UNIVERSITI PUTRA MALAYSIA

ROBUST ESTIMATION TECHNIQUE AND ROBUST AUTOCORRELATION DIAGNOSTIC FOR MULTIPLE LINEAR REGRESSION MODEL WITH AUTOCORRELATED ERRORS

LIM HOCK ANN

FS 2014 9

ROBUST ESTIMATION TECHNIQUE
AND
ROBUSTAUTOCORRELATION DIAGNOSTIC
FOR
MULTIPLE LINEAR REGRESSION MODEL
WITH AUTOCORRELATED ERRORS

LIM HOCK ANN

DOCTOR OF PHILOSOPHY
UNIVERSITI PUTRA MALAYSIA

2014

ROBUST ESTIMATION TECHNIQUE AND ROBUST AUTOCORRELATION DIAGNOSTIC FOR MULTIPLE LINEAR REGRESSION MODEL WITH AUTOCORRELATED ERRORS

By

LIM HOCK ANN

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

August 2014

DEDICATION

I would like to dedicate this dissertation work to

- *my respectful parents, who have taught me a lot on the meaning of persistency in life.*

- *my beloved wife, for being so patient and understanding throughout my doctoral pursue. Her love has always been my greatest support and achor in my life.*

- *my precious newborn son, Isaac. He is such a joy and pride to me and my wife.*

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Doctor of Philosophy

ROBUST ESTIMATION TECHNIQUE AND ROBUST AUTOCORRELATION DIAGNOSTIC FOR MULTIPLE LINEAR REGRESSION MODEL WITH AUTOCORRELATED ERRORS

By

LIM HOCK ANN

August 2014

Chairperson: Professor Habshah Midi, Ph.D.

Faculty: Science

Autocorrelated errors cause the Ordinary Least Squares (OLS) estimators to become inefficient. Hence, it is very essential to detect the autocorrelated errors. The Breusch-Godfrey (BG) test is the most commonly used test for detection of autocorrelated errors. Since this test is easily affected by high leverage points, the robust Modified Breusch-Godfrey (MBG) test is proposed. The results of the study indicate that the MBG test is a robust test to detect the autocorrelated errors.

Thus far, there is no specific method proposed to identify high leverage points in linear model with autocorrelated errors. Hence, the Diagnostic Robust Generalized Potentials Based on Index Set Equality (DRGP(ISE)) is proposed to close the gap in the literature. The findings indicate that DRGP(ISE) is an excellent and fast identification method to detect the high leverage in linear model with autocorrelated errors.

High leverage points have tremendous effect in regression analysis. In this study we verified that high leverage points is another cause of autocorrelation.

Not much research has been done to investigate autocorrelation-influential observations. Hence, the Robust Autocorrelation-Influential Measure based on DRGP (RAIM(DRGP)) is formulated to identify the autocorrelation-influential observations in autocorrelated data. The RAIM(DRGP) is found to do a credible job to identify the high leverage autocorrelation-enhancing and autocorrelation-reducing observations and autocorrelation-influential observations.

Cochrane-Orcutt Prais-Winsten (COPW) iterative method is the most commonly used remedial measure to rectify the autocorrelation problems. However, this procedure is extremely vulnerable in the presence of high leverage points. On the other hand, the autocorrelation may be caused by the presence of high leverage points. The Robust Cochrane-Orcutt Prais-Winsten (RCOPW) iterative method is

therefore developed. The results of the study show that RCOPW estimation is a robust remedial measure in the case when the autocorrelated data come together with the presence of high leverage points and also in the autocorrelation caused by the high leverage points.

The existing diagnostic plot does not take into the consideration of autocorrelated errors. Thus, the robust remedial of autocorrelated errors - Robust Cochrane-Orcutt Prais-Winsten (RCOPW) is incorporated in the diagnostic plot to form the Diagnostic Plot for Autocorrelation Based on Standardized Cochrane-Orcutt Prais-Winsten Residuals (DPA-RCOPW). The results based on simulated autocorrelated data and well known outlying datasets show that DPA-RCOPW successfully identifies and classifies the outlying observations according to its types precisely.

In this study, an alternative method of finding confidence intervals of regression parameters for autocorrelated data in the presence of high leverage points and for autocorrelation caused by high leverage points is proposed. The findings provide strong evidences that the Diagnostic Before Bootstrap based on Robust Cochrane-Orcutt Prais-Winsten estimate (DBB RCOPW) estimate is a robust procedure and consistently provides close answers to the actual confidence intervals of the regression parameters for data with autocorrelated errors in the presence of high leverage points and for autocorrelation caused by high leverage points.

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

TEKNIK PENGANGGARAN TEGUH DAN DIAGNOSTIK AUTOKORELASI TEGUH BAGI MODEL LINEAR BERGANDA RALAT BERAUTOKORELASI

Oleh

LIM HOCK ANN

Ogos 2014

Pengerusi : Profesor Habshah Midi, Ph.D.

Fakulti: Sains

Masalah autokorelasi menyebabkan anggaran Pengangar Kuasadua Terkecil Biasa (OLS) tidak cekap. Oleh sebab itu, adalah sangat penting untuk mengesan masalah ralat berautokorelasi. Ujian Breusch-Godfrey (BG) adalah antara ujian yang paling biasa digunakan untuk mengesan ralat berautokrelasi. Oleh kerana ujian ini senang dipengaruhi oleh titik tuasan tinggi, ujian Terubahsuai Breucsh-Godfrey (MBG) dicadangkan. Keputusan kajian menunjukkan ujian MBG adalah ujian pengesahan yang teguh dalam pengesahan ralat berautokorelasi.

Setakat ini, tidak ada kaedah spesifik yang dicadangkan untuk mengenal pasti titik tuasan tinggi dalam model linear ralat berautokorrelasi. Degan itu, Kaedah Teguh Berdiagnostik Potensi Teritlak Berasaskan Indeks Set Kesaksamaan (DRGP(ISE)) dicadangkan untuk menutup jurang kesusasteraan ini. Hasil kajian menunjukkan bahawa DRGP (ISE) adalah kaedah yang sangat baik and pantas untuk mengesan titik tuasan tinggi dalam model linear bermasalah ralat berautokorelasi.

Titik tuasan tinggi mempunyai kesan yang amat besar dalam analisis regresi. Dalam kajian ini kita mengesahkan bahawa titik tuasan tinggi adalah penyebab lain autokorelasi.

Tidak banyak penyelidikan telah dijalankan untuk menyiasat cerapan autokorelasi-berpengaruh. Dengan itu, Pengukur Autokorerlasi-Berpengaruh Teguh berdasarkan DRGP (RAIM(DRGP)) dibina bagi mengenal pasti cerapan autokorelasi berpengaruh dalam data berautokorelasi. Didapati RAIM(DRGP) telah menunjukkan kecemerlangan untuk mengenal pasti cerapan tuasan tinggi autokorelasi-meningkat dan autokorelasi-menurun dan cerapan autokorelasi berpengaruh.

Kaedah lelaran Cochrane-Orcutt Prais-Winsten (COPW) adalah langkah pemulihan yang paling biasa digunakan untuk membetulkan masalah autokorelasi. Walau bagaimanapun, prosedur ini sangat terdedah dengan kehadiran titik tuasan tinggi.

Pada masa yang sama, masalah autokorelasi mungkin disebabkan oleh kehadiran titik tuasan tinggi. Dengan itu, Kaedah lelaran Cochrane-Orcutt Prais-Winsten Teguh (RCOPW) dicadangkan. Keputusan kajian menunjukkan anggaran RCOPW langkah pembaikan yang teguh dalam kes data berautokorelasi bersama dengan kehadiran titik tuasan tinggi dan masalah autokorelasi disebabkan oleh titik tuasan tinggi.

Plot diagnostik sedia ada tidak mengambil kira masalah autokerelasi. Dengan itu, kaedah pemulihan teguh masalah autokorelasi - Cochrane-Orcutt Prais-Winsten Teguh (RCOPW) digabungkan ke dalam plot diagnostic untuk membentuk Plot Diagnostik Untuk Autokorelasi Berdasarkan Piawaian Sisa Cochrane-Orcutt Prais-Winsten (DPA-RCOPW). Keputusan berdasarkan simulasi dan data terpencil terkenal menunjukkan DPA-RCOPW berjaya mengenal pasti dan mengklasifikasikan cerapan terpencil mengikut jenisnya dengan tepat.

Dalam kajian ini, kaedah alternatif mencari selang keyakinan parameter regresi bagi data bermasalah autokorelasi dengan kehadiran titik tuasan tinggi dan autokorelasi yang disebabkan oleh titik tuasan tinggi dicadangkan. Hasil kajian menunjukkan Diagnostik Sebelum Bootstrap berdasarkan anggaran Cochrane-Orcutt Prais-Winsten Teguh (DBB RCOPW) adalah prosedur yang mantap dan konsisten dalam memberikan jawapan dekat dengan selang keyakinan sebenar parameter regresi untuk data ralat berautokorelasi dengan kehadiran titik tuasan tinggi dan autokorelasi disebabkan oleh titik tuasan tinggi.

ACKNOWLEDGEMENTS

First and foremost, I would like to give thanks to my God, who have provided me His strength and grace to throughout my doctoral pursue.

Heartfelt appreciation also goes to my committee chairperson, Prof. Dr. Habshah Midi for her constant inspiration, efficient guidance, and constructive feedback rendered. I am deeply honoured to have the opportunity to complete my degree under her supervision.

I would also like to thank my internal co-supervisors, Assoc. Prof. Dr. Jayanthi Arasan and Dr. Md. Sohel Rana for all their guidance provided and also knowledge shared.

A special word of thanks to Professor Dr. Rahmatullah Imon, who is a professor of statistics from the Department of Mathematical Sciences, Ball State University, U.S.A, for his valuable time in sharing with me some of the insightful ideas during his visit to my campus.

I would like to extend my gratitude to all the wonderful people such as Mazlina, Shafie, Balqish and others. Their presence have indeed enriched my journey in completing my doctoral pursue. Appreciation also extended to all members of Graduate School and Faculty of Science, who have worked hard in creating a conducive environment for all post graduate students. I am glad to be able to graduate from this institution.

My sincere regards to all my siblings, who have continuously encourage me not to loose heart in all that I am pursuing, both mentally and also spiritually.

Last but not least, my special thanks go to my beloved wife, for standing by with me patiently with her never ending encouragement, prayers and support throughout my doctoral pursue. My newborn son, Isaac has also been very cooperative and well behaved for allowing me to complete my dissertation. Thank you.

APPROVAL

I certify that a Thesis Examination Committee has met on 26 August 2014 to conduct the final examination of Lim Hock Ann on his thesis entitled "Robust Estimation Technique and Robust Autocorrelation Diagnostic for Multiple Linear Regression Model with Autocorrelated Errors" in accordance with 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.

Member of the Thesis Examination Committee were as follows:

Noor Akma binti Ibrahim, PhD
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)

Muhammad Hanif Mian, PhD
Professor
Lahore University of Management Science
Pakistan
(External Examiner)

NORITAH OMAR, Ph.D.
Associate Professor and Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 19 September 2014

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

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

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

Md. Sohel Rana, PhD
Senior Lecturer
Faculty of Science
Universiti Putra Malaysia
(Member)

BUJANG BIN KIM HUAT, Ph.D.
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

DECLARATION

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 other 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.:____

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) are adhered to.

Signature: ______________
Name of Chairman of Supervisory Committee: _______________

Signature: ______________
Name of Member of Supervisory Committee: _______________

Signature: ______________
Name of Member of Supervisory Committee: _______________

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

OLS	Ordinary Least Squares
BLUE	Best Linear Unbiased Estimators
MSE	Mean Square Errors
SSE	Sum of Squares Errors
SSR	Sum of Squares Regression
MVUE	Minimum Variance Unbiased Estimator
IF	Influence Function
MVE	Minimum Volume Ellipsoid
MCD	Minimum Covariance Determinant
CME	Covariance Matrix Equality
ISE	Index Set Equality
LAV	Least Absolute Values
MSAE	Minimum Sum of Absolute Errors
LAR	Least Absolute Residuals
LAD	Least Absolute Deviations
LMS	Least Median of Squares
LTS	Least Trimmed Squares
MAD	Median Absolute Deviation
MADN	Normalized Median Absolute Deviation
WLS	Weighted Least Squares
IWLS	Iterative Weighted Least Squares
RLS	Reweighted Least Squares
IRLS	Iteratively Reweighted Least Squares
SD	Standard Deviation
MD	Mahalanobis Distance
RMD	Robust Mahalanobis Distance
RMD (MVE)	Robust Mahalanobis Distance based on the Minimum Volume Ellipsoid
ASE	Asymptotic Standard Error
CLRM	Classical Linear Regression Model
LM	Lagrange Multiplier
MA	Moving Average
AR	Autoregressive
BG	Breusch-Godfrey
MBG	Modified Breusch-Godfrey
LM	Lagrange Multiplier
CDF	Cumulative Density Function
PDF	Probability Density Function
GLS	Generalized Least Square
COPW	Cochrane-Orcutt Prais-Winsten
RCOPW	Robust Concrane-Orcutt Prais-Winsten
GP	Generalized Potentials

DRGP (MVE)	Diagnostic Robust Generalized Potential based on Minimum Volume Ellipsoid
DRGP (ISE)	Diagnostic Robust Generalized Potential based on Index Set Equality
RAIM(DRGP)	Robust Autocorrelation-Influential Measure based on DRGP(ISE)
AEIO	Autocorrelation Enhancing-Influential Observations
ARIO	Autocorrelation Reducing-Influential Observations
CAIM	Classical Autocorrelation-Influential Measure
RDP	Robust Diagnostic Plot
DPA-COPW	Diagnostic Plot for Autocorrelation Based on Standardized Cochrane-Orcutt Prais-Winsten Residuals
DPA-RCOPW	Diagnostic Plot for Autocorrelation Based on Standardized Robust Cochrane-Orcutt Prais-Winsten Residuals
DBB	Diagnostic Before Bootstrap
DBB OLS	Diagnostic Before Bootstrap based on OLS
DBB COPW	Diagnostic Before Bootstrap based on COPW
DBB RCOPW	Diagnostic Before Bootstrap based on RCOPW

CHAPTER 1

INTRODUCTION

1.1 Introduction and Background of the Study

Linear regression is widely used in all areas of human efforts. It was the primary regression analysis to be studied rigorously. Modeling and analysis using linear regression is comparatively easier than non-linear regression as the properties of parameters estimate is easier to be determined in linear regression. It has become a traditional practice to regress linear regression models using the predominant Ordinary Least Squares (OLS) estimator. The reason for the universally acceptance of OLS is because of its computational simplicity. However, the OLS estimate has its optimum properties only when all the underlying model assumptions are met. Unfortunately, in reality the assumption of random and uncorrelated errors is always violated. The classical model assumes that the error term relating to any observation is not influenced by the error term relating to any other observation. However, the errors might be correlated with the previous errors which means that $E(u_i, u_j) \neq 0$ or $\text{cov}(u_i, u_j) = 0$ for $i \neq j$. Although autocorrelated errors do not cause any biasness in the OLS coefficients estimates, but the OLS coefficients estimates become less efficient in the presence of autocorrelated errors. The standard errors of the parameters estimate tend to be underestimated and this lead to misleading conclusion about the statistical significance of the estimated regression coefficients.

On the other hand, the OLS estimate which minimizes the sum of squared between the fitted values and the observed responses in the dataset is obviously affected by high leverage points. Research done by Harter (1974) confirmed that squaring of the residual causes the least square becomes extremely weak to the presence of high leverage points. Thus, it caused the violation from the least squares assumption. At the same time, routine dataset cannot be guaranteed free from outlying observations such as outliers and high leverage points. It is a necessity to introduce the robust methods in linear regression to address both autocorrelation and high leverage points problems.

1.2 Importance and Motivation of the Study

Autocorrelation violates the important properties of the OLS estimates (White and Brisbon, 1980). The parameters estimates obtained by the OLS estimation procedure no longer the Best Linear Unbiased Estimators (BLUE) in the sense that we are able to obtain the parameters estimate with lower standard errors. As the result, the usual t and F test of significance are no longer convincing as the tests tend to be statistically significant when in fact it is not. In addition, the coefficient of determination, R^2 becomes inflated, the estimators would look more accurate as compared to its actual values. In short, the existence of autocorrelated errors will most likely causing the wrong conclusions about the statistical significance of the estimated regression coefficients.

(Gujarati and Porter, 2009). Therefore, detection of autocorrelation problems is very critical. Breusch-Godfrey (BG) test (Breusch, 1978; Godfrey, 1978) is the most general test to detect the presence of autocorrelated errors in economics. However, this test is based on OLS estimate which is not robust, the poor performance of BG test is anticipated in the presence of high leverage points. High leverage points may be defined as the data points which are bulky different from the rest of the data points in X-direction. Many robust literatures have pointed out that high leverage points have great impact on the OLS estimates. (Habshah et al., 2009; Rana et al., 2008; Norazan, 2008).This motivates us to develop a robust autocorrelation detection method which shall perform equally good as BG test for detecting the autocorrelation problems in clean time series and cross sectional datasets. At the same time, it can detect the autocorrelation problems in the contaminated high leverage time series and cross sectional datasets. This is certainly a first attempt in statistics to develop robust autocorrelation detection technique which is resisting of the influence of high leverage points.

When the OLS estimate is applied for fitting the linear regression line, the resulting residuals are function of the leverages and true errors. The masking effect occurs when the high leverage points pull the fitted regression line in a way that the fitted residuals corresponding to that high leverage points. Similarly, the swamping effect happens when the residuals corresponding to inliers are too large to cause the case to be declared as high leverage cases. Péna and Yohai (1995) pointed out that high leverage points are the cause of masking and swamping of data points in linear regression. Therefore, identifying high leverage points in the data is very essential before any inferential is made. Although much works have been done on the identification of high leverage points in linear regression such as leverages method, Hadi's Potential (Hadi, 1992), Mahalanobis Distance (Mahalanobis, 1936) and Diagnostic-Robust Generalized Potentials (Habshah et al., 2009) but no specific method was proposed to identify the high leverage points in linear regression with autocorrelated errors. In this thesis, we would like to take up the challenge to find out the most reliable approach in identifying high leverage points in linear regression with autocorrelated errors.

The recent researches done by Bagheri et al. (2012) and Riazoshams et al. (2010) have further confirmed that high leverage points have tremendous effect on the OLS estimates. However, the effect of high leverage points in data with autocorrelated errors has not been fully discussed. No study is done to justify the autocorrelation in time series and cross sectional data is due to the presence of high leverage points. This literature gap motivates us to go a step further to verify that the high leverage points are the cause of autocorrelation in time series and cross sectional data.

Bagheri et al. (2012) proposed a novel method for collinearity-influential observation diagnostic measure based on group deletion approach to measure the contribution of each observation towards the collinearity in the dataset. However, to the best of our knowledge, no research has been done to study the autocorrelation-influential observations diagnostic measures in linear model. The existing diagnostic measure only focused on time series model where the observations are viewed in the time domain. An observation is omitted and the resultant effect on the interested statistic values is noted. Observations which give relatively large changes in the calculated values are deemed to be the influential observations. This diagnostic measure can only be applied to time series model. Since no diagnostic measure has been proposed to evaluate the autocorrelation-influential observations in linear model, in this thesis we take the initiative to develop a novel robust diagnostic measure for identification of autocorrelation-influential observations in linear model to close the gap in the literature. On the other hand, high leverage points are discovered as a new source of autocorrelation, it may be considered to be a special case of the autocorrelation–enhancing influential observations. It is reasonable to conclude that autocorrelation-influential measure which observes the influential effect of an observation at a time may not be efficient in the presence of high leverage points as high leverage points have unduly effect on the classical estimates. In addition, an autocorrelated dataset may change its nature to a non-autocorrelated dataset in the presence of high leverage points. To our knowledge, nothing has yet been done to diagnose autocorrelation reducing-influential points. It is also interesting to find out whether all the autocorrelation-influential observations are caused by high leverage points and also whether all the high leverage points in the autocorrelated data are the high leverage autocorrelation-influential observations. These further encouraged us to develop a novel robust diagnostic measure for identification of autocorrelation-enhancing and reducing-influential observations for linear model with autocorrelated errors in the presence of high leverage points.

This thesis also addresses the parameter estimation of linear model with autocorrelated errors. A large number of novel works in the literatures about the parameter estimation of linear model with autocorrelated errors. Cochrane-Orcutt Prais-Winsten iterative method (COPW) iterative method (Prais and Winsten, 1954) is the most popular remedial measure in econometrics to obtain estimators with the optimum Best Linear Unbiased Estimators (BLUE) properties. However, the COPW iterative method is based on the OLS estimate which is expected to be easily affected by high leverage points. The shortcoming of COPW iterative procedure has inspired us to develop a robust parameter estimation method to get rid both the autocorrelation and high leverage points problems in the time series and cross sectional datasets. To the best of our knowledge, this is indeed the first attempt to remedy the autocorrelation problems in the presence of high leverage points. At the same time, we also examined the usefulness of this proposed robust parameter estimation in rectifying the autocorrelation caused by high leverage points. The proposed robust parameter estimation is indeed working well in rectifying both autocorrelation and high leverage points problems. This is also another new discovery in statistics to remedy the autocorrelation caused by high leverage points.

According to Hampel et al. (1986), a normal dataset usually contains about 1 to 10 percent outlying observations. There is no guarantee that the high quality data will be free from outlying observations. The outlying observations in univariate dataset with autocorrelated errors may be detected by visual inspection of scatter plot. However, the identification of outlying observations based on scatter diagram is not convincing enough. In addition, the graphical method does not work in high dimensional datasets.

Hubert et al. (2008) also pointed out that the outlying observations are more likely to occur in datasets with many variables. Thus, we need specify statistical method to identify the outlying observations. Many outlying observations detection methods are available in the literatures (Mishra, 2008; Maronna et al., 2006; Rocke and Woodruff, 1996; Kashyap and Maiyuran, 1993). However, not much studies have been carried out in classifying outlying observations according to its inference locations. Although Hubert et al. (2008) have proposed a robust diagnostic plot of classifying outlying observations. However, the method proposed does not take into the consideration of autocorrelated errors in time series and cross sectional data. The autocorrelation problems remain as it is without any concern. In the autocorrelated dataset, the residuals are correlated with the previous errors which means $E(u_i, u_j) \neq 0$ for $i \neq j$. An observation may be far from the bulk, but due to the autocorrelated errors, it may not really an outlying observation in the autocorrelated data. This inspires us to design a first ever exclusive diagnostic plot which incorporates the corrective action of autocorrelation to classify the outlying observations according to it types in the presence of autocorrelated errors in time series and cross sectional data. Since the outlying observations are presence in the dataset, the robust methods must be incorporated in the procedures of designing this comprehensive diagnostic plot.

Confidence interval is one of the favorite topics in linear regression analysis. It is used to indicate the reliability of an estimate. The classical confidence interval is constructed based on the sample finding. Thus, it is too obviously affected by the sample with unusual observations. At the same time, the distribution assumptions need to be made for the classical approach of finding the confidence interval. In contrast, bootstrap methods have a practical point that it does not require normality assumption of the parameters estimate. At the same time, it also enjoys the benefit of not requiring any theoretical calculations to estimate the standard errors of complicated model. This encourages us to find an alternative ways of finding confidence interval of regression parameters using bootstrap methods which do not subject to the statistical distribution requirement and applicable in unwell behaved dataset. The focus here is on the linear model with autocorrelation problems. We have seen that high leverage points have tremendous effect on the parameters estimate. The study here discusses the robust bootstrapping alternative approaches of finding the confidence intervals of regression parameters for data with autocorrelated errors in the presence of high leverage points. Autocorrelation may be due to the presence of high leverage points. Thus, in this study, some robust bootstrapping alternatives of finding the confidence intervals of regression parameters for autocorrelation due to the presence of high leverage in time series and cross sectional datasets are also examined.

1.3 Research Objectives

The main purpose of this thesis is to investigate the autocorrelation problems in linear regression model. Currently, the diagnostic and estimation methods dealing with autocorrelated errors are based on OLS estimates. Unfortunately, OLS estimate is easily effected by high leverage points. It will be a big success in statistics if we can have robust identification and estimation methods for autocorrelated data in the presence of high leverage points. Moreover, the autocorrelation may be caused by the presence of high leverage points. It will be interesting to have the autocorrelation correction measures to remedy the existence of autocorrelation because of the presence of high leverage points. Nevertheless, identification of autocorrelation influential observations is very essential in linear regression analysis. A comprehensive approach has yet to be developed to identify the autocorrelation influential observations in the presence of high leverage points. In addition, detection and classification of outlying observations is an interesting area in robust statistics. It would be great if we could have customised methods for identifying and classifying outlying observations in data with autocorrelated errors. Moreover, robust alternative approach of finding the confidence interval for regression coefficients in autocorrelated data is also an interesting area to be explored.

The main objectives of this research can be outlined systematically as follows:

1. To formulate a robust autocorrelation diagnostic method and to develop a reliable high leverage identification technique for linear model with autocorrelated errors in the presence of high leverage points.
2. To develop a diagnostic measure of autocorrelation influential observation which can successfully distinguish the autocorrelation-enhancing and autocorrelation-reducing observations for linear model with autocorrelated errors in the presence of high leverage points.
3. To develop a robust parameter estimation method of autocorrelated data in the presence of high leverage points and autocorrelation caused by high leverage points.
4. To construct a diagnostic plot which is able to identify and classify the outlying observations according to their inferential locations in data with autocorrelated errors.
5. To develop a robust bootstrapping alternative approach of finding the confidence intervals of the regression coefficients of autocorrelated data in the presence of high leverage points and autocorrelation caused by high leverage points.

1.4 Significance of Study

Linear regression is used extensively in many areas of studies such as business, engineering, education, medicine and social science. It has many practical applications. The foremost application of linear regression is to make a prediction of the dependence variable based of the fitted model. Linear regression models are often fitted using the OLS estimator. The OLS estimates have optimum properties if all the underlying model assumptions are met. Unfortunately, in reality the assumption of random and uncorrelated errors is always violated. On the other hand, the OLS estimates is not a robust estimates, it is easily effected by high leverage points. Many researchers are unaware of violation of autocorrelation and the effect of high leverage points on the linear regression parameters estimates. The robust autocorrelation diagnostic and estimation methods developed in this thesis are working well in good and contaminated autocorrelated data. Their excellence performances were verified by the assessments done by Monte Carlo simulation study together with some real time series and cross sectional datasets.

This research also pointed out that the high leverage points are the cause of autocorrelation problems. Therefore, the identification of high leverage points in linear regression is very crucial before any remedial action is taken. A credible diagnostic measure was also developed for identifying autocorrelation–influential observations in autocorrelated dataset in the presence of high leverage points. The diagnostic measure working excellently in detecting all the autocorrelation enhancing and reducing influential observations and other autocorrelation influential observations which are not the high leverage points.

In this research, a comprehensive diagnostic plot was also designed for the first time in statistics specifically to identify and classify the outlying observations according to their inferential location in autocorrelated data. The designed diagnostic plot performs superb in identifying and classifying the outlying observations according to their types in autocorrelated data.

Robust alternative approach of finding the confidence intervals of regression parameters in autocorrelated data was also proposed in this study. For all these discoveries, we expect there will be a good application for researchers and industry experts in the future.

1.5 Scope and Limitation of the Study

Robust statistics is still a new area in statistics. Thus, not many statistical software are equipped with robust statistics applications. For the existing software with robust statistics applications, the applications are not really diversified. Most of the time, there is no direct method to get the solution of the desired robustified method. Writing our own programming codes are most of the time required in this case. Although we may get the desired results, but we cannot guarantee that the programming codes are perfect without mistake.

Again, since the robust statistics is a newly developed field of statistics studies, not many well referred outlying datasets are available in the literature for discussion purpose. Not to mention that the outlying datasets with autocorrelation problems. Thus, the same datasets are used repeatedly in this thesis for difference objectives of study.

Alciaturi et al. (2005) proposed the use of the autocorrelation function with lag 1 residual in model selection. Following their suggestion, in this thesis we only focus autocorrelation problems at first-order autoregressive AR(1).

There are many existing robust estimators such as S-estimator, M-estimator, Least Median Squares estimator and etc. In this study, we concentrate only on MM-estimator because it is a bounded influence estimator has high breakdown point (50 percent) and high efficiency (approximately 95 percent) relative to the OLS under the Gauss-Markov assumptions.The MM-estimator is incorporated into the existing procedures in the formulation of robustified methods in the topics of the study.

1.6 Outline of the Thesis

In accordance with the objectives and the scope of the study, the contents of this thesis are organized in such a way that the research objectives are apparent and are conducted in the sequence outlined.

Chapter Two: This chapter presents a brief literature review of the OLS estimations of linear regression parameters and the violations from least squares assumptions. The review on autocorrelation problems and its consequences, diagnostic methods, remedial actions and the sources of autocorrelation problems are also discussed. Moreover, basic concepts of robust regression and some important existing robust regression methods are also highlighted. Diagnostic methods of outlying observations are also reviewed. Finally, bootstrapping methods are discussed briefly.

Chapter Three: This chapter presents the failure of autocorrelation diagnostic using the Breusch-Godfrey (BG) test developed by Breusch (1978) and Godfrey (1978) in the presence of high leverage points in time series and cross sectional data. The BG test is then robustified by incorporating the high efficient and high breakdown point MM-estimator (Yohai, 1978) in the BG test procedure. The merit of using the Modified Breusch-Godfrey (MBG) test is studied through Monte Carlo simulation, time series and cross sectional datasets.

Chapter Four: In this chapter we suggests the Diagnostic Robust Generalized Potential Based on Index Set Equality (DRGP(ISE)) for identifying high leverage points in linear regression with autocorrelated errors. The advantages of using this proposed method is supported by the evidence from the Montle Carlo simulation and real time series and cross sectional datasets.

Chapter Five: This chapter investigates high leverage observations as a cause of autocorrelation. Study through Monte Carlo simulation and some well-referred time series and cross sectional datasets were supported the finding that the existence of autocorrelation was due to the presence of high leverage points.

Chapter Six: In this chapter we propose to use the Robust Autocorrelation-Influential Measure based on DRGP (RAIM(DRGP)) to identify the autocorrelation-influential observations in autocorrelated data in the presence of high leverage points. The merit and the excellent performance of RAIM(DRGP) is assessed by using Monte Carlo simulation experiments and so well-known datasets.

Chapter Seven: This chapter deals with the development of robust parameters estimation to address the autocorrelation and high leverage points problems. Data with autocorrelated errors may be contaminated by the high leverage points. On the other hand, autocorrelation may be due to the presence of high leverage points. The Concrane-Orcutt Prais-Winsten (COPW) iterative method performs miserably in correcting autocorrelation problems in the presence of high leverage points in time series and cross sectional datasets. The Robust Concrane-Orcutt Prais-Winsten (RCOPW) iterative method is then proposed to remedy both autocorrelation and high leverage points problems. The performance of RCOPW procedure is evaluated by using Monte Carlo simulation experiments and real datasets.

Chapter Eight: This chapter discussed the disadvantages of Robust Diagnostic Plot (RDP) proposed by Hubert et al. (2008) in identifying and classifying the outlying observations in data with autocorrelated errors. In this chapter we designed a comprehensive plot which is able to identify and classify the outlying observations according to its inferential locations accurately for data with autocorrelated errors. The plot is called Diagnostic Plot for Autocorrelation Based on Standardized Robust Cochrane-Orcutt Prais-Winsten Residuals (DPA-RCOPW). It is a plot of Standardized Robust Residuals obtained by Robust Cochrane-Orcutt Pais-Winsten (RCOPW) iterative method versus the leverages computed from Diagnostic Robust Generalized Potentials based on Index Set Equality (DRGP(ISE)). The excellency of DPA-RCOPW is tested using Monte Carlo simulation and some famous robust statistics datasets.

Chapter Nine: This chapter introduced a robust alternative of finding confidence intervals of regression parameters for autocorrelation data in the presence of high leverage points and autocorrelation caused by high leverage points. The Diagnostic Before Bootstrap (DBB) is incorporated in the bootstrapping residuals based on Robust Cochran-Orcutt Prais-Winsten (RCOPW) procedure to form the DBB RCOPW confidence intervals. The DBB RCOPW confidence intervals constantly provide fairly close intervals to the benchmark confidence intervals for autocorrelation data in the presence of high leverage points and autocorrelation due to the presence of high leverage points.

Chapter Ten: This chapter provides summary and detailed discussions of the thesis conclusions. Some areas of future studies are also tabulated.

REFERENCES

Anwar, F. and Habshah, M.(2010). Diagnostic robust generalized potentials for identifying high leverage points in mediation analysis. *World Applied Sciences Journal*, 11(8):979-987.

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

Andrews, D. F. (1974). A robust method for multiple linear regression. *Technometrics*.16:523-531.

Alciaturi, C.E., Escobar, M.E., Estéves, I.(2005). The use of the autocorrelation function in modeling of multivariate data. *Analytica Chimica Acta*, 553: 134-140.

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

Barnett, V. and Lewis, T. (1994). *Outliers in Statistical Data*. 3rd edition. New York: Wiley.

Beaton, A.E. and Tukey, J.W. (1974). The fitting of power series, meaning polynomials, illustrated on band-spectroscopic data. *Technometrics*.16: 147-185.

Bernard, L. (1991). Detecting over-influential observations in time series. *Biometrika*. 78 (1) :91-99.

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

Bowerman, B. L. and O'Connell, R.T. (2003). *Business Statistics in Practice*. 2nd edition. New York: McGraw-Hill.

Breusch, T. S. (1978). Testing for autocorrelation in dynamic model. *Australian Economic Papers*. 17:334-355.

Chatterjee, S. and Hadi, A.S. (2006). *Regression Analysis by Example*. 4th edition. New York: Wiley.

Chernick, M. R., Downing, D. J. and Pike, D. H. (1982). Detecting outliers in time series data. *Journal of American Statistical Association.* 77: 743-7.

Choulakian, V., Lockhart, R.A. and Stephens, M.A. (1994). Cramer-von Mises statistics for discrete distributions. *The Canadian Journal of Statistics*. 22(1): 125-137.

Cochrane, D. and Orcutt, G.H. (1949). Application of least squares regression to relationships containing auto-correlated error terms. *Journal of American Statistical Association*. 44:32-61.

Croux, C., Rousseeuw, P. J. and Hössjer, O. (1994). Generalized S-estimators. *Journal of the American Statistical Association.* 89:1271–1281.

Draper, N. R. and Smith, H. (1998). *Applied Regression Analysis*. New York:Wiley.

Durbin. J. and Watson G. S. (1951). Testing for serial correlation in least squares regression II. *Biometrika*. 38:159-178.

Edgeworth, F. Y. (1887). On observations relating to several quantities. *Hermathena*. 6:279-285.

Efron, B. (1979), Bootstrap methods: Another look at the jackknife, Annals of Statistics, 7, 1-26.

Efron, B. (1987). Better bootstrap confidence intervals. *Journal of the American Statistical Association.* 82: 171–185.

Ellenberg, J.H. (1976). Testing for a single outlier from a general linear regression. *Biometrics*. 32: 637-645.

Geary, R.C. (1970). Relative efficiency of count sign changes of assessing residual autoregression in least squares regression. *Biometrika*. 57: 123-127.

Godfrey. L. G. (1978). Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables. *Econometrica*. 46: 1293-1301.

Grassian, A. and Boer, E.S. (1980). Some methods of growth curve fitting. *Math. Scientist*. 5: 91-103.

Greene, W. H. (2008). *Econometric analysis*. 6th edition. Upper Saddle River. New jersey: Prentice Hall.

Gujarati, D.N. and Porter, D.C. (2009). *Basic Econometrics*. 5th edition. New York : McGraw-Hill.

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

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

Hampel, F. R. (1974). The influence curve and its role in robust estimation. *Journal of the American Statistical Association*. 69: 383-393.

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

Harter, H.L. (1974). The method of least squares and some alternatives-Part II. *International Statistics Review*. 42: 235-264.

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.

Hill, R. W. and Holland, P. W. (1977). Two robust alternatives to robust regression. *Journal of the American Statistical Association.* 72: 828–833.

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

Hosking, J. R. M. (1980). The multivariate portmanteau statistics. *Journal of American Statistical Association*. 75: 602-608.

Hosking, J. R. M. (1981). Lagrange-multiplier tests of multivariate time series model. *Journal of the Royal Statistical Society*. B43: 261-262.

Huber, P. J. (1964). Robust estimation of location parameters. *Annals of Mathematical Statistics.* 35:73–101.

Huber, P. J. (1973). Robust regression: asymptotic, conjectures, and Monte Carlo. *The Annals of Statistics*. 1: 799-821.

Huber, P.J. (2004). *Robust Statistics*. New York: John Wiley & Sons.

Hubert, M., Rousseeuw, P.J. and Van Aelst, S. (2008). *High-breakdown robust multivariate methods*. Statistical Science, 23(1), 92-119.

Iman, R.L. (1994). *A Data-Based Approach to Statistics*. Belmont California : Duxbury Press.

Imon, A.H.M.R. (2002). Identifying multiple high leverage points in linear regression. *Journal of Statistical Studies*. *Special Volume in Honour of Professor Mir Masoom Ali*. 3: 207–218.

Imon, A.H.M.R. and Ali, M.M. (2005) Bootstrapping regression residual. J*ournal of Korean Data & Information Science Society*, 16(3). 665-682.

Imon, A.H.M.R. and Das, K.K. (2005). A comparative study on the estimation of regression errors by bootstrap techniques. *Pakistan Journal of Statistics*. 21(1):109-122.

Imon, A. H. M. R., and Khan, M. A. I. (2003). A solution to the problem of multicollinearity caused by the presence of multiple high leverage points. *International Journal of Staistical Sciences.* 2: 37-50.

Jaggia, S. and Kelly, A. (2008). Practical considerations when estimating in the presence of autocorrelation. *CS-BIGS*. 2(1): 21-27.

Johnson, R.A., Wichern, D.W. (2002). Applied Multivariate *Statistical Analysis*. 5th edition. Upper Saddle River: Prentice Hall.

Kashyap, R.L and Maiyuran, S. (1993). Robust regression and outlier set estimation using likelihood reasoning. *Electrical and Computer Engineering ECE Technical Report, TR-EE 93-8, Purdue University School of Electrical Engineering. http://doc.lib.purdue.edu/ecetr/33/*

Khan, M.J. (2006). Robust Linear Model Selection for High-Dimensional Datasets, Unpublished Ph.D. Thesis, The University of British Columbia, United Kingdom.

Kutner, M.H., Nachtsheim, C.J., Neter, J. and Li, W. (2005). *Applied Linear Regression Models*. 5th edition. New York: MacGraw-Hill.

Li, W. K. and Hui, Y. V. (1987). On the empirical influence functions of residual autocorrelations in time series. *Proc. Bus. Econ. Statist. Sect., Am. Statist. Assoc.*: 465-468.

Li, G. (1985). *Robust regression in Exploring Data Tables*, *Trends and Shapes*. New York: Wiley.

Lim, H.A. and Midi, H. (2011). The performance of robust modification of Breusch-Godfrey test in the presence of outliers. Pertanika, Journal of Science and Technology. (Accepted)

Lim, H.A. and Habshah, M. (2012a). Robust autocorrelation testing in multiple linear regression. *International Journal of Mathematics and Computers in Simulation*. 6(1): 119-126.

Lim, H.A. and Habshah, M. (2012b). The effect of high leverage points on non-autocorrelated data. *Fundamental Science Congress 2012, "Fundamental Sciences : Merging Science, Industry and Society"*. (24-25)

Maddala, G.S. (2001). *Introduction to Econometrics*, 3rd edition. New York: John Wiley and sons.

Mahalanobis, P.C. (1936). On the generalized distance in statistics. *Proceeding of the National Institute of Sciences of India*. 12: 49-55.

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

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

Maurice G., Kendall and William, R.B. (1971). *A Dictionary of Statistical Terms*. New York : Hafner Publishing Company.

McClave, P. and James, T. (2008). *Statistics for Business and Economics*. 10th edition. Upper Saddle River: Prentice Hall.

McClave, J.T., Benson, P.G. and Sincich, T. (2011). *Statistics for Business and Economics*. 11th edition. United Sates of America: Prentice Hall.

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

Mishra, S.K. (2008). A new method of robust linear regression analysis : some monte carlo experiments. *Munich Personal RePEc Archive*. Paper no. 9445.

Montgomery, D. C., Peck, E. A. and Viving, G.G. (2001). *Introduction to linear regression Analysis*. 3rd edition. New York: John Wiley and sons.

Murray, M. (2006). *Econometrics A Modern Introduction*. 2nd edition. United States of America: Pearson Addison-Wesley.

Mustafa, M. N. (2005). Overview of current road safety situation in Malaysia. *Highway planning Unit, Road Safety Section, Ministry of Works*, 5-9.

Newbold, P., Carlson, W.L. and Thorne, B. (2007).6th edition. *Statistics for Business and Economics*. New Jersey: Pearson Education.

Norazan, M.R. (2008). *Weighted Maximum Median Likelihood Estimation for Parameters in Multiple Linear Regression Model*, Unpublished Ph.D. Thesis, Universiti Putra Malaysia, Malaysia.

Norazan, M.R., Habshah, M. and Imon, A.H.M.R. (2009). Estimating regression coefficients using weighted bootstrap with probability. *WSEAS Transactions on Mathematics*. 8(7):362-371.

Peña, D. and Yohai, V.J. (1995). The detection of influential subsets in linear regression by using an influence matrix. *Journal of Royal Statistical Society*. B 57: 18–44.

Pison, G., Van Aelst, S. and Willems, G. (2002). Small sample correction for LTS and MCD. *Metrika*. 55:111-123.

Prais, S. J. and Winsten, C. B. (1954). Trend estimators and serial correlation. *Cowles Commission Discussion Paper*. *Statistics* (*No. 383*).

Preston, K. R. and Tipples, K. H. (1978). An Ultracentrifuge absorption method. *The American Association of Cereal Chemist*. 55(1):96-101.

Politis, D. N., & Romano, J. P. (1994). The stationary bootstrap. *Journal of the American Statistical Association*. 89(428):1303-1313.

Rahman, M., Pearson, L.M., and Heien, H.C. (2006). A modified Anderson-Darling test for uniformity. *Bulletin of the Malaysian Mathematical Sciences Society* (2). 29(1): 11-16.

Rana, M.S., Midi, H. and Imon, A. A. H. M. R. (2008). A robust modification of the Goldfeld-Quandt test for the detection of heteroscedasticity in the presence of outliers. *Journal of Mathematics and Statistics*. 4(4): 277-283.

Riazoshams, H., Midi, H. and Sharipov, O.(2010). The performance of robust two-stage estimator in nonlinear regression with autocorrelated error. *Communications in Statistics. Simulation and Computation*. 39(6): 1236-1253.

Rocke, D.M. and Woodruff, D.L (1996). Identification of outliers in multivariate data. *Journal of American Statistical Association*. 91: 1047-1061.

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.

Rousseeuw, P.J. and Leroy, A.M. (1987). *Robust Regression and Outlier Detection*. New York: Wiley.

Rousseeuw, P. J. (1983). Multivariate estimation with high breakdown point. *Mathematical Statistics and Applications*. Vol (B): 283-297.

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

Rousseeuw, P.J. (1985). Multivariate estimation with high breakdown point. *Mathematical and Statistical Applications*. B: 283-297.

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

Rousseeuw P.J. and Van Driessen, K. (1999). A fast algorithm for the minimum covariance determinant estimator. *Technometrics.* 41:212–223.

Rousseeuw, P. and Van Zomeren, B. (1990).Unmasking multivariate outliers and leverage points. *Journal of American Statistical Associations*. 85: 633-639.

Rousseeuw, P.J. and Van Zomeren, B.C. (2006). Computing LTS regression for large data sets. *Data Mining and Knowledge Discovery*. 12: 29-45. MR2225526.

Rousseeuw, P. J. and Yohai, V. (1984). Robust regression by means of S-estimators, Robust and Nonlinear Time series Analysis. *Lecture Notes in Statistics*. 26: 256-272.

Rupert, D. and Carrol, R.J. (1980). Trimmed least squares estimation in the linear model. *Journal of American Statistical Association*. 75: 828-838.

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

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.

Sharma, S.C. (1987). The effects of autocorrelation among errors on the consistency property of OLS estimator. *Journal of Statistical Computation and Simulation*. 28(1): 43-52.

Sharpe, N.R., De Veaux, R.D. and Velleman, P. F. (2010). *Business Statistics*. Person International Edition. Upper Saddle River, New Jersey.

Srikantan, K.S. (1961). Testing for the single outlier in a regression mode *Sankyā* Series A, 23: 251–260.

Stromberg, A. J. (1993). Computation of high-breakdown nonlinear regression parameters. *Journal of the American Statistical Association.* 88:237–244.

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.

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

White, G. C. and Brisbon, I. L. (1980). Estimation and comparison of parameters in stochastic growth model for barn owls. *Growth*. 44: 97-111.

Yohai, V.J. (1987). High breakdown point and high efficiency robust estimates for regression. *The Annals of Statistics*. 15: 642-656.