.01514	0.01333	0.86102	0.01263	-7.80e-6	0.00287	0.91401	0.00431
RSh-S_n	0.01367	0.01325	0.89929	0.01198	0.00099	0.00208	0.93521	0.00234
		$n = 200$	δ = 20%	$p = 4$	$q = 8$	diag = 0.1		
S_n	0.00664	0.00637	0.94831	0.00628	0.00036	-2.70e-5	0.96723	0.00535
MCD	0.00735	0.00653	0.88217	0.00655	0.00017	5.53e-5	0.93195	0.00459
Comedian	0.00659	0.00627	0.94099	0.00627	0.00031	0.00028	0.96357	0.00457
OGK	0.00690	0.00647	0.90551	0.00634	1.61e-5	0.00259	0.94473	0.00447
RSh-S_n	0.00675	0.00637	0.92623	0.00618	-0.00089	0.00310	0.95562	0.00431
		$n = 500$	δ = 20%	$p = 4$	$q = 8$	diag = 0.1		
S_n	0.00261	0.00247	0.95209	0.00253	-8.20e-5	-9.50e-5	0.97302	0.00129
MCD	0.00269	0.00251	0.93088	0.00256	-0.00019	-0.00012	0.96203	0.00057
Comedian	0.00262	0.00251	0.95425	0.00259	0.00039	-0.00074	0.97421	0.00023
OGK	0.00267	0.00253	0.92965	0.00257	-1.00e-5	8.10e-5	0.96142	0.0014
RSh-S_n	0.00254	0.00246	0.94409	0.00252	0.00011	0.00088	0.96899	9.38e-6

Table B.18: Robustness table-Contaminants from N($\mathbf{2}\sqrt{\boldsymbol{\chi}^{\mathbf{2}}_{\boldsymbol{p+q},\mathbf{0.99}}}, 0.1 \times \mathbf{I}_q$),$(p, q) =$ (8,4) and $\delta = 20\%$

	Slope MSE	Intercept MSE	diagonal element MSE	non-diagonal elements MSE	slope bias	intercept bias	diagonal element bias	non-diagonal element bias
		$n = 50$	$\delta = 20\%$	$p = 8$	$q = 4$	diag $= 0.1$		
S_n	0.03374	0.03345	0.69862	0.02149	0.00061	0.00055	0.80972	-0.00483
MCD	0.06381	0.04222	0.50153	0.02572	0.00069	0.00162	0.67153	-0.00264
Comedian	0.03402	0.03093	0.70009	0.02088	0.00139	0.00281	0.80951	-0.0049
OGK	0.03992	0.03290	0.61369	0.02133	0.00193	0.00396	0.75319	-0.00154
RSh-S_n	0.03344	0.03091	0.67365	0.02055	-0.00067	-0.00028	0.74131	-0.00433
		$n = 100$	$\delta = 20\%$	$p = 8$	$q = 4$	diag=0.1		
S_n	0.01487	0.01405	0.80891	0.01198	0.00035	-0.00207	0.88622	0.00762
MCD	0.01943	0.01453	0.668330	0.01385	-0.00033	0.00242	0.800258	0.00817
Comedian	0.01472	0.01376	0.80632	0.01159	-0.00083	-0.00145	0.88429	0.00715
OGK	0.01583	0.01418	0.76464	0.01214	-0.00105	0.00111	0.86020	0.01186
RSh-S_n	0.01481	0.01335	0.77938	0.01139	-0.00018	0.00218	0.86949	0.00567
		$n = 200$	$\delta = 20\%$	$p = 8$	$q = 4$	diag $= 0.1$		
S_n	0.00684	0.00638	0.88094	0.00623	-0.00059	5.36e-5	0.93169	0.00291
MCD	0.00739	0.00642	0.81891	0.00663	-0.00014	0.00087	0.89748	0.00333
Comedian	0.00688	0.00639	0.87029	0.00611	-0.00074	-0.00126	0.92622	0.00334
OGK	0.00714	0.00636	0.83691	0.00637	-0.00024	0.00124	0.90767	0.00625
RSh-S_n	0.00681	0.00634	0.85894	0.000627	0.000105	0.00212	0.92005	0.00233
		$n = 500$	$\delta = 20\%$	$p = 8$	$q = 4$	diag $= 0.1$		
S_n	0.00269	0.00248	0.90819	0.00254	0.00024	0.00044	0.95025	0.00282
MCD	0.00271	0.00246	0.89083	0.00260	0.00034	0.00010	0.94093	0.00129
Comedian	0.00268	0.00251	0.91519	0.00258	0.00047	0.00101	0.95395	0.00099
OGK	0.00273	0.00255	0.89548	0.00262	-0.0002	0.00071	0.94343	0.00379
RSh-S_n	0.00261	0.00246	0.90256	0.00254	-0.00044	1.56e-3	0.94725	0.00167

Table B.19: Robustness table-Contaminants from N($\mathbf{2}\sqrt{\boldsymbol{\chi}^{\mathbf{2}}_{\boldsymbol{p+q},\mathbf{0.99}}}, 0.1 \times \mathbf{I}_q$),$(p, q) =$ (10,10) and δ = 20%

	Slope MSE	Intercept MSE	diagonal element MSE	non-diagonal elements MSE	slope bias	intercept bias	diagonal element bias	non diagonal element bias
		$n = 50$	δ = 20%	$p = 10$	$q = 10$	diag = 0.1		
S_n	0.03994	0.03749	0.62827	0.02142	0.00063	0.00494	0.75840	0.03005
MCD	0.08538	0.10234	1.99508	1.38534	0.06619	0.02868	0.81063	0.19452
Comedian	0.03685	0.03416	0.64861	0.02095	0.00058	0.00427	0.77916	0.02745
OGK	0.04426	0.04086	0.58296	0.02201	-0.00116	0.00674	0.73376	0.03467
RSh-S_n	0.03686	0.03417	0.62789	0.02058	-0.00056	0.00424	0.75979	0.02731
		$n = 100$	δ = 20%	$p = 10$	$q = 10$	diag = 0.1		
S_n	0.01586	0.01493	0.76322	0.01169	-0.00032	0.00245	0.86402	0.01219
MCD	0.02425	0.01867	0.63574	0.01409	0.01374	-0.00094	0.77980	0.02225
Comedian	0.01597	0.01414	0.79717	0.01157	0.00016	0.00165	0.87988	0.01213
OGK	0.01679	0.01506	0.73413	0.01162	-0.00038	0.00291	0.84276	0.01287
RSh-S_n	0.01515	0.01418	0.76848	0.01128	4.85e-5	0.00119	0.86394	0.01204
		$n = 200$	δ = 20%	$p = 10$	$q = 10$	diag = 0.1		
S_n	0.00715	0.00662	0.85347	0.00614	-2.50e-5	0.00082	0.91701	0.00569
MCD	0.00826	0.00719	0.77997	0.00667	0.00168	-0.0004	0.87561	0.00688
Comedian	0.00684	0.00667	0.88076	0.00603	0.00021	-7.50e-5	0.93209	0.00518
OGK	0.00735	0.00689	0.83814	0.006201	-0.0003	0.001344	0.90866	0.00658
RSh-S_n	0.00697	0.00671	0.86112	0.00596	-5.00e-5	0.00193	0.92139	0.00574
		$n = 500$	δ = 20%	$p = 10$	$q = 10$	diag = 0.1		
S_n	0.00269	0.00259	0.91858	0.00254	-0.00012	0.00128	0.95578	0.00251
MCD	0.00277	0.00262	0.901796	0.00261	2.45e-4	-4.10e-5	0.94686	0.00252
Comedian	0.00262	0.00264	0.93416	0.00251	-3.70e-5	-0.00032	0.96389	0.00224
OGK	0.00271	0.00255	0.90908	0.00252	-4.40e-5	-4.20e-5	0.95079	0.00236
RSh-S_n	0.00261	0.00255	0.91939	0.00248	0.00015	0.00049	0.95522	0.00206

Table B.20: Robustness table-Contaminants from N($\mathbf{2}\sqrt{\boldsymbol{\chi}^{\mathbf{2}}_{\boldsymbol{p+q,0.99}}}, 0.1 \times \mathbf{I}_q$), (p, q) = (10,10) and δ = 40%

	Slope MSE	Intercept MSE	diagonal element MSE	non-diagonal elements MSE	slope bias	intercept bias	diagonal element bias	non diagonal element bias
		$n = 50$	$\delta = 40\%$	$p = 10$	$q = 10$	diag = 0.1		
S_n	0.05844	0.05399	0.57708	0.02813	-0.00063	0.00351	0.72277	0.04215
MCD	2.4997	19.171	210.80	201.92	0.05781	4.0571	13.554	13.225
Comedian	0.05584	0.05243	0.58585	0.02813	-0.00035	-0.00265	0.72996	0.04069
OGK	0.17829	0.53677	26.002	23.626	0.00808	0.21098	2.1958	1.5336
RSh-S_n	0.05579	0.05097	0.57641	0.02753	-0.00089	0.00219	0.72227	0.04067
		$n = 100$	$\delta = 40\%$	$p = 10$	$q = 10$	diag = 0.1		
S_n	0.02148	0.02055	0.73583	0.01469	0.00010	-0.00014	0.84066	0.01669
MCD	1.43679	32.895	332.13	323.11	0.04217	5.5703	17.722	17.467
Comedian	0.02091	0.01974	0.75097	0.01494	-0.0002	-0.00162	0.84909	0.01676
OGK	0.14387	2.00717	136.05	127.88	0.00899	0.68192	6.35096	5.5712
RSh-S_n	0.02103	0.02060	0.74497	0.01489	-0.00081	-0.00059	0.84568	0.01622
		$n = 200$	$\delta = 40\%$	$p = 10$	$q = 10$	diag = 0.1		
S_n	0.009555	0.009109	0.846444	0.007927	0.00023	-0.00096	0.911263	0.007491
MCD	0.673562	53.76655	367.2779	360.0604	0.024282	7.294028	19.03284	18.84639
Comedian	0.009385	0.009103	0.854696	0.008016	-0.00013	-0.00075	0.915971	0.006963
OGK	0.143872	2.007173	136.0479	127.8799	0.008989	0.68192	6.350955	5.571183
RSh-S_n	0.009336	0.009174	0.849648	0.007916	0.000283	-0.00081	0.913315	0.007676
		$n = 500$	$\delta = 40\%$	$p = 10$	$q = 10$	diag = 0.1		
S_n	0.00355	0.00353	0.91521	0.00334	0.00011	0.000211	0.9532	0.00294
MCD	0.20230	79.031	271.08	266.38	7.79e-3	8.88e+0	16.401	16.260
Comedian	0.00354	0.00344	0.91997	0.00330	4.59e-5	-0.00011	0.95565	0.00311
OGK	0.06209	7.3105	532.14	504.957	3.03e-2	2.35e+0	20.2159	19.4717
RSh-S_n	0.00352	0.00348	0.92139	0.00329	-0.00018	0.00029	0.95642	0.00262

Table B.21: Robustness table-Contaminants from N($\mathbf{2}\sqrt{\boldsymbol{\chi}^{\mathbf{2}}_{\boldsymbol{p+q},\mathbf{0.99}}}, 0.1 \times \mathbf{I}_q$),$(p, q) =$ (8,4) and $\delta = 40\%$

	Slope MSE	Intercept MSE	diagonal element MSE	non-diagonal elements MSE	slope bias	intercept bias	diagonal element bias	non-diagonal element bias
		$n = 50$	$delta = 40\%$	$p = 8$	$q = 4$	diag = 0.1		
S_n	0.67604	4.7197	189.57	173.68	0.02041	1.93489	12.526	11.876
MCD	0.22249	0.72785	4.68828	4.22266	0.09912	0.16406	0.87398	0.40646
Comedian	0.08008	0.19894	6.76004	5.36965	0.00618	0.07377	1.2466	0.48109
OGK	0.47698	2.4181	97.769	88.577	0.0519	1.05517	7.2745	6.5795
RSh-S_n	0.05105	0.04801	0.63371	0.02721	-0.00129	-0.00042	0.760148	-0.00735
		$n = 100$	$delta = 40\%$	$p = 8$	$q = 4$	diag = 0.1		
S_n	0.30009	5.8859	290.31	268.66	0.00647	2.3327	16.444	15.771
MCD	0.05243	0.06361	0.69358	0.43759	0.10117	0.01547	0.51527	0.07769
Comedian	0.04822	0.41848	21.122	18.769	0.00071	0.17219	2.0267	1.1927
OGK	0.23910	3.63738	181.25	166.69	0.03763	1.50101	10.970	10.241
RSh-S_n	0.02046	0.01989	0.76972	0.01564	-0.00085	-0.00198	0.85912	0.00857
		$n = 200$	$delta = 40\%$	$p = 8$	$q = 4$	diag = 0.1		
S_n	0.13794	6.49049	323.99	300.54	0.01517	2.53085	17.8287	17.1642
MCD	0.02638	0.01075	0.28965	0.00773	0.10119	0.01033	0.52628	0.04516
Comedian	0.02253	0.49061	24.8709	22.2412	0.00659	0.19842	2.2407	1.3537
OGK	0.12026	4.7165	237.46	219.63	0.04184	1.8786	13.5439	12.8255
RSh-S_n	0.00927	0.00905	0.85261	0.00857	-0.00052	0.00115	0.91442	0.00469
		$n = 500$	$delta = 40\%$	$p = 8$	$q = 4$	diag = 0.1		
S_n	0.05049	6.7196	294.32	273.06	0.04917	2.5916	17.09053	16.4615
MCD	0.01503	0.00370	0.37857	0.00479	0.10117	0.00943	0.61087	0.04819
Comedian	0.01964	1.6889	77.773	71.273	0.01879	0.67522	5.2369	4.3807
OGK	0.050407	5.844764	249.5629	231.1968	0.061419	2.2787	14.96317	14.31361
RSh-S_n	0.00363	0.00331	0.89875	0.00327	0.00025	4.42e-4	0.94438	0.00127

Table B.22: Robustness table-Contaminants from N($\mathbf{2}\sqrt{\boldsymbol{\chi}^{\mathbf{2}}_{\boldsymbol{p+q},\mathbf{0.99}}}$, $0.1 \times \mathbf{I}_q$), (p, q) = (4,8) and δ = 40%

	Slope MSE	Intercept MSE	diagonal element MSE	non-diagonal elements MSE	slope bias	intercept bias	diagonal element bias	non diagonal element bias
		$n = 50$	$\delta = 40\%$	$p = 4$	$q = 8$	diag = 0.1		
S_n	0.04276	0.04142	1.00681	0.14032	0.00145	0.00236	0.91515	0.01164
MCD	1.61146	19.865	213.19	203.72	0.16089	4.24079	13.9311	13.5893
Comedian	0.04213	0.03914	0.86906	0.03128	-0.00198	3.15e-5	0.89736	-0.00523
OGK	0.34543	2.92446	125.41	115.05	0.08919	1.13075	7.9352	7.1756
RSh-S_n	0.04203	0.03844	0.86211	0.03121	-0.00217	0.00069	0.89494	-0.00511
		$n = 100$	$\delta = 40\%$	$p = 4$	$q = 8$	diag = 0.1		
S_n	0.01878	0.01762	0.90864	0.01637	0.00040	0.00051	0.93626	0.00254
MCD	0.79948	30.776	237.63	230.44	0.11793	5.5050	15.282	15.0549
Comedian	0.01867	0.03220	1.7074	0.74729	0.00051	0.00488	0.97882	0.04113
OGK	0.25305	5.8210	253.85	235.62	0.12353	2.1851	14.5669	13.9005
RSh-S_n	0.01876	0.01779	0.91178	0.01639	-9.50e-5	-0.00079	0.93767	0.003157
		$n = 200$	$\delta = 40\%$	$p = 4$	$q = 8$	diag = 0.1		
S_n	0.00907	0.00831	0.94042	0.00839	0.00074	-0.00114	0.96081	0.00623
MCD	0.32752	40.052	219.84	214.55	0.09014	6.3186	14.784	14.609
Comedian	0.00889	0.00841	0.94039	0.00836	0.00079	3.14e-5	0.96104	0.00689
OGK	0.14479	7.1026	282.49	263.29	0.17926	2.6188	16.513	15.919
RSh-S_n	0.00886	0.00854	0.93573	0.00826	-0.00122	0.00054	0.95858	0.00650
		$n = 500$	$\delta = 40\%$	$p = 4$	$q = 8$	diag = 0.1		
S_n	0.00357	0.00339	0.95067	0.00344	2.40e-4	-1.32e-3	0.97150	0.00135
MCD	0.12316	48.173	184.25	180.13	0.06909	6.9366	13.552	13.40124
Comedian	0.00347	0.00339	0.95503	0.00339	0.00042	0.00012	0.97365	7.16e-5
OGK	0.08471	7.4702	277.98	259.51	2.10e-1	2.72e+0	16.618	16.052
RSh-S_n	0.00349	0.00339	0.95122	0.00337	0.00022	-0.0002	0.97182	9.28e-4

Table B.23: Robustness table-Contaminants from N($\mathbf{2}\sqrt{\boldsymbol{\chi}^{\mathbf{2}}_{\boldsymbol{p+q},\mathbf{0.99}}}$, $0.1 \times \mathbf{I}_q$), (p,q) = (4,4) and δ = 40%

	Slope MSE	Intercept MSE	diagonal element MSE	non-diagonal elements MSE	slope bias	intercept bias	diagonal element bias	non diagonal element bias
		$n = 50$	δ = 40%	$p = 4$	$q = 4$	diag = 0.1		
S_n	0.34236	2.86147	107.304	95.8018	0.00815	1.3031	8.3696	7.6065
MCD	0.20895	0.6777	5.7594	5.1498	0.23056	0.17052	1.0439	0.53301
Comedian	0.07304	0.20331	8.2389	6.2862	0.01107	0.09460	1.4986	0.61259
OGK	0.37440	2.90728	105.80	94.846	0.07036	1.3161	8.3262	7.5881
RSh-S_n	0.04171	0.03936	0.84139	0.03175	0.000396	-0.00082	0.88122	-0.00816
		$n = 100$	δ = 40%	$p = 4$	$q = 4$	diag = 0.1		
S_n	0.16657	3.34178	126.76	114.233	0.03881	1.5040	9.5526	8.8165
MCD	0.09897	0.24681	2.68858	2.2208	0.23034	0.07856	0.79908	0.29075
Comedian	0.04405	0.58265	22.0002	19.1608	0.01630	0.26029	2.4131	1.5347
OGK	0.16882	3.80005	138.98	125.73	0.07553	1.6884	10.413	9.7134
RSh-S_n	0.01891	0.01788	0.87993	0.01652	0.00164	-0.00066	0.91983	0.00648
		$n = 200$	δ = 40%	$p = 4$	$q = 4$	diag = 0.1		
S_n	0.10285	3.6858	134.2794	121.1064	0.06407	1.66273	10.2703	9.5508
MCD	0.07246	0.01184	0.49299	0.02680	0.23952	0.02759	0.68878	0.13367
Comedian	0.03068	0.94529	31.8636	28.16068	0.03452	0.41796	3.18009	0.00467
OGK	0.10302	4.5527	151.715	137.491	0.13086	2.00746	11.6645	11.0154
RSh-S_n	0.00890	0.00871	0.91137	0.00847	0.00042	-0.00096	0.94556	0.00389
		$n = 500$	δ = 40%	$p = 4$	$q = 4$	diag = 0.1		
S_n	0.05775	4.29006	159.64	144.47	0.07819	1.91e+0	11.78005	11.0825
MCD	0.06499	0.00494	0.56339	0.02514	0.24577	0.02804	0.74579	0.14715
Comedian	0.03586	2.48335	78.2417	70.36449	0.10027	1.10e+0	6.6505	5.87e+0
OGK	0.05742	5.15367	159.75	145.08	1.66e-1	2.25e+0	12.521	11.917
RSh-S_n	0.00349	0.00339	0.93097	0.003426	2.55e-4	-0.00032	0.96137	0.00223

Table B.24: Efficiency table

	Slope	Intercept	Σ_{diag}	$\Sigma_{offdiag}$
	$n = 200$	$p = 4$	$q = 4$	
RSh-S_n	1.085685	1.019474	2.182346	1.095446
MLE	1.065408	1.081035	2.157384	1.139241
Comedian	1.16082	1.075198	2.532731	1.237662
MCD	1.310039	1.23455	3.135779	1.444395
OGK	1.158311	1.101057	2.477163	1.264569
S_n	1.077284	1.042145	2.115124	1.11251
	$n = 200$	$p = 4$	$q = 8$	
RSh-S_n	1.049875	1.045491	2.150233	1.088122
MLE	1.055721	1.046016	2.134762	1.06304
Comedian	1.126149	1.048017	2.458079	1.199587
MCD	1.261398	1.219316	2.735495	1.35651
OGK	1.13887	1.073517	2.40479	1.212931
S_n	1.049497	1.017444	2.055503	1.079099
	$n = 200$	$p = 8$	$q = 4$	
RSh-S_n	1.104645	1.018578	2.310298	1.161796
MLE	1.081907	1.101591	2.19617	1.152949
Comedian	1.170313	1.17985	2.584451	1.277201
MCD	1.392171	1.324717	3.285776	1.644481
OGK	1.183097	1.148603	2.584904	1.179488
S_n	1.099407	1.074058	2.24533	1.120149
	$n = 200$	$p = 10$	$q = 10$	
RSh-S_n	1.115416	1.026495	2.246079	1.133814
MLE	1.085531	1.062867	2.225373	1.098707
Comedian	1.180156	1.066586	2.417354	1.232591
MCD	1.448076	1.384457	3.186626	1.597878
OGK	1.186367	1.086911	2.585714	1.257713
S_n	1.11442	1.02788	2.22118	1.11389

Table B.25: Efficiency table

	Slope	Intercept	Σ_{diag}	$\Sigma_{offdiag}$
	$n = 500$	$p = 4$	$q = 4$	
$RSh\text{-}S_n$	1.067141	0.993516	2.291747	1.102374
MLE	1.05921	1.032378	2.275096	1.114671
Comedian	1.093142	1.039526	2.361259	1.23466
MCD	1.110777	1.07855	2.508448	1.275869
OGK	1.085246	1.100821	2.521479	1.210516
S_n	1.037956	1.02024	2.217112	1.089853
	$n = 500$	$p = 4$	$q = 8$	
$RSh\text{-}S_n$	1.039417	1.044132	2.084362	1.05972
MLE	1.036092	1.038818	2.118347	1.060237
Comedian	1.087771	1.040496	2.325303	1.162874
MCD	1.098278	1.09345	2.279732	1.163004
OGK	1.134125	1.041724	2.430282	1.234794
S_n	1.036082	0.982349	2.174474	1.080231
	$n = 500$	$p = 8$	$q = 4$	
$RSh\text{-}S_n$	1.083061	0.992563	2.245992	1.129318
MLE	1.056213	1.014551	2.321965	1.098386
Comedian	1.099334	1.111282	2.501702	1.193836
MCD	1.119814	1.143253	2.470724	1.279161
OGK	1.101297	1.080718	2.317278	1.220443
S_n	1.057181	1.000702	2.253193	1.107232
	$n = 500$	$p = 10$	$q = 10$	
$RSh\text{-}S_n$	1.061849	1.004919	2.193867	1.101598
MLE	1.055777	1.056073	2.146041	1.073717
Comedian	1.090472	1.054135	2.311923	1.142357
MCD	1.119162	1.106028	2.389408	1.204313
OGK	1.096527	1.051755	2.301928	1.15537
S_n	1.079007	1.06211	2.225079	1.104945

Table B.26: Efficiency table

	Slope	Intercept	Σ_{diag}	$\Sigma_{offdiag}$
	$n = 50$	$p = 4$	$q = 4$	
$RSh\text{-}S_n$	1.172021	1.020086	2.287704	1.181081
MLE	1.156195	1.083698	2.37104	1.12325
Comedian	1.361004	1.196674	2.831552	1.366031
MCD	3.324486	2.382297	5.907785	3.015436
OGK	1.469819	1.161894	3.082958	1.417595
S_n	1.151859	1.046601	2.339094	1.155314
	$n = 50$	$p = 4$	$q = 8$	
$RSh\text{-}S_n$	1.140247	1.099694	2.219859	1.110488
MLE	1.142621	1.085648	2.233066	1.107241
Comedian	1.354765	1.185182	2.654161	1.143647
MCD	2.738309	2.108277	4.648089	2.400629
OGK	1.525824	1.257324	3.159617	1.467343
S_n	1.151209	1.071138	2.188045	1.125286
	$n = 50$	$p = 8$	$q = 4$	
$RSh\text{-}S_n$	1.306255	1.117222	2.64119	1.287287
MLE	1.273597	1.220601	2.467596	1.223631
Comedian	1.504346	1.331331	2.927176	1.494309
MCD	4.288275	3.149337	6.679908	3.628854
OGK	1.643489	1.236376	3.41832	1.680805
S_n	1.277141	1.17973	2.529524	1.29809
	$n = 50$	$p = 10$	$q = 10$	
$RSh\text{-}S_n$	1.409954	1.303259	2.673938	1.310709
MLE	1.321088	1.242969	2.464703	1.245198
Comedian	1.503693	1.387149	3.023365	1.38795
MCD	2.95578	2.484511	4.99337	2.432618
OGK	1.854153	1.623738	3.582416	1.667471
S_n	1.554822	1.323204	2.996569	1.373973

Table B.27: Robustness table-Contaminants from N($\mathbf{2}\sqrt{\boldsymbol{\chi}^2_{\boldsymbol{p+q,0.99}}}, 0.1 \times \mathbf{I}_q$),$(p, q) =$ (10,10) and $\delta = 10\%$

	Slope MSE	Intercept MSE	diagonal element mse	non diagonal elements mse	slope bias	intercept bias	diagonal element bias	non diagonal element bias
		$n = 50$	$\delta = 10\%$	$p = 10$	$q = 10$	diag=0.1		
S_n	0.035379	0.034299	0.624847	0.019899	0.000192	1.87e-3	0.764586	0.027986
MCD	0.052203	0.044448	0.537251	0.023585	-0.00015	0.000292	0.701854	0.035432
Comedian	0.03354	0.030931	0.679281	0.019648	0.000477	5.23e-4	0.800382	2.50e-2
OGK	0.04065	0.037191	0.582545	0.02048	2.08e-4	6.03e-3	0.735208	0.032768
RSh-S_n	0.03277	0.030403	0.639593	0.018907	2.51e-4	0.001972	0.765733	0.023259
		$n = 100$	$\delta = 10\%$	$p = 10$	$q = 10$	diag=0.1		
S_n	0.014004	0.013292	0.762375	0.010382	5.78e-4	7.34e-4	0.860593	0.011466
MCD	0.02302	0.017691	0.608272	0.013155	6.72e-5	0.001166	0.763183	0.01508
Comedian	0.013894	0.012616	0.806761	0.010174	-0.00035	-0.00078	0.886806	1.02e-2
OGK	0.01515	0.013676	0.721858	0.010639	-3.70e-4	1.83e-3	0.836498	0.012231
RSh-S_n	0.013372	0.012622	0.773067	0.010031	0.000441	0.001128	0.867663	1.08e-2
		$n = 200$	$\delta = 10\%$	$p = 10$	$q = 10$	diag=0.1		
S_n	0.00634	0.005954	0.854939	0.005492	1.40e-5	0.000272	0.918432	0.00517
MCD	0.007377	0.006179	0.775626	0.006039	8.47e-5	-0.00035	0.873666	0.004764
Comedian	0.006177	0.005827	0.892463	0.005505	0.000834	-0.00095	0.938882	0.004963
OGK	0.006604	0.006088	0.830516	0.005617	-0.00017	0.002482	0.905209	0.005521
RSh-S_n	0.00613	0.005825	0.858965	0.005326	-0.00024	1.28e-3	0.920977	0.004581
		$n = 500$	$\delta = 10\%$	$p = 10$	$q = 10$	diag=0.1		
S_n	0.002376	0.002208	0.919783	0.002225	-0.00011	-5.20e-5	0.9566	0.001779
MCD	0.002418	0.002309	0.900211	0.002273	1.25e-4	2.41e-4	0.94635	0.001675
Comedian	0.002348	0.002271	0.936253	0.002213	-1.20e-4	-0.00116	0.965302	0.002007
OGK	0.002436	0.002339	0.903601	0.002292	-1.50e-4	3.31e-4	0.948139	0.0024
RSh-S_n	0.002154	0.0021931	0.679281	0.0019648	0.000477	0.000523	0.800382	0.0012496

Table B.28: Robustness table-Contaminants from N($\mathbf{2}\sqrt{\boldsymbol{\chi}^{\mathbf{2}}_{\boldsymbol{p+q,0.99}}}$, $0.1 \times \mathbf{I}_q$), (p, q) = (8,4) and δ = 10%

	Slope MSE	Intercept MSE	diagonal element MSE	non diagonal elements MSE	slope bias	intercept bias	diagonal element bias	non diagonal element bias
		$n = 50$	δ = 10%	$p = 8$	$q = 4$	diag = 0.1		
S_n	0.030071	0.029982	0.710073	0.020074	-0.00024	1.39e-3	0.818126	-0.00602
MCD	0.071754	0.040567	0.452425	0.025876	-0.00251	0.002572	0.632953	-0.00422
Comedian	0.03016	0.0283	0.702836	0.019552	0.000189	1.69e-3	0.814418	-4.97e-3
OGK	0.036603	0.029532	0.601802	0.019494	1.02e-3	7.87e-3	0.746853	0.002305
RSh-S_n	0.021185	0.02856	0.6768	0.019388	1.49e-4	0.000577	0.797973	-0.00343
		$n = 100$	δ = 10%	$p = 8$	$q = 4$	diag = 0.1		
S_n	0.01468	0.011625	0.822838	0.010568	2.83e-4	-1.67e-3	0.895149	0.005459
MCD	0.018414	0.012494	0.632756	0.012816	-6.80e-4	0.00015	0.778974	0.005181
Comedian	0.013838	0.011703	0.814563	0.010544	0.000571	0.001131	0.890538	5.30e-3
OGK	0.014344	0.012224	0.749691	0.011096	5.24e-6	3.00e-4	0.852167	0.009678
RSh-S_n	0.013151	0.011495	0.786627	0.010442	0.000276	0.001099	0.875251	7.94e-3
		$n = 200$	δ = 10%	$p = 8$	$q = 4$	diag = 0.1		
S_n	0.006087	0.005765	0.887469	0.005593	-9.50e-5	0.000937	0.936072	0.003351
MCD	0.006688	0.005708	0.854231	0.005953	-1.50e-4	0.000788	0.889898	0.002215
Comedian	0.006099	0.005683	0.87817	0.00559	0.000193	0.000788	0.930824	0.005072
OGK	0.006438	0.005796	0.833996	0.00575	-0.00015	-0.00057	0.90655	0.00307
RSh-S_n	0.006072	0.005668	0.859239	0.005493	-0.00018	-7.10e-4	0.920615	0.001929
		$n = 500$	δ = 10%	$p = 8$	$q = 4$	diag = 0.1		
S_n	0.002327	0.002167	0.919339	0.00227	0.000489	-4.70e-4	0.956496	0.000819
MCD	0.006688	0.005708	0.804231	0.005953	-1.50e-4	7.88e-4	0.889898	0.002215
Comedian	0.002319	0.002195	0.919471	0.002271	-1.20e-4	0.000267	0.956583	0.000463
OGK	0.002443	0.002253	0.885856	0.002292	4.43e-4	3.71e-5	0.938705	0.002677
RSh-S_n	0.002652	0.002503	0.89919	0.00255	-0.00014	-0.00028	0.94555	0.00143

Table B.29: Robustness table-Contaminants from N($\mathbf{2}\sqrt{\boldsymbol{\chi}^{\mathbf{2}}_{\boldsymbol{p+q,0.99}}}, 0.1 \times \mathbf{I}_q$), (p, q) = (4,8) and δ = 10%

	Slope MSE	Intercept MSE	diagonal element MSE	non diagonal elements MSE	slope bias	intercept bias	diagonal element bias	non diagonal element bias
		$n = 50$	$\delta = 10\%$	$p = 4$	$q = 8$	diag = 0.1		
S_n	0.026368	0.025535	0.898845	0.021827	-0.00054	-2.40e-3	0.924452	-0.00526
MCD	0.051203	0.036295	0.687359	0.029132	-0.00179	-0.00134	0.793941	-0.00536
Comedian	0.026599	0.02533	0.887046	0.02133	-0.00068	3.81e-3	0.917734	-1.95e-3
OGK	0.032623	0.029198	0.783143	0.023087	1.40e-3	5.66e-3	0.858097	-0.00088
RSh-S_n	0.02599	0.024524	0.849369	0.020729	-1.23e-3	0.003604	0.898265	-0.00183
		$n = 100$	$\delta = 10\%$	$p = 4$	$q = 8$	diag = 0.1		
S_n	0.01207	0.011584	0.935365	0.011002	5.20e-4	1.62e-3	0.955791	0.001446
MCD	0.017647	0.013201	0.750804	0.013175	4.76e-4	-0.00092	0.850689	0.002867
Comedian	0.012224	0.011975	0.928054	0.010863	0.000587	0.000461	0.951575	2.23e-3
OGK	0.013814	0.011857	0.845729	0.011493	-8.30e-4	2.74e-3	0.906208	0.004591
RSh-S_n	0.011325	0.011505	0.884337	0.01069	-0.00068	0.001695	0.928422	1.77e-3
		$n = 200$	$\delta = 10\%$	$p = 4$	$q = 8$	diag = 0.1		
S_n	0.005859	0.005726	0.943262	0.005523	6.59e-4	-0.00039	0.965365	0.004226
MCD	0.006482	0.005635	0.870071	0.005925	1.57e-4	0.000395	0.926094	0.004519
Comedian	0.005881	0.005699	0.944197	0.005567	-0.00016	-9.90e-5	0.965948	0.004217
OGK	0.006364	0.005658	0.892352	0.005751	-0.00016	0.000539	0.938439	0.004844
RSh-S_n	0.005851	0.005604	0.91682	0.00553	0.000377	1.42e-4	0.951738	0.003961
		$n = 500$	$\delta = 10\%$	$p = 4$	$q = 8$	diag = 0.1		
S_n	0.002333	0.002207	0.957668	0.002265	0.000236	3.51e-4	0.976269	0.00084
MCD	0.002394	0.002244	0.924019	0.002312	-4.50e-4	5.88e-4	0.958785	0.00059
Comedian	0.002303	0.002211	0.954446	0.002251	-4.00e-4	0.000591	0.974639	0.001185
OGK	0.002394	0.002237	0.922555	0.002337	6.94e-4	1.02e-3	0.958018	0.000802
RSh-S_n	0.00263	0.00248	0.94284	0.00255	0.00042	0.00029	0.96843	0.00039

Table B.30: Robustness table-Contaminants from N($\mathbf{2}\sqrt{\boldsymbol{\chi}^{\mathbf{2}}_{\boldsymbol{p+q,0.99}}}$, $0.1 \times \mathbf{I}_q$),$(p, q) =$ (4,4) and $\delta = 10\%$

	Slope MSE	Intercept MSE	diagonal element MSE	non diagonal elements MSE	slope bias	intercept bias	diagonal element bias	non diagonal element bias
		$n = 50$	$\delta = 10\%$	$p = 4$	$q = 4$	diag = 0.1		
S_n	0.027062	0.024811	0.869399	0.021403	0.003809	1.82e-3	0.908377	-0.00498
MCD	0.056792	0.029701	0.605927	0.02885	-0.00122	-7.10e-5	0.740357	-0.00533
Comedian	0.027862	0.024672	0.852523	0.021707	0.002757	2.44e-4	0.898504	-3.92e-3
OGK	0.032135	0.025488	0.75953	0.022854	9.46e-4	3.93e-3	0.842867	-0.00075
RSh-S_n	0.027537	0.02368	0.803693	0.01998	5.57e-4	0.005104	0.87283	-0.00295
		$n = 100$	$\delta = 10\%$	$p = 4$	$q = 4$	diag = 0.1		
S_n	0.012804	0.011781	0.903699	0.011783	-1.70e-4	-1.53e-3	0.938678	0.006895
MCD	0.015236	0.01145	0.750362	0.01302	-1.77e-3	0.000362	0.850676	0.008832
Comedian	0.012505	0.011895	0.902279	0.011021	-0.00057	-0.00269	0.93753	6.86e-3
OGK	0.014642	0.013542	0.847704	0.012663	8.48e-4	7.61e-4	0.906497	0.008697
RSh-S_n	0.01248	0.0113	0.864418	0.010806	0.000346	0.002428	0.917793	6.03e-3
		$n = 200$	$\delta = 10\%$	$p = 4$	$q = 4$	diag = 0.1		
S_n	0.005973	0.005682	0.926016	0.005729	-8.00e-5	0.000141	0.956202	0.003477
MCD	0.006437	0.00573	0.846856	0.006063	2.98e-4	2.51e-6	0.913761	0.003293
Comedian	0.005955	0.005523	0.923031	0.005722	-5.90e-5	-9.10e-6	0.954904	0.00261
OGK	0.013524	0.011409	0.81934	0.011265	0.001151	0.003127	0.891392	0.00972
RSh-S_n	0.00591	0.005506	0.898654	0.005729	0.000355	-4.40e-4	0.941665	0.002543
		$n = 500$	$\delta = 10\%$	$p = 4$	$q = 4$	diag = 0.1		
S_n	0.002309	0.002166	0.93471	0.002198	4.65e-5	-2.50e-4	0.964442	0.001258
MCD	0.002404	0.00217	0.905404	0.002333	-1.90e-5	5.97e-4	0.949063	0.002317
Comedian	0.00231	0.002195	0.933463	0.002302	-1.90e-4	0.001192	0.963769	0.001277
OGK	0.002426	0.002143	0.893785	0.002358	-4.90e-5	2.73e-4	0.94283	0.000895
RSh-S_n	0.00262	0.00239	0.92092	0.00252	0.00033	0.00052	0.95693	0.00046

Table B.31: Robustness table-Contaminants from N($\mathbf{2}\sqrt{\boldsymbol{\chi}^{\mathbf{2}}_{\boldsymbol{p+q},\mathbf{0.99}}}$, $\mathbf{I}_q$),(p, q) = (4,4) and δ = 10%

	Slope MSE	Intercept MSE	diagonal element MSE	non diagonal elements MSE	slope bias	intercept bias	diagonal element bias	non diagonal element bias
		$n = 50$	$\delta = 10\%$	$p = 4$	$q = 4$			
S_n	0.027104	0.026141	0.871365	0.022099	-0.00017	-1.70e-3	0.909667	-0.00548
MCD	0.054961	0.030504	0.820277	0.028164	-2.90e-5	-2.78e-3	0.836767	-0.001
Comedian	0.032238	0.026643	0.861882	0.022298	0.002052	3.70e-3	0.845044	-1.03e-3
OGK	0.032483	0.025911	0.846421	0.022184	9.58e-4	4.30e-3	0.836388	0.001282
RSh-S_n	0.026918	0.024661	0.826234	0.021195	4.00e-4	0.001171	0.883227	-0.00357
		$n = 100$	$\delta = 10\%$	$p = 4$	$q = 4$			
S_n	0.01262	0.011744	0.909175	0.011397	6.22e-4	5.36e-4	0.941562	0.006297
MCD	0.015602	0.01156	0.866	0.01277	-2.80e-5	-0.00137	0.859509	0.005811
Comedian	0.013078	0.011632	0.821849	0.011563	0.000731	0.001822	0.893043	8.89e-3
OGK	0.013281	0.011338	0.825212	0.011239	1.36e-3	-7.20e-4	0.895182	0.007572
RSh-S_n	0.012423	0.011437	0.814166	0.010954	0.000505	0.00206	0.817553	5.99e-3
		$n = 200$	$\delta = 10\%$	$p = 4$	$q = 4$			
S_n	0.006075	0.00561	0.927595	0.005794	-2.20e-4	-0.00045	0.956908	0.002803
MCD	0.006341	0.005631	0.86313	0.005914	1.24e-4	1.23e-3	0.922112	0.003357
Comedian	0.006279	0.005663	0.87052	0.005857	-3.30e-4	8.01e-4	0.926773	0.004161
OGK	0.006323	0.00565	0.866304	0.006013	-1.70e-5	0.000297	0.924135	0.005033
RSh-S_n	0.006041	0.005542	0.813558	0.00544	-0.00019	3.30e-4	0.919649	0.002496
		$n = 500$	$\delta = 10\%$	$p = 4$	$q = 4$			
S_n	0.002324	0.002114	0.93284	0.002287	7.79e-5	-3.00e-5	0.963423	0.002273
MCD	0.002454	0.002223	0.904046	0.002375	-1.80e-5	-5.80e-4	0.948241	0.002253
Comedian	0.002425	0.002141	0.92591	0.002396	-4.80e-4	0.001091	0.942233	0.002711
OGK	0.002397	0.002234	0.922206	0.002297	-4.10e-5	6.97e-4	0.957776	0.00141
RSh-S_n	0.002267	0.00207	0.904538	0.002251	0.000522	0.000228	0.923882	0.00139

Table B.32: Robustness table-Contaminants from N($\mathbf{2}\sqrt{\boldsymbol{\chi}^{\mathbf{2}}_{\boldsymbol{p+q},\mathbf{0.99}}}, \mathbf{I}_q$), (p, q) = (4,8) and δ = 10%

	Slope MSE	Intercept MSE	diagonal element mse	non diagonal elements mse	slope bias	intercept bias	diagonal element bias	non diagonal element bias
		$n = 50$	δ = 10%	$p = 4$	$q = 8$			
S_n	0.026736	0.024352	0.898834	0.021601	0.000235	6.76e-5	0.925397	-0.00177
MCD	0.052629	0.036653	0.87425	0.028786	3.13e-3	-4.60e-4	0.904536	-0.00185
Comedian	0.026819	0.02595	0.895646	0.021803	-7.60e-5	-1.50e-4	0.922931	-2.96e-3
OGK	0.032762	0.027617	0.871278	0.022477	1.77e-3	3.45e-3	0.905827	-0.00057
RSh-S_n	0.026652	0.02444	0.852299	0.020924	-7.90e-4	0.004284	0.898662	-0.00315
		$n = 100$	δ = 10%	$p = 4$	$q = 8$			
S_n	0.01207	0.011492	0.93025	0.010913	9.31e-4	-2.04e-3	0.953118	0.00125
MCD	0.017246	0.012694	0.895446	0.013573	2.71e-4	-0.00177	0.91355	0.00309
Comedian	0.013746	0.012192	0.897849	0.011448	0.000796	0.002605	0.912437	3.12e-3
OGK	0.013819	0.012194	0.890424	0.011384	5.36e-4	2.83e-3	0.91312	0.004237
RSh-S_n	0.012228	0.011432	0.885593	0.010725	4.20e-6	-0.00143	0.909856	1.53e-3
		$n = 200$	δ = 10%	$p = 4$	$q = 8$			
S_n	0.005922	0.005727	0.951432	0.00569	2.68e-5	0.000311	0.969582	0.004976
MCD	0.006504	0.005943	0.873553	0.005972	-2.10e-4	2.88e-4	0.927869	0.004817
Comedian	0.006353	0.005859	0.886311	0.005809	-3.90e-4	1.94e-3	0.935359	0.005867
OGK	0.006369	0.005742	0.835858	0.005715	-6.10e-4	0.001927	0.907761	0.004887
RSh-S_n	0.00598	0.005617	0.819595	0.005549	-0.00052	1.08e-3	0.953175	0.004802
		$n = 500$	δ = 10%	$p = 4$	$q = 8$			
S_n	0.002315	0.002237	0.955434	0.00224	-4.50e-4	6.80e-4	0.975091	0.000792
MCD	0.002346	0.002193	0.94329	0.002277	-9.80e-5	7.93e-4	0.98386	0.000411
Comedian	0.002286	0.002301	0.955494	0.002256	3.59e-5	-0.00134	0.975122	0.00014
OGK	0.002397	0.002234	0.942206	0.002297	-4.10e-5	6.97e-4	0.97776	0.00134
RSh-S_n	0.002227	0.00219	0.938934	0.002256	0.000349	0.001011	0.96662	0.000289

Table B.33: Robustness table-Contaminants from N($\mathbf{2}\sqrt{\boldsymbol{\chi}^{\mathbf{2}}_{\boldsymbol{p+q},\mathbf{0.99}}}$, $\mathbf{I}_q$),(p, q) = (8,4) and δ = 10%

	Slope MSE	Intercept MSE	diagonal element mse	non diagonal elements mse	slope bias	intercept bias	diagonal element bias	non diagonal element bias
		$n = 50$	$\delta = 10\%$	$p = 8$	$q = 4$			
S_n	0.040228	0.027811	0.709027	0.019555	0.000525	2.45e-3	0.817944	-0.00341
MCD	0.071942	0.042172	0.766495	0.026127	2.23e-3	-1.87e-3	0.642774	-0.00385
Comedian	0.036868	0.029341	0.602844	0.019968	-2.20e-4	2.41e-3	0.74797	1.45e-3
OGK	0.037403	0.030711	0.606282	0.019967	5.17e-4	6.06e-3	0.751147	-0.00039
RSh-S_n	0.030355	0.027533	0.66581	0.019232	-1.20e-4	0.001425	0.791538	-0.00179
		$n = 100$	$\delta = 10\%$	$p = 8$	$q = 4$			
S_n	0.014977	0.012155	0.819943	0.010506	5.67e-4	-1.40e-3	0.894061	0.006268
MCD	0.018887	0.012857	0.742024	0.013133	1.03e-3	-0.0009	0.884671	0.008915
Comedian	0.014252	0.012338	0.749359	0.010878	-0.00105	0.000931	0.892171	9.80e-3
OGK	0.014383	0.012097	0.749174	0.010909	6.13e-4	-2.00e-4	0.89233	0.009278
RSh-S_n	0.013304	0.011938	0.791517	0.010294	2.01e-4	0.003988	0.877971	4.62e-3
		$n = 200$	$\delta = 10\%$	$p = 8$	$q = 4$			
S_n	0.006925	0.005856	0.878332	0.005491	-7.90e-5	0.000477	0.931262	0.003079
MCD	0.006612	0.005592	0.809249	0.006022	4.23e-4	-2.70e-3	0.912674	0.003315
Comedian	0.006405	0.005996	0.827204	0.005495	3.07e-4	1.76e-3	0.903246	0.003814
OGK	0.006369	0.005942	0.835858	0.005715	-6.10e-4	0.001927	0.907761	0.004887
RSh-S_n	0.006235	0.00583	0.807884	0.005459	-0.00031	1.02e-4	0.900022	0.003575
		$n = 500$	$\delta = 10\%$	$p = 8$	$q = 4$			
S_n	0.00234	0.00226	0.916931	0.002298	-4.70e-4	-5.70e-4	0.95514	0.001272
MCD	0.002443	0.002189	0.886638	0.002319	6.42e-5	-4.80e-4	0.939115	0.000909
Comedian	0.002436	0.002232	0.88465	0.0023	-2.40e-4	-0.00032	0.937966	0.001095
OGK	0.002441	0.002285	0.901099	0.002305	-1.20e-4	7.61e-4	0.946845	0.002174
RSh-S_n	0.002157	0.002154	0.877109	0.00224	2.94e-5	-0.00028	0.944794	0.000946

Table B.34: Robustness table-Contaminants from N($\mathbf{2}\sqrt{\boldsymbol{\chi}^{\mathbf{2}}_{\boldsymbol{p+q},\mathbf{0.99}}}, \mathbf{I}_q$), (p, q) = (10,10) and δ = 10%

	Slope MSE	Intercept MSE	diagonal element MSE	non diagonal elements MSE	slope bias	intercept bias	diagonal element bias	non diagonal element bias
		$n = 50$	$\delta = 10\%$	$p = 10$	$q = 10$			
S_n	0.035186	0.033186	0.628093	0.019893	-0.00029	1.92e-3	0.766835	0.026861
MCD	0.05331	0.045228	0.648384	0.024382	-4.90e-4	9.81e-4	0.798176	0.036003
Comedian	0.039939	0.030156	0.674917	0.019562	-2.60e-4	-4.90e-4	0.797826	2.40e-2
OGK	0.041184	0.036058	0.681502	0.020469	1.89e-4	3.19e-3	0.794167	0.030117
RSh-S_n	0.032333	0.029723	0.637209	0.018516	-5.00e-4	0.001362	0.774981	0.022447
		$n = 100$	$\delta = 10\%$	$p = 10$	$q = 10$			
S_n	0.013988	0.013156	0.796434	0.010543	-2.20e-5	9.67e-4	0.863244	0.011695
MCD	0.023567	0.017868	0.802532	0.013419	8.36e-4	0.000879	0.886948	0.016278
Comedian	0.014097	0.012964	0.818361	0.010321	-0.00027	2.16e-5	0.893156	1.11e-2
OGK	0.015392	0.01365	0.790919	0.010768	-1.90e-4	4.50e-3	0.895836	0.01286
RSh-S_n	0.013361	0.012747	0.771019	0.010076	2.26e-4	0.002149	0.866525	9.51e-3
		$n = 200$	$\delta = 10\%$	$p = 10$	$q = 10$			
S_n	0.006289	0.005989	0.875448	0.005516	3.12e-4	0.002718	0.919868	0.006935
MCD	0.007359	0.00635	0.873152	0.006109	3.28e-4	2.85e-5	0.872256	0.005688
Comedian	0.006457	0.005785	0.886198	0.005354	-8.00e-5	1.56e-3	0.935541	0.006612
OGK	0.006611	0.006031	0.892817	0.005625	1.22e-4	0.001466	0.90463	0.005839
RSh-S_n	0.006175	0.005768	0.860194	0.005271	-0.0003	1.43e-3	0.921546	0.005657
		$n = 500$	$\delta = 10\%$	$p = 10$	$q = 10$			
S_n	0.002381	0.002312	0.91882	0.002253	-3.20e-4	4.00e-4	0.95622	0.001973
MCD	0.002433	0.002329	0.902059	0.002315	-4.00e-4	-6.70e-4	0.947311	0.002049
Comedian	0.00244	0.002267	0.935814	0.002231	-8.50e-5	0.000351	0.965056	0.001924
OGK	0.002441	0.002385	0.910099	0.002305	-1.20e-4	7.61e-4	0.946845	0.002174
RSh-S_n	0.002338	0.002314	0.901812	0.00223	-1.60e-5	0.001107	0.957193	0.001919

Table B.35: $\mathbf{y}$ - affine equivariance table

	slope	intercept	slope	intercept	slope	intercept	slope	intercept
δ = 0%	p = 4,q = 4		p = 6,q = 10		p = 10,q = 6		p = 10,q = 10	
n = 50	3.08e-5	8.80e-5	1.67e-5	6.62e-5	1.60e-5	9.53e-5	1.19e-5	8.59e-5
n = 100	7.67e-6	2.21e-5	3.68e-6	1.48e-5	2.63e-6	1.92e-5	2.61e-6	1.69e-5
n = 200	1.71e-6	5.71e-6	7.03e-7	3.42e-6	5.37e-7	4.13e-6	4.25e-7	3.60e-6
n = 500	3.24e-7	1.13e-6	9.88e-8	5.39e-7	1.17e-7	9.51e-7	6.41e-8	5.16e-7
δ = 10%	p = 4,q = 4		p = 6,q = 10		p = 10,q = 6		p = 10,q = 10	
n = 50	4.07e-5	9.28e-5	2.21e-5	6.86e-5	1.30e-5	1.15e-4	1.21e-5	8.00e-5
n = 100	1.16e-5	3.22e-5	3.64e-6	1.23e-5	3.10e-6	2.13e-5	2.29e-6	1.59e-5
n = 200	4.63e-6	1.69e-5	8.22e-7	3.53e-6	6.20e-7	4.77e-6	4.89e-7	3.59e-6
n = 500	6.00e-6	6.27e-6	1.40e-7	6.60e-7	2.14e-7	6.08e-6	9.14e-8	6.73e-7
δ = 20%	p = 4,q = 4		p = 6,q = 10		p = 10,q = 6		p = 10,q = 10	
n = 50	0.000104	0.002708	4.94e-6	1.46e-5	5.47e-6	0.00021	4.03e-6	2.70e-5
n = 100	4.96e-5	0.002438	1.31e-6	4.29e-6	2.74e-6	0.00026	9.84e-7	5.26e-6
n = 200	6.42e-5	0.003637	3.18e-7	1.15e-6	1.28e-6	1.93e-4	1.55e-7	1.21e-6
n = 500	8.49e-5	0.004147	6.71e-8	2.38e-7	1.88e-6	9.19e-4	3.54e-8	4.32e-6

Table B.36: $\mathbf{x}$ - affine equivariance table

	slope	intercept	slope	intercept	slope	intercept	slope	intercept
$\delta = 0\%$	$p = 4$,$q = 4$		$p = 6$,$q = 10$		$p = 10$,$q = 6$		$p = 10$,$q = 10$	
$n = 50$	1.33e-1	4.11e-4	1.45e-1	1.94e-4	1.75e-1	5.19e-4	1.29e-1	2.91e-4
$n = 100$	1.06e-1	8.88e-5	1.01e-1	5.39e-5	1.14e-1	1.02e-4	9.18e-2	8.18e-5
$n = 200$	6.33e-2	2.79e-5	5.79e-2	1.40e-5	8.21e-2	2.52e-5	5.60e-2	1.85e-5
$n = 500$	6.53e-2	4.73e-6	4.32e-2	2.27e-6	3.37e-2	3.25e-6	3.72e-2	2.63e-6
$\delta = 10\%$	$p = 4$,$q = 4$		$p = 6$,$q = 10$		$p = 10$,$q = 6$		$p = 10$,$q = 10$	
$n = 50$	2.10e-1	4.25e-4	1.43e-1	2.22e-4	2.04e-1	5.88e-4	1.42e-1	2.93e-4
$n = 100$	8.75e-2	3.93e-5	9.87e-2	6.58e-5	1.42e-1	1.99e-4	1.05e-1	1.03e-4
$n = 200$	1.06e-1	4.51e-5	7.96e-2	1.90e-5	8.38e-2	5.43e-5	5.19e-2	3.13e-5
$n = 500$	9.93e-2	1.04e+0	2.78e-2	3.18e-6	5.63e-2	9.95e-6	4.14e-2	4.98e-6
$\delta = 20\%$	$p = 4$,$q = 4$		$p = 6$,$q = 10$		$p = 10$,$q = 6$		$p = 10$,$q = 10$	
$n = 50$	0.266032	0.000148	1.57e-1	1.23e-4	2.25e-1	0.00024	1.85e-1	1.57e-4
$n = 100$	1.46e-1	6.29e-5	1.15e-1	2.84e-5	1.59e-1	9.06e-5	1.07e-1	5.22e-5
$n = 200$	9.45e-2	9.08e-6	5.58e-2	9.44e-6	1.13e-1	2.91e-5	5.48e-2	1.37e-5
$n = 500$	7.71e-2	1.28e-6	5.89e-2	1.17e-6	5.67e-2	2.28e-6	2.79e-2	2.34e-6

UNIVERSITY OF CALICUT

CERTIFICATE ON PLAGIARISM CHECK

<table>
<tr><td>1.</td><td>Name of the Research Scholar</td><td colspan="2">Lakshmi R</td></tr>
<tr><td>2.</td><td>Title of thesis / dissertation</td><td colspan="2">A study on Robust Regression Methods</td></tr>
<tr><td>3.</td><td>Name of the Supervisor</td><td colspan="2">Dr. Sajesh T A</td></tr>
<tr><td>4.</td><td>Department/Institution</td><td colspan="2">Department of Statistics,
St. Thomas' College (Autonomous), Thrissur</td></tr>
<tr><td rowspan="4">5.</td><td rowspan="2"></td><td>Non-core</td><td>Core</td></tr>
<tr><td>Introduction/ Theoretical overview/Review of literature/ Materials & Methods/ Methodology</td><td>Analysis/Result/Discussion/ Summary/Conclusion/ Recommendations</td></tr>
<tr><td>Similar content (%) identified</td><td>8%</td><td>3%</td></tr>
<tr><td>Acceptable maximum limit (%)</td><td>10</td><td>10</td></tr>
<tr><td>6.</td><td>Software used</td><td colspan="2">iThenticate</td></tr>
<tr><td>7.</td><td>Date of verification</td><td colspan="2">Nov 23, 2024</td></tr>
</table>

**Report on plagiarism check, specifying included/excluded items with % of similarity to be attached.*

Checked by (with name , designation & signature)

Dr. Nasirudheen. T
Assistant Librarian
University of Calicut, Kerala.

Name and signature of the Researcher LAKSHMI R.

Name and signature of the Supervisor.

Dr. SAJESH. T.A, PhD
HEAD OF THE DEPARTMENT
DEPARTMENT OF STATISTICS
ST. THOMAS COLLEGE (AUTONOMOUS)
THRISSUR, KERALA - 680 001

The Doctoral Committee* has verified the report on plagiarism check with the contents of the thesis, as summarized above and appropriate measures have been taken to ensure originality of the Research accomplished herein.

Name & Signature of the HoD/HoI (Chairperson of the Doctoral Committee)

Dr. Martin K. A.
Principal-in-Charge
St. Thomas College (Autonomous)
Thrissur - 680 001

**In case of languages like Malayalam, Tamil etc..on which no software is available for plagiarism check, a manual check shall be made by the Doctoral Committee, for which an additional certificate has to be attached.*