UNIVERSITI PUTRA MALAYSIA

ROBUST DIAGNOSTIC AND ROBUST ESTIMATION METHODS FOR FIXED EFFECT PANEL DATA MODEL IN PRESENCE OF HIGH LEVERAGE POINTS AND MULTICOLLINEARITY

SHELAN SAIED ISMAEEL

FS 2018 1

ROBUST DIAGNOSTIC AND ROBUST ESTIMATION METHODS FOR FIXED EFFECT PANEL DATA MODEL IN PRESENCE OF HIGH LEVERAGE POINTS AND MULTICOLLINEARITY

By

SHELAN SAIED ISMAEEL

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfillment of the Requirements for the Degree of Doctor of Philosophy

December 2017

DEDICATION

To

Prophet of Almighty Allah, Prophet Muhammad (PBUH).

My father and mother

For their encouragement

And

To my husband (Azad) and my kids (Warezh & Aroj)

For their great patience

My brothers (Ramazan and Mohammed) and my sisters(Jiyan and Hizert), you

have been my

inspiration and my soul mates

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfillment of the requirements for the degree of Doctor of Philosophy

ROBUST DIAGNOSTIC AND ROBUST ESTIMATION METHODS FOR FIXED EFFECT PANEL DATA MODEL IN PRESENCE OF HIGH LEVERAGE POINTS AND MULTICOLLINEARITY

By

SHELAN SAIED ISMAEEL

December 2017

Chairman : Professor Habshah Midi, PhD
Faculty : Science

The Diagnostic Robust Generalized Potential based on Minimum Volume Ellipsoid (MVE) is proposed in linear regression to detect high leverage points (HLPs). However, it takes a very long computational running time and also has small rate of swamping and masking effects. Hence the Improvised Diagnostic Robust Generalized Potential based on Index Set Equality (IDRGP (ISE)) is proposed to linear and fixed effect panel data model. The results indicate that IDRGP(ISE) successfully identify high leverage points with the reduction in the rate of swamping and masking effects and has less computational running time.

To date no research has been done to identify HLPs for panel data. Hence, to close the gap in the literature we propose Within Group Improvised Diagnostic Robust Generalized Potential (WIDRGP). It is very successful in detecting HLPs and relatively fast to compute.

The Generalized M-estimator (GM6) is the widely used method to overcome the problem of HLPs for multiple linear regression model. However, this method is less efficient since it is based on Robust Mahalanobis Distance RMD- MVE as an initial π–weight function. Its efficiency decreases as the number of good leverage points increases. Hence, the Generalized M-estimator (GM) based on Fast Improvised Generalized Studentized Residuals (FIMGT), denoted as (GM-FIMGT) is developed. The results show that the GM-FIMGT is highly efficient and relatively fast. A robust Within Group GM estimator based on FIMGT estimator (WGM-FIMGT) for fixed effect panel data model is proposed. The findings indicate that the WGM-FIMGT is very efficient compared to the existing estimators.

Thus far, no research has been done on the detection of multicollinearity for fixed effect panel data models in the presence of HLPs. Hence, Robust Variance Inflation Factor based on GM-FIMGT (RVIF(GM-FIMGT)) is formulated. The results of the study show that it is very effective in detecting multicollinearity in the presence of HLPs.

The Jackknife ridge regression is one of the commonly used method to remedy the problem of multicollinearity. Nonetheless, it is very sensitive to outliers and HLPs. Hence Robust Jackknife ridge regression based on FIMGT (RJFIMGT) is developed to rectify the combined problem of multicollinearity and high leverage points. The results of the study indicate that the RJFIMGT is the most efficient method when multicollinearity problem come together with the presence of HLPs.

Still no research has been done on the parameter estimation of fixed effect panel data model in the presence of multicollinearity and HLPs. Thus the within Group Robust Jackknife ridge regression based on FIMGT (WRJFIMGT) is developed to close the gap in the literature. The findings signify that WRJFIMGT provides the best estimates when multicollinearity and HLPs are present in a data set

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Doktor Falsafah

KAEDAH DIAGNOSTIK TEGUH DAN ANGGARAN TEGUH UNTUK MODEL LINEAR BERGANDA DAN MODEL DATA PANEL KESAN TETAP DENGAN KEHADIRAN TITIK TUASAN TINGGI DAN MULTIKOLINEARAN

Oleh

SHELAN SAIED ISMAEEL

Disember 2017

Pengerusi : Profesor Habshah Midi, PhD
Fakulti : Sains

Kaedah teguh berdiagnostik potensi teritlak (DRGP) berasaskan isipadu minimum ellipsoid (MVE) telah dicadangkan untuk mengesan titik tuasan tinggi (HLPs). Walau bagaimanapun, ianya mengambil masa yang lama dari segi masa pengiraan dan juga mempunyai kadar yang kecil bagi kesan swamping dan masking. Oleh itu kaedah teguh berdiagnostik potensi teritlak tertambahbaik berasaskan indeks set kesaksamaan (IDRGP(ISE)) dicadangkan untuk model linear dan model data panel kesan tetap. Keputusan menunjukkan bahawa (IDRGP(ISE)) berjaya mengesan titik tuasan tinggi dengan penurunan kadar kesan swamping dan masking dan juga masa pengiraan dapat dikurangkan.

Setakat ini, tidak ada penyelidikan yang telah dijalankan untuk mengesan HLPs bagi data panel. Oleh itu, untuk menutup jurang kesusasteraan ini, kami mencadangkan kaedah teguh berdiagnostik potensi teritlak tertambahbaik dalam kumpulan (WIDRGP). Kaedah ini sangat berjaya dalam mengesan HLPs dan pengiraannya pantas.

Kaedah penganggar M-teritlak GM6 digunakan secara meluas untuk mengatasi masalah HLPs bagi model linear regresi berganda. Walau bagaimanapun, kaedah ini kurang cekap kerana ianya berasaskan RMD-MVE sebagai fungsi pemberat -π permulaan. Kecekapannya berkurangan apabila bilangan titik tuasan yang baik meningkat. Oleh itu penganggar GM berasaskan reja teritlak student tertambahbaik (FIMGT) dinamakan (GM-FIMGT) di bangunakan. Keputusan menunjukkan bahawa GM-FIMGT adalah sangat cekap dan pantas secara relatif.

Penganggar teguh GM dalam kumpulan berasaskan penganggar FIMGT (WGM-FIMGT) dicadangkan bagi model data panel kesan tetap. Hasil kajian menunjukkan bahawa WGM-FIMGT adalah sangat cekap berbanding dengan kaedah penganggar yang sedia ada.

Setakat ini, tidak ada kajian telah dijalankan untuk mengesan multikolinearan bagi model data panel kesan tetap dengan kehadiran titik tuasan tinggi. Oleh itu, Faktor Inflasi Varians Teguh berasaskan WGM-FIMGT (RVIF(WGM-FIMGT)) telah diformulasikan. Keputusan menunjukkan bahawa kaedah ini sangat berkesan untuk mengenalpasti multikolinearan dengan kehadiran HLPs.

Kaedah regresi Jackknife Ridge adalah salah satu kaedah yang biasa digunakan untuk pemulihan masalah multikolinearan. Walau bagaimanapun, ianya sangat sensitif terhadap titik terpencil dan HLPs. Oleh itu, regresi teguh Jackknife Ridge berasaskan FIMGT (RJFIMGT) telah dibangunkan untuk menyelesaikan masalah gabungan multikolinearan dengan HLPs. Keputusan kajian menunjukkan bahawa RJFIMGT adalah kaedah yang sangat berkesan bagi masalah multikolinearan bersama dengan kehadiran HLPs.

Masih tiada kajian telah dijalankan untuk menganggar parameter bagi model data panel kesan tetap dengan kehadiran multikolinearan dan HLPs. Oleh itu, regresi Jackknife ridge teguh dalam kumpulan berasaskan FIMGT (WRJFIMGT) telah dibangunkan untuk menutup jurang kesusasteraan. Hasil kajian menunjukkan WRJFIMGT adalah penganggar terbaik apabila kehadiran multikolinearan dan HLPs dalam set data.

ACKNOWLEDGEMENTS

First of all, I would like to thank Allah for the strength and energy he has given to me. This task could not have been completed without his grace and mercy. To him I owe everything.

The person to whom I would like to express my gratitude is my supervisor Prof. Dr. Habshah Midi. Your wonderful guidance over the years made this thesis possible and I am be forever grateful. I am deeply honored to have the opportunity to complete my degree under her supervision. Thanks for being my supervisor.

I am grateful to my committee members Associate Prof. Dr. Jayanthi A/P Arasan, Dr. Mohd Shafie Bin Mustafa for all their support and guidance provided.

I would like to thank all the students, professors, and the staff members at Department of Mathematics, Universiti Putra Malaysia for all of their help, guidance, and advice.

I would also like to thank my mother and father for their prayers and their never-ending unconditional support.

My special thanks go to my beloved husband, Azad, for his support and encouragement. I would like to extend my gratitude to our two sons, Warezh and Aroj; they are a great blessing to us.

Finally, I am thankful for the financial support during my studies at the Universiti Putra Malaysia, which was kindly provides by Kurdistan Regional Government (KRG). I do not know what I could do without these supports.

I certify that a Thesis Examination Committee has met on 21 December 2017 to conduct the final examination of Shelan Saied Ismaeel on his thesis entitled "Robust Diagnostic and Robust Estimation Methods for Fixed Effect Panel Data Model in Presence of High Leverage Points and Multicollinearity" in accordance with the Universities and University Colleges Act 1971 and the Constitution of the Universiti Putra Malaysia [P.U.(A) 106] 15 March 1998. The Committee recommends that the student be awarded the Doctor of Philosophy.

Members of the Thesis Examination Committee were as follows:

Leong Wah June, PhD
Associate Professor
Faculty of Science
Universiti Putra Malaysia
(Chairman)

Mohd Rizam bin Abu Bakar, PhD
Associate Professor
Faculty of Science
Universiti Putra Malaysia
(Internal Examiner)

Abdul Ghapor bin Hussin, PhD
Professor
Universiti Pertahanan Nasional Malaysia
Malaysia
(External Examiner)

Md. Nurul Haque Mollah, PhD
Professor
University of Rajshahi
Bangladesh
(External Examiner)

NOR AINI AB. SHUKOR, PhD
Professor and Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 29 January 2018

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfillment of the requirement for the degree of Doctor of Philosophy. The members of the Supervisory Committee were as follows:

Habshah Bt Midi, PhD
Professor
Faculty of Science
Universiti Putra Malaysia
(Chairman)

Mohd Shafie Bin Mustafa, PhD
Faculty of Science
Universiti Putra Malaysia
(Member)

Jayanthi A/P Arasan, PhD
Associate Professor
Faculty of Science
Universiti Putra Malaysia
(Member)

ROBIAH BINTI YUNUS, PhD
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

Declaration by graduate student

I hereby confirm that:

- this thesis is my original work;
- quotations, illustrations and citations have been duly referenced;
- this thesis has not been submitted previously or concurrently for any other degree at any institutions;
- intellectual property from the thesis and copyright of thesis are fully-owned by Universiti Putra Malaysia, as according to the Universiti Putra Malaysia (Research) Rules 2012;
- written permission must be obtained from supervisor and the office of Deputy Vice-Chancellor (Research and innovation) before thesis is published (in the form of written, printed or in electronic form) including books, journals, modules, proceedings, popular writings, seminar papers, manuscripts, posters, reports, lecture notes, learning modules or any other materials as stated in the Universiti Putra Malaysia (Research) Rules 2012;
- there is no plagiarism or data falsification/fabrication in the thesis, and scholarly integrity is upheld as according to the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) and the Universiti Putra Malaysia (Research) Rules 2012. The thesis has undergone plagiarism detection software

Signature: ____________________ Date: ________________

Name and Matric No: Shelan Saied Ismaeel, GS42235

Declaration by Members of Supervisory Committee

This is to confirm that:

- the research conducted and the writing of this thesis was under our supervision;
- supervision responsibilities as stated in the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) were adhered to.

Signature: ______________________
Name of Chairman of Supervisory Committee: Professor Dr. Habshah Bt Midi

Signature: ______________________
Name of Member of Supervisory Committee: Dr. Mohd Shafie Bin Mustafa

Signature: ______________________
Name of Member of Supervisory Committee: Associate Professor Dr. Jayanthi A/P Arasan

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

BLUE	Best Linear Unbiased Estimators
CVIF	Classical Variance Inflation Factor
CIO	Collinearity-Influential Observations
DRGP(MVE)	Diagnostic Robust Generalized Potential based on Minimum Volume Ellipsoid
DRGP(ISE)	Diagnostic Robust Generalized Potential based on Index Set Equality
FGM	Fast Generalized M-estimator
GM	Generalized M-estimator
GM-FIMGT	Fast Improvised Diagnostic Robust Generalized Potential
HLP	High Leverage Point
HB	Hadi's Potential
HLCIO	High Leverage Collinearity-Influential Observations
ISE	Index Set Equality
IDRGP(MVE)	Improvised Diagnostic Robust Generalized Potential based on Minimum Volume Ellipsoid
IDRGP(ISE)	Improvised Diagnostic Robust Generalized Potential based on Index Set Equality
RMD	Robust Mahalanobis Distance
LP	Leverage Point
LTS	Least Trimmed Squares
MAD	Median Absolute Deviation
MCD	Minimum Covariance Determinant
MD	Mahalanobis Distance
MM	Modify of M-estimator
MSE	Mean Square Error
MVE	Minimum Volume Ellipsoid
OLS	Ordinary Least Squares
OLSc	Ordinary Least Squares for clean data
OLScont	Ordinary Least Squares for contamination data
RR^2(MM)	Robust coefficient determinations (R^2) based on MM estimator

RMD-MVE	Robust Mahalanobis Distance based on Minimum Volume Ellipsoid
RVIF(GM-FIMGT)	Robust Variance Inflation Factor based on Fast Improvised Diagnostic Robust Generalized Potential
RJFIMGT	Robust Jack-knife Ridge Regression based on GM-FIMGT estimator
RJMM	Robust Jack-knife Ridge Regression based on MM-estimator
RJGM2	Robust Jack-knife Ridge Regression based on GM2-estimator
SVR	Support Vector Regression
SRM	Structural Risk Minimization
WHB	Hadi's Potential in panel data
WIDRGP(MVE)	Improvised Diagnostic Robust Generalized Potential based on Minimum Volume Ellipsoid in panel data.
WIDRGP(ISE)	Improvised Diagnostic Robust Generalized Potential based on Index Set Equality in panel data.
RVIF(WGM-FIMGT)	Robust Variance Inflation Factor based on Fast Improvised Diagnostic Robust Generalized Potential in panel data.
WRJFIMGT	Robust Jack-knife Ridge Regression based on GM-FIMGT estimator in panel data
WRJMM	Robust Jack-knife Ridge Regression based on MM-estimator in panel data
W RJGM2	Robust Jack-knife Ridge Regression based on GM2-estimator in panel data

CHAPTER 1

INTRODUCTION

1.1 Background and Purposes

Regression analysis is a statistical process for estimating the linear relationships between two or more variables. It involves several techniques for modeling and analyzing several variables. The earliest form of regression was the Least Squares, commonly known as Ordinary Least Squares (OLS) method which was introduced by the two famous statisticians Legendre and Gauss (Maronna et al.,2006). To this day, empirical researchers use OLS and its generalizations since the Gauss-Markov theorem asserts that OLS provides the Best Linear Unbiased Estimator (BLUE) for the parameters of the standard linear model. The estimator is best, in the sense that it is the most efficient one among the linear estimators when the data are smooth or normally distributed, i.e. do not contain outlying observations. Rousseeuw and Van Zomeren (1990) distinguished outliers (high leverage points and/or vertical outliers) when defining these anomalous observations. Data items for which the independent variable lies far from the majority of the explanatory observations are called leverage points. On the other hand, the data points which are far away from the majority of the data point in Y-direction are refers to as vertical outlier. This is the reason why observations corresponding to very large residuals are treated as residual outliers. These observations deserve special attention from the statistician since they may invalidate classical statistical inference (Maronna, 1976;Maronna et al.,2006; Tyleret al.,1994; Hubber, 2005). OLS is inefficient and produce dramatically different estimates even when a single outlying observation is added or present in a data set. The survival of OLS for about two centuries in empirical studies is not justified by its performance on contaminated data (Belsley et al., 1980; Hocking and Pendelton, 1983; Rousseeow and Leroy, 1987).

Similarly, an observation can be both a vertical outlier and a high leverage point (HLP) or can be horizontal outlying without being a vertical outlier when it perfectly fits the linear relation between the independent variable and the explanatory variables which referred to as good leverage points. Good leverage points may have no effects in OLS technique.

Habshah et al. (2009) developed the Diagnostic Robust Generalized Potential (DRGP) based on MVE to improve the rate of detection of high leverage points. Also, Mohammed and Midi (2015) proposed MGti-DRGP which is very successful to classifying observations into regular observation, vertical outliers, good and bad leverage points. Imon et al. (2015) introduced a robust influence distance that can identify multiple IOs, and propose a six-fold plotting technique based on the well-known group deletion approach to classify regular observations, outliers, high leverage points and IOs simultaneously in linear regression.

Imon and Khan (2003) verified that HLPs is a new source of multicollinearity. HLPs have high impact on the OLS estimates in regression model and also responsible for causing multicollinearity problem as they may increase (enhancing observation) or decrease (reducing observation) as explained by Midi et al. (2011). Multicollinearity exist in a data set when two or more independent variables are highly correlated.

Moreover, panel data regression model is one of the most widely used models especially in finance and economics due to the advantage it has over cross-sectional and time-series model. A data collected over time and over the same individual is referred to as Panel data. It is usually analyzed by running a regression over this two-dimensions (cross section/time series) using a classical least square called pooled OLS (pooling of time series and cross-sectional observations). Two methods of analysis exist for panel data models i.e. fixed effect and random effect model. The major difference between these two models is the definition of the time invariant variable in the data set. Therefore, the same problem of multicollinearity and outlier affect panel data estimator as in classical linear regression.

Therefore, in general the existence of multicollinearity and anomalous points causes OLS deviate from the normality assumption in regression analysis. This was addressed by many researchers in articles and books. In recent years, several techniques and methods which deal separately with multicollinearity and outliers are available such as: Imon and Khan (2003) attempted to show how generalized potentials can be used as a remedy to multicollinearity problem due to HLP. Bagheri and Midi (2011) developed variance inflation factor to be more resistant to HLPs. They proposed robust variance inflation factor based on GM(DRGP) for detection of multicollinearity when the source is due to HLPs. Nonetheless, there is not much significant work reported in the literature which takes into account the presence of both multicollinearity and outlying observations problems concurrently (see Johnston, 1984; Montgomery et al., 2015; Gujarati, 2002; Kutner et al., 2004; Chatterjee and Hadi, 2006; Kamruzzaman and Imon, 2002; Imon, 2005).

1.2 Importance and Motivation of the Study

The presence of outlying observation in a data set has serious problem in the parameter estimation. It is now evident that in the presence of outliers, inferential procedure will produce an invalid inferential statement, for example, the OLS performs poorly in the presence of outliers (Rousseeuw and Leroy, 1987). It is very important to detect them so that appropriate measure can be taken. Moreover, many statistics practitioners are not aware that erroneous estimation may occur due to the presence of outliers. Detection of outliers is very important in statistical estimation. However, most of the method of detecting outliers may not perform well due to the swamping and masking effect. Different methods have been developed in the literature to detect the existence of outliers in a data set.

The existences of HLPs in a data set are responsible for causing masking and swamping effect of outliers in linear regression (Pena and Yohai, 1995). The HLPs cause multicollinearity problem and also have a great effect on the potential values (Hadi, 1992). The Hadi's method detect a single leverage point but they are not successful to identify multiple leverage points (Ruppert and Simpson, 1990; Imon, 2005, Habshah et al., 2009). This problem was addressed by Imon (1996), where he proposed a generalized potentials (GP) as a diagnostic technique for identifying multiple HLPs. The generalized potential method provides an extension from a single detection to multiple detection of HLPs. Also, the GP suffers much set back due to its inefficiency to successfully identifying the correct number of HLPs due to the masking effects (Habshah et al. 2009). The remedial measure of this problem is the provision of another step by Habshah et al. (2009) to confirm whether all the detected HLPs in Step 1 is a genuine HLPs. They developed a new method called Diagnostic Robust Generalized Potential (DRGP) which is very successful in identifying HLPs and also reduce the rate of masking and swamping effect. Nonetheless, the DRGP technique is much successful for small sample size and high percentage of contamination. Midi and Mohammed (2015) added another step to the DRGP algorithm termed as Improvised DRGP (IDRGP) in order to improve the efficiency of DRGP method and to reduce the rate of masking and swamping. The problem of this IDRGP is that the procedure is very time consuming as it employed minimum volume ellipsoid (MVE) in it computation. Lim and Midi (2016) proposed diagnostic robust generalized potential based on Index Set Equality ISE (DRGP(ISE)) which is less time consuming. However, the method of Lim and Midi (2016) still has some percentage of swamping and masking effect. We are aware from these discussion that several shortcomings can be seen in some of these methods, such as having swamping and masking, longer running times and computational complexity.

Therefore, their work has motivated us to propose a Fast Improvised Diagnostic Robust Generalized Potential based on Index Set Equality (IDRGP (ISE)), which is expected to be more efficient in the identification of HLPs and able to reduce the rate of swamping, masking and computational time. To date, no research has been done to identify outliers /HLP for panel data. Hence, we also propose a new method of identification of HLPs in panel data, an extension identification method from linear regression model to fixed panel data model.

The thesis also considered an efficient estimation technique in linear and fixed effect panel data model when there is HLPs in the data set. The GM6 estimation method proposed by Rousseeuw (1985) is the widely used method to overcome the HLPs. However, this method is less efficient since it is based on robust Mahalanobis Distance Square which utilized minimum volume ellipsoid (MVE) as an initial π–weight function. The shortcoming of MVE is that it is not only tends to swamp some low leverage points as high leverage, it attempts to identify high leverage points without taking into consideration whether they are good or bad leverage points. Hence, the GM6 considers the good leverage point as bad leverage points and its efficiency tends to decrease as the number of good leverage points increases. Dhhan et al. (2016) have successfully developed a new GM estimator that satisfied all the three properties of good robust estimator. However, the method is based on support vector regression which is quite complicated and difficult to understand by non-expert SVR statistics practitioners. Their work has motivated us to develop another version of GM estimator which is relatively simple and easy to understand compared to GM-SVR (Dhhan et al., 2016) and uses less computational running time. The new proposed GM estimator termed Fast GM estimator (FIMGT) which is quite fast and only down weight vertical outliers and bad leverage points. The good high leverage points are not down weighted because they have no impact or little effect on the parameter estimates and may contribute to the precision of the estimates. We also proposed the Within Group WGM-FIMGT estimator for fixed effect panel data model an extension from the linear regression model method. To the best of our knowledge no such method has been proposed in panel data model.

The thesis also addressed the problem of multicollinearity in parameter estimation in the presence of HLPs. The OLS estimator suffers tremendous effect in the presence of multicollinearity. The presence of high leverage points also effect multicollinearity.

The traditional multicollinearity diagnostic methods cannot correctly detect the existence of multicollinearity when there is HLPs in a data (Rosen, 1999). It is now evident that the traditional diagnostic measure Variance Inflation Factor (VIF) cannot correctly detect multicollinearity in the presence of HLPs. Bagheri and Midi (2011) proposed RVIF(MM) and RVIF(GM(DRGP)) to diagnose multicollinearity. Nonetheless, the RVIF(MM) is not efficient (Bagheri and Midi, 2011). The RVIF(GM(DRGP)) is also less effective because it is formulated based on DRGP which is less efficient as it down weight all detected HLPs irrespective of whether it is good or bad. The shortcoming of this method has inspired us to propose new robust VIF, namely the RVIF(GM-FIMGT) which is anticipated to be more efficient and reliable as it was developed based on IDRGP(ISE) which is relatively fast and only down weight vertical outliers and bad leverage point. Again we would like to develop robust VIF for panel data since such measure has not been focused by any researcher in fixed effect panel data model.

This thesis also considered the parameter estimation to remedy multicollinearity in the presence of HLPs. The Jackknife ridge estimator (JRR) and Latent Root Regression (LRR) which have small bias is put forward to remedy this problem. Nevertheless, it is now evident that these classical estimation methods perform poorly when outliers exist in a data set. In literature, not much work is available for the combined problems of multicollinearity and the existence of outliers in linear as well as panel data model. Jadhav and Kashid (2011) suggested using a Jackknife ridge M-estimator to overcome multicollinearity and outliers in the Y direction. Mohammed and Midi (2015) integrated MM-estimator and the GM2-estimator in the JRR algorithm for the establishment of the improvised versions of JRR. The suggested method is called jackknife ridge MM based denoted by JRMM and the jackknife ridge GM2denoted by JRGM2.As already mentioned, the shortcoming of GM2 is that it formulation is based on DRGP which down weight HLPs without considering whether it is bad or good leverage points. In order to improve this estimator, we propose new Jackknife ridge (GM-FIMGT), denoted by JRFIMGT which is based on FIMGT with reasons already mentioned. This method simultaneously rectifies the problems of multicollinearity and outliers.

As already started earlier, this thesis also addressed the same issues for panel data model. Some robust estimation technique for panel data were proposed, such as Bramati and Croux (2007) who applied the robust Generalized M estimator and also combine of S and M-estimates to provide alternatives to the classical Within Group estimator using median-centering data transformation, in which the data are centered within the time series by using the median instead of the mean in order to eliminate the fixed effect in a robust method. The within GM-estimator and Within MS estimator moderately achieved low breakdown points. Recently, Verardi and Wagner (2011) used S-estimator for another robust Within Group estimator. They used the same method of robust data transformation in their studies. Unfortunately, median centering method is found to produce nonlinearity to the resulting data and make the equivariance properties of the robust estimators redundant (Bakar and Midi, 2015).

More recently, Bakar and Midi (2015) used different centering approach whereby data are centered by MM-estimate of location and then employed robust MM and robust GM6 within group estimator. This robust approach is maintained not only to bring linearity back to the transformed data but also to enhance their performances. The weaknesses of their method is that it is based on MM estimate which is not bounded influence and also based on GM6 which down weight all HLPs irrespective of whether they are Good HLPs or Bad HLPs. This motivated us to employ a new proposed method namely WGM-FIMGT to the transformed data following Bramati and Croux (2007) who applied the GM6 estimator to the transformed data. Our transformed data is based on MM centering but Bramati and Croux (2007) based on median centering which has low efficiency under normal distribution.

This thesis also focuses on both detection of multicollinearity and estimation of parameter in fixed effect panel data model. To the best of our knowledge, to date no work has been done for simultaneously taking care of multicollinearity and outliers in panel data. Hence, this inspired us to extend all methods that we developed earlier in linear model to panel data model.

1.3 Objective of Thesis

The foremost objective of our research can be outlined systematically as follows:

1. To develop a new fast method for detecting HLPs in linear regression and fixed effect panel data model.
2. To develop a simple version of high breakdown, high efficiency, bounded influence and fast GM estimator based on Index Set Inequality in both linear regression and fixed effect panel data model.
3. To formulate a new method for detecting multicollinearity in the presence of HLPs in both linear regression and fixed effect panel data.
4. To establish a new parameter estimation method to remedy multicollinearity in both linear and fixed effect panel data model in the presence of HLPs.

1.4 Scope and Limitation of study

Panel data is still anew area in robust statistics. It is widely used in many field of study such as economics, finance and social science. Since robust statistic is relatively new technique in panel data model, there are not so many algorithms and statistical softwares related to panel data are available. Writing our own programming codes is the most challenging job.

Since not much robust work is developed in panel data model, not many well referred outlying datasets and references are variable in the literature for discussion purposes.

Not to mention that the outlying datasets with multicollinearity problems in panel data. Thus, generated data are used to apply our proposed method in panel datasets.

1.5 Outline of the Thesis

In accordance with the objectives and the scope of the study, the contents of this thesis are organized in the nine chapters. The thesis chapters are structured so that the study objectives are apparent and are conducted in the sequence outline.

Chapter Two : This chapter briefly presents the literature reviews of the ordinary least squares estimation method and basic concepts of robust regression. Diagnostic methods of vertical outliers and high leverage points are reviewed. Moreover, important existing robust regression methods for estimation parameter in the presence of HLPs and vertical outliers are also presented. The literature reviews on multicollinearity diagnostic methods and remedial techniques are presented. Some literatures in panel data also include in this chapter.

Chapter Three : This chapter discusses the existing Fast Improvised Diagnostic Robust Generalized Potential (IDRGP) which is developed by Midi and Mohammed (2015). The new proposed Fast Improvised Diagnostic Robust Generalized Potential method based on Index Set Equality (IDRGP (ISE)) for identification of HLPs is presented. Finally, a Monte Carlo simulation study is discussed to evaluate the performance of the proposed method.

Chapter Four : This chapter deals with development of the GM-estimator denoted by GM-FIMGT in multiple linear regressions. Monte Carlo simulations are presented to assess the performance of the proposed method.

Chapter Five : In this chapter, we extend our proposed (IDRGP(ISE)) method that have discussed in Chapter three to panel data setting for the identification of HLPs. Also, the effect of high leverage points before and after transformation is presented.

Chapter Six : In this chapter, we used our proposed method that is discussed in Chapter four for estimation parameter in panel data set. Monte Carlo simulation studies and numerical example is carried out to assess the performance of the proposed method.

Chapter Seven : This chapter is divided into two sections;

First Section : deals with a new proposed VIF based on GM-FIMGT denoted by RVIF(GM-FIMGT) method to diagnose multicollinearity in multiple linear regression model.

Second Section : In this section, we extend our new proposed method in the first section to panel data setting to detect the multicollinearity problem in the presence of HLPs. Monte Carlo simulation studies and two artificial data sets are carried out to assess the performance of the proposed method.

Chapter Eight : we divided this chapter into two sections;

First Section : deals with robust jackknife ridge regression estimation method named (RJFIMGT) to remedy problem of multicollinearity in the presence of high leverage points in linear regression. A Monte Carlo simulation study and some numerical examples are given to assess the performance of our proposed method.

Second Section : In this section, we extend our proposed method in the first section to panel data setting to remedy problem of multicollinearity in the presence of high leverage points. A Monte Carlo simulation study is presented to assess the performance of our method.

Chapter Nine : This chapter provides summary and detailed discussions of the thesis conclusions. Areas for future research are also recommended.

REFERENCES

Akdeniz Duran, E. and F. Akdeniz (2012). "Efficiency of the modified jackknifed Liu-type estimator." Statistical Papers 53(2): 265-280.

Atkinson, D. (1985). "Effects of some plant growth regulators on water use and the uptake of mineral nutrients by tree crops." V International Symposium on Growth Regulators in Fruit Production 179.

Alguraibawi, M., Midi, H., and Imon, A. (2015). "A new robust diagnostic plot for classifying good and bad high leverage points in a multiple linear regression model." Mathematical Problems in Engineering. Volume 2015, Article ID 279472, pp 1-12

Andersen, R. (2008). "Modern methods for robust regression." The United States of America: Sara Miller McCune. SAGE publications.

Aquaro, M. and P. Cizek (2010). "One-step robust estimation of fixed-effects panel data models."

Arellano, M. and S. Bond (1991). "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations." The review of economic studies 58(2): 277-297.

Arkin, R. G. and D. C. Montgomery (1980). "Augmented robust estimators." Technometrics 22(3): 333-341.

Bagheri, A. (2011). "Robust Estimation Methods And Robust Multicollinearity Diagnostics For Multiple Regression Model in the Presence of High Leverage Collinearity-Influential Observations." Unpublished doctorial thesis. University Putra Malaysia .

Bagheri, A., Habshah M., and Imon, R. H. M. R. (2012). "A novel collinearity-influential observation diagnostic measure based on a group deletion approach." Communications in Statistics-Simulation and Computation, 41(8), 1379-1396.

Bagheri, A. and H. Midi (2015). "Diagnostic plot for the identification of high leverage collinearity-influential observations." SORT 39(1): 51-70.

Bagheri, A., Habshah M., and Imon, R. H. M. R. (2010). "The effect of collinearity-influential observations on collinear data set: A monte carlo simulation study." Journal of Applied Sciences 10: 2086-2093.

Bakar, N. M. A. and H. Midi (2015). "Robust centering in the fixed effect panel data model." Pakistan Journal of Statistics 31(1).

Baltagi, B. (2008). "Econometric analysis of panel data. "John Wiley & Sons.

Baltagi, B. H. (2014). "The Oxford handbook of panel data." Oxford Handbooks.

Baltagi, B. H. and M. D. Giles (1998). "Panel data methods." Statistics Textbooks and Monographs 155: 291-324.

Baltagi, B. H. and D. Li (2002). "Series estimation of partially linear panel data models with fixed effects." Annals of economics and finance 3(1): 103-116.

Baltagi, B. H., and W. Koh. (2003). "Testing panel data regression models with spatial error correlation." Journal of econometrics 117(1): 123-150.

Batah F. S., Ramanathan T. V. and Gore S. D. (2008). "The efficiency of modified jackknife and ridge type regression estimators: a comparison, Surv. Math. Appl, 3:111-122.

Belsley, D.A., Kuh, E. and Welsch, R.E. (1980). "Regression Diagnostics: Identifying Influential Data and Sources of Collinearity." New York :Wiley.

Blundell, R. and S. Bond (1998). "Initial conditions and moment restrictions in dynamic panel data models." Journal of econometrics 87(1): 115-143.

Bramati, M. C. and C. Croux (2007). "Robust estimators for the fixed effects panel data model." The econometrics journal 10(3): 521-540.

Brown, P. J. (1977). "Centering and scaling in ridge regression." Technometrics 19(1): 35-36.

Campbell, N. A., Lopuhaä, H. P. and Rousseeuw, P. J. (1998) ." On the calculation of a robust S-estimator of a covariance matrix." Statistics in Medicine. 17(23): 2685-2695.

Chatterjee, S. and A. S. Hadi (2006). "Simple linear regression." Regression Analysis by Example, Fourth Edition: 21-51.

Chatterji, S. and A. Hadi (1988). "Sensitivity Analysis in Linear Regression." Wiley, New York.

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

Croux, C. and G. Haesbroeck (2003). "Implementing the Bianco and Yohai estimator for logistic regression." Computational Statistics & Data Analysis 44(1): 273-295.

Dhhan, W., Rana, S., & Midi, H. (2015). "Non-sparse ϵ-insensitive support vector regression for outlier detection." Journal of Applied Statistics, 42(8), 1723-1739

Dhhan, W., Rana, S., & Midi, H. (2016). "A high breakdown, high efficiency and bounded influence modified GM estimator based on support vector regression." Journal of Applied statistics 44(4): 700-714.

Donoho, D. L. and P. J. Huber (1983). "The notion of breakdown point." A festschrift for erich l. lehmann 157184.

Groβ, J. (2003). "Linear regression." Lecture Notes in Statistics 175: 12.

Grieser, W. D., & Hadlock, C. J. (2015). "Panel data estimation in finance: parameter consistency and the standard error sideshow. " Working Paper.

Gujarati, D. and D. Porter (2002). "Basic econometrics." McGraw-Hill/Irwin.

Gujarati, D. N. (2009). "Basic econometrics." Tata McGraw-Hill Education.

Habshah, M., Norazan, M.R. and Imon, A.H.M.R. (2009). "The performance of Diagnostic-Robust Generalized Potentials for the identification of multiple high leverage points in linear regression." Journal of Applied Statistics. 36(5): 507-520.

Hadi, A. S. (1988). "Diagnosing collinearity-influential observations." Computational Statistics & Data Analysis 7(2): 143-159.

Hadi, A. S. (1992). "A new measure of overall potential influence in linear regression." Computational and Statistical Data Analysis. 14:1-27.

Hadi, A. S. (1992). "A new measure of overall potential influence in linear regression." Computational Statistics & Data Analysis 14(1): 1-27.

Hadi, A. S. and J. S. Simonoff (1993). "Procedures for the identification of multiple outliers in linear models." Journal of the American Statistical Association 88(424): 1264-1272.

Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J. and Stahel, W. A. (1986). "Robust Statistics. " New York: Wiley.

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

Hekimoğlu, S. and R. C. Erenoglu (2013). "A new GM-estimate with high breakdown point." Acta Geodaetica et Geophysica 4(48): 419-437.

Hinkley, D. V. (1977). "Jackknifing in unbalanced situations." Technometrics, 19(3), 285-292.

Hill, R. W. (1977). "Robust regression when there are outliers in the carriers." Harvard University.

Hill, R. W. and P. W. Holland (1977). "Two robust alternatives to least-squares regression." Journal of the American Statistical Association 72(360a): 828-833.

Hoaglin, D. C and Welsch, R. E. (1978). "The hat matrix in regression and ANOVA." American Statistician. 32:17-22.

Hocking, R. and O. Pendleton (1983). "The regression dilemma." Communications in Statistics-Theory and Methods 12(5): 497-527.

Hoerl, A. E. and R. W. Kennard (1970). "Ridge regression: Biased estimation for nonorthogonal problems." Technometrics 12(1): 55-67.

Hoerl, A. E. and Kennard, R. W. (1970a)." Ridge regression: applications to non-orthogonal problems." Technometrics. 12:69-82.

Hoerl, A. E., and Kennard, R. W. (1970b)." Ridge regression: biased estimation for non- orthogonal problems." Technometrics.12:55-67.

Hsiao, C. (2014). "Analysis of panel data." Cambridge university press.

Hsieh, F., Lavori P. and Cohen H. (2003). "An overview of variance inflation factors for sample-size calculation." Evaluation & the Health Professions 26(3): 239-257.

Huber, P. and E. Ronchetti (1981). "Robust Statistics, ser." Wiley Series in Probability and Mathematical Statistics. New York, NY, USA, Wiley-IEEE 52: 54.

Huber, P. J. (2005). "Robust statistics." John Wiley & Sons.

Imon, A. (2002). "Identifying multiple high leverage points in linear regression." Journal of Statistical Studies 3: 207-218.

Imon, A. and M. A. I. Khan (2003). "A solution to the problem of multicollinearity caused by the presence of multiple high leverage points." Int. J. Stat. Sci 2: 37-50.

Jadhav, N. and D. Kashid (2011). "A jackknifed ridge M-estimator for regression model with multicollinearity and outliers." Journal of statistical theory and practice 5(4): 659-673.

Johnston, J. (1984). "Econometric methods ." 3rd Edn.. New York: McGraw Hill.

Kamruzzaman, M. and A. Imon (2002). "High leverage point: another source of multicollinearity." PAKISTAN JOURNAL OF STATISTICS-ALL SERIES- 18(3): 435-448.

Katz, M. H. (2006). "Multivariate Analysis: a Practical Guide for Clinicians." UK: Cambridge University Press.

Kiviet, J. F. (1995). "On bias, inconsistency, and efficiency of various estimators in dynamic panel data models." Journal of econometrics 68(1): 53-78.

Krasker, W. S. and R. E. Welsch (1982). "Efficient bounded-influence regression estimation." Journal of the American Statistical Association 77(379): 595-604.

Kutner, M. H., Neter J and Nachtsheim C.J. and Wasserman W. (2004). "Applied linear regression models." McGraw-Hill/Irwin.

Lawrence, K. and J. Arthur (1990). "Robust nonlinear regression." Robust Regression. Analysis and Applications. New York, NY: Marcel Dekker Ine: 59-86.

Leroy, A. M. and P. J. Rousseeuw (1987). "Robust regression and outlier detection." Wiley Series in Probability and Mathematical Statistics, New York: Wiley, 1987.

Li, G., and Chen, Z. (1985). "Projection-Pursuit Approach to Robust Dispersion Matrices and Principal Components: Primary Theory and Monte Carlo." Journal of the American Statistical Association. 80: 759–766.

Lim, H. A. and H. Midi (2016). "Diagnostic Robust Generalized Potential Based on Index Set Equality (DRGP (ISE)) for the identification of high leverage points in linear model." Computational Statistics 3(31): 859-877.

Lukman, A., Arowolo O. and Ayinde K. (2014). "Some Robust Ridge Regession for Handling Multicollinearity and Outliers." International Journal of Sciences: Basic and Applied Research (IJSBAR) 16(2): 192-202.

Lyu, Y. (2015). "Detection of outliers in panel data of intervention effects model based on variance of remainder disturbance." Mathematical Problems in Engineering 2015.

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

Manuel, A. and B. Stephen (1991). "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations." Review of Economic Studies.

Maronna, R.A., Martin, R.D. and Yohai, V.J. (2006). "Robust Statistics Theory and Methods." New York: Willy and sons.

Maronna, R. A. and V. J. Yohai (1976). "Robust estimation of multivariate location and scatter." Wiley StatsRef: Statistics Reference Online.

Marquaridt, D. W. (1970). "Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation." Technometrics 12(3): 591-612.

McDonald, G. C. and D. I. Galarneau (1975). "A Monte Carlo evaluation of some ridge-type estimators." Journal of the American Statistical Association 70(350): 407-416.

Meyers, L. S., Gamst, G. and Guarino, A. J. (2006). "Applied Multivariate Research: Design and Interpretation." Sage publications, INC.

Midi, H. (1999). "Preliminary estimators for robust non-linear regression estimation." Journal of Applied statistics 26(5): 591-600.

Habshah M., Bagheri, A., and Imon, A. H. M. (2010). " The Application of Robust Multicollinearity Diagnostic Method Based on Robust Coefficient Determination to a Non-Collinear Data." Journal of Applied Sciences, 10(8):611-619.

Habshah M. Bagheri, A. and Imon, R. H. M. R. (2011). "A monte carlo simulation study on high leverage collinearity-enhancing observation and its effect on multicollinearity pattern." Sains Malaysiana 40(12): 1437-1447.

Midi, H. and A. Mohammed (2015). "The Identification of Good and Bad High Leverage Points in Multiple Linear Regression Model." Mathematical Methods and System in Science and Engineering.

Midi, H. and M. Zahari (2007). "A simulation study on ridge regression estimators in the presence of outliers and multicollinearity." Jurnal Teknologi 47: 59-74.

Mohd Salleh, R. (2013)." A robust estimation method of location and scale with application in monitoring process variability." Universiti Teknologi Malaysia.

Mohammed A. (2015). "Robust diagonostic methods and robust parameter estimation techniques for linear regression model having multicollinearity problem and outliers". Unpublished doctorial thesis, university putra Malaysia.

Moller, S.F., Frese, J.V. and Bro, R. (2005). "Robust Methods for multivariate data analysis." Journal of Chemometrics. 19(10): 549-563.

Montgomery, D. C., Peck, E. A., and Vining, G. G. (2015)." Introduction to linear regression analysis." John Wiley & Sons.

Montgomery, E., Bronner, M. P., Goldblum, J. R., Greenson, J. K., Haber, M. M., Hart, J. & Washington, K. (2001)." Reproducibility of the diagnosis of dysplasia in Barrett esophagus: a reaffirmation. Human pathology, 32(4), 368-378.

Myers, R. H. (1990). "Classical and modern regression with applications".2nd edition. CA: Duxbury press.

Myers, R. H. and D. C. Montgomery (1997). "A tutorial on generalized linear models." Journal of Quality Technology 29(3): 274.

Peña, D. and V. J. Yohai (1995). "The detection of influential subsets in linear regression by using an influence matrix." Journal of the Royal Statistical Society. Series B (Methodological): 145-156.

Pison, G., Rousseeuw, P. J., Filzmoser, P., and Croux, C. (2003). "Robust factor analysis." Journal of Multivariate Analysis, 84(1), 145-172.

Quenouille, M. H. (1956). "Notes on bias in estimation. "*Biometrika*, 353-360.

Rahmatullah Imon, A. (2005). "Identifying multiple influential observations in linear regression." Journal of Applied statistics 32(9): 929-946.

Raj, B. and B. H. Baltagi (2012). "Panel data analysis." Springer Science & Business Media.

Riazoshams, H. and H. B. Midi (2016). "The Performance of a Robust Multistage Estimator in Nonlinear Regression with Heteroscedastic Errors." Communications in Statistics-Simulation and Computation 45(9): 3394-3415.

Rohayu, M.S. (2013). "A Robust Estimation Method of Location and Scale with Application in Monitoring Process Variability." Unpublished Ph.D. Thesis, Universit Teknologi Malaysia, Malaysia.

Rojo-Álvarez, J. L., Martínez-Ramón, M., Figueiras-Vidal, A. R., García-Armada, A., & Artés-Rodríguez, A. (2003). "A robust support vector algorithm for nonparametric spectral analysis." IEEE Signal Processing Letters 10(11), pp 320-323.

Rosen, D. (1999). "The diagnosis of colinearity. A Monte Carlo simulation method." Dissertation.

Rousseeuw, P. J. (1984). "Least median of squares regression." Journal of the American Statistical Association 79(388): 871-880.

Rousseeuw, P. J. (1985). "Multivariate estimation with high breakdown point." Mathematical statistics and applications 8: 283-297.

Rousseeuw, P. J. and C. Croux (1993). "Alternatives to the median absolute deviation." Journal of the American Statistical Association 88(424): 1273-1283.

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

Ruppert, D. (1992). "Computing S estimators for regression and multivariate location/dispersion." Journal of Computational and Graphical Statistics 1(3): 253-270.

Ruppert, D. and D. G. Simpson (1990). "Unmasking multivariate outliers and leverage points: comment." Journal of the American Statistical Association 85(411): 644-646.

Ryan, T. P. (1997). "Modern Regression Methods. "NewYork: Wiley.

Sengupta, D. and P. Bhimasankaram (1997). "On the roles of observations in collinearity in the linear model." Journal of the American Statistical Association 92(439): 1024-1032.

Simpson, D. G., Ruppert, D. and Carroll, R. J. (1992). " On one-step GM estimates and stability of influences in linear regression." Journal of the American statistical association. 87: 439-450.

Simpson, J. R. (1995). "New Methods and Comparative Evaluations for Robust and Biased-Robust Regression Estimation." Unpublished Ph.D. thesis, Arizona State University, The United States of America.

Simpson, J. R. and D. C. Montgomery (1998). "The development and evaluation of alternative generalized m-estmation techniques." Communications in Statistics-Simulation and Computation 27(4): 999-1018.

Singh B., Chaubey Y. P. and Dwivedi T. D. (1986), “An almost unbiased ridge estimator”, *The Indian Journal of Statistics*,(48): 342-346.

Stine, R. A. (1995). "Graphical interpretation of variance inflation factors." The American Statistician 49(1): 53-56.

Stromberg, A. J., Hossjer, O. and Hawkins, D. M. (2000). "The least trimmed differences regression estimator and alternatives." Journal of the American Statistical Association. 95: 853-864.

Tyler, D. E. (1994). "Finite sample breakdown points of projection based multivariate location and scatter statistics." The Annals of Statistics: 1024-1044.

Vellman, P.F. and Welsch, R.E. (1981). "Efficient computing of regression diagnostics." American Statistician. 27: 234-242.

Verardi, V. and J. Wagner (2011). "Robust estimation of linear fixed effects panel data models with an application to the exporter productivity premium." Jahrbücher für Nationalökonomie und Statistik 231(4): 546-557.

Víšek, J. Á. (2015). "Estimating the Model with Fixed and Random Effects by a Robust Method." Methodology and Computing in Applied Probability 17(4): 999-1014.

Wilcox, R. R. (2005). "Introduction to robust estimation and hypothesis testing (Statistical Modeling and Decision Science." Academic press.

Woodruff, D. L. and Rocke, D. M. (1994). "Computable robust estimation of multivariate location and shape in high dimension using compound estimators." Journal of the American Statistical Association. 89:888–896

Wooldridge, J. M. (1995). "Selection corrections for panel data models under conditional mean independence assumptions." Journal of econometrics 68(1): 115-132.

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

Yohai, V. J. and R. H. Zamar (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.

Zaman, A., et al. (2001). "Econometric applications of high-breakdown robust regression techniques." Economics Letters 71(1): 1-8.