6-1997

Estimating IBNR Reserves with Robust Statistics

Daniel Cheung
Western Michigan University

ESTIMATING IBNR RESERVES
WITH ROBUST STATISTICS

by

Daniel Cheung

A Dissertation
Submitted to the
Faculty of The Graduate College
in partial fulfillment of the
requirement for the
Degree of Doctor of Philosophy
Department of Mathematics and Statistics

Western Michigan University
Kalamazoo, Michigan
June 1997

ESTIMATING IBNR RESERVES
WITH ROBUST STATISTICS

Daniel Cheung, Ph.D.

Western Michigan University, 1997

There is often a considerable time lag between an incurred of an accident, such as medical malpractice or product liability, and the time it is reported to the insurance company. These Incurred But Not Reported (IBNR) losses need to be predicted in order to determine the necessary loss reserves. Many actuarial methods have been developed for IBNR reserves estimation. However, none of the methods being used for loss reserving is robust to outliers, nor do they provide adequate statistical inferences to support the actuarial decisions. The rank-based method proposed in this thesis is robust to outliers, and provides statistical inference for testing hypotheses. This rank-based method also calculates the R^2 , an indicator of goodness of fit, and approximates the standard error for calculating the confidence interval of IBNR.

UMI Number: 9732879

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

ACKNOWLEDGMENTS

I would like to express my gratitude for all those who contributed to the completion of my dissertation. The following are just a few who helped me in finishing this dissertation. I must thank Dr. Joshua D. Naranjo, my dissertation advisor, for guiding me through each research step. Dr. Joseph W. McKean who was most helpful for his inspiration, his papers and his book on the subject of Robust Statistics. I also would like to thank the other members of my dissertation committee, Dr. Gerald L. Sievers, Dr. J. C. Wang, and Dr. Thomas J. Vidmar for their time and insight. I am very grateful to my good friends, Dr. Kenneth Kaminski and Mr. Todd Gruenhagen for their inputs from an actuarial point of view.

A very special thank you to my wife, Corjena, who encouraged me to go back to finish this degree after 11 years away from school. For her love, support, encouragement and standing by me throughout the completion of my final destination, I owe it all to she.

Daniel Cheung

TABLE OF CONTENTS

CHAPTER

Table of Contents -Continued

CHAPTER

APPENDICES

LIST OF TABLES

List of Tables - Continued

LIST OF FIGURES

CHAPTER I

LOSS RESERVING

1.1 Introduction

One of the major tasks that an actuary routinely performs is Loss Reserve Analysis. The objective of Loss Reserve Analysis is to estimate the financial liability of an insurance company. Insurance companies rely on the results of Loss Reserve Analysis to make important financial decisions such as investment, pricing, and corporate planning. The insurance commissioner of each state relies on the results of Loss Reserve Analysis to determine the financial strength of an insurance company. If an insurance company is found to be inadequately reserved, it is not allowed to continue selling insurance in that state. Insurance companies also need to submit their results of Loss Reserve Analysis to the state insurance commissioner's office in order to request any insurance price increase. Outside investors also use the results of Loss Reserve Analysis to determine the financial strength of an insurance company for investment purposes. In short, results of Loss Reserve Analysis are major instruments for making financial decisions concerning an insurance company.

Some claims tend to have a short reporting lag, i.e. the time between an accident's occurrence and the date reported to the insurance company. With the experiences of claim adjusters and with a large number of similar claims, it is relatively easy to estimate the reserves for reported claims. However, some claims tend to have much longer reporting lag, as long as 5 or more years. In fact, the reporting lags for claims such as medical malpractice or product liability can be longer than 10 years. Claims Incurred But Not Reported are generally referred to as IBNR claims in the actuarial profession. Estimating the loss reserves needed for the IBNR claims is always a great challenge to actuaries.

The most commonly used method for estimating IBNR claims reserve needed is called the Chain Ladder Method which is taking the averages of loss development patterns from the past to predict the future loss incurred development. The advantage of the chain ladder method is its simplicity of the estimation process and relative ease of interpreting its results. Anyone who can perform simple arithmetic can complete the IBNR reserve estimation.

While I was working as a loss reserve analyst at The St. Paul Companies, one of the largest casualty property insurance companies, the chain ladder method was the main loss reserving method the company relied on. While doing loss reserve analysis using the chain ladder method, I frequently experienced frustrations as follows:

One major problem of the Chain Ladder Method is that it heavily depends on the average of the loss development pattern from the past. Since it heavily depends on the average of the past it is not robust to outliers. If there is one very large outlier in the past history, it can possibly skew the prediction significantly. This method also assumes that the losses incurred for accident years are independent. It measures only the loss development pattern within each accident year. It does not evaluate any trends for loss incurred along the accident years. If there is a skewed loss development pattern caused by outliers and is multiplied to the accident year's latest incurred loss data, it can tremendously over or under predicts the IBNR reserves for that accident year.

Another deficiency of the Chain Ladder Method is it does not use any statistical procedures to predict the IBNR reserve needed. Results of the method can not be tested with any statistical tests. Without the values of dispersion, estimate of scale, standard errors, or R^2 analysts who use this method are not able to determine the reliability of the results nor can they do any hypotheses testing. All they can do is make comments like "it looks good" or "it doesn't look good.".

Since the chain ladder method relies on analysts to select loss development factors based on averages, the results of the method would depend on many human factors. Results of the loss reserve analysis tend to vary among analysts. Since results are not estimated statistically, even if it was analyzed by one analyst, the results could still be varied by human factors.

In addition, if the same loss reserve estimation for each line of business is needed to be done routinely every 3 or 6 months, why not develop a statistical routine to estimate the results automatically and allow the loss reserve analysts to monitor the estimation results only.

Over the years, other statistical methods have been developed to estimate IBNR reserves with different forms of regression analyses. One that is commonly known is called De Vylder's Least Square Method (1978). De Vylder's Least Square Method (DVLS) takes a triangular data matrix to estimate the loss incurred patterns in two directions, following the loss development years and the accident years, with least square estimations. This method requires iteration to find the unique solution. DVLS method generally converge and does not take many iterations to find the unique solution.

One deficiency in this method is it is not a robust procedure. Since it relies on least square estimation, if there is one very large outlier, it will skew both the loss development pattern and the ultimate loss incurred. In fact, this method is so sensitive to outlier that only one large outlier is sufficient to generate unacceptable results. In addition, this method is not able to test the results of the estimation. Hence, analysts are not able to determine the goodness of fit nor testing any hypotheses.

Since this method relies on least square estimation, it requires a relatively large data set to produce reliable results. Somehow it is not possible that an actuary will always be guaranteed to have a large data set to perform Loss Reserve Analysis. Least square estimation on small data sets are more effected by an outlier than larger data sets.

Throughout the years that I was doing loss reserve analysis for The St. Paul Companies, DVLS method had been used often to predict IBNR reserves needed. As for the reasons stated above, its results were generally acceptable when estimating IBNR reserves for loss data with stable distribution and free of outlier such as Workers' Compensation or Personal Property Damage. But when it was used to estimate IBNR reserves for loss data with unstable or long tail distributions such as Product Liability or Professional Liability, DVLS method did not generate acceptable results.

Other published statistical methods for Loss Reserve Analysis tend to have the same problems the DVLS method has. They are not robust to outliers, and require large data sets in order to produce reliable results. They also do not provide statistical inference for testing hypotheses.

For those line of businesses where the existing loss reserving methods are not able to accurately estimate the IBNR reserves, insurance companies can possibly over reserve or under reserve them. If one line of business is over reserved, consumers are over charged; if it is under reserved, consumers are not getting the proper protection. That is a lose/lose situation for consumers. This is why the state insurance commissioner requires insurance companies to set the reserve for each line of business adequately. However, if there is not a loss reserving method that helps the state insurance commissioner to monitor those line of businesses with unstable distributions or with outliers, insurance companies can easily take advantage of the situation. There were numerous reported incidences that some insurance companies were found guilty of over charging consumers for policies such as medical malpractice or professional liability and consequently were forced to return premiums back to their customers. Without a loss reserving method accurately predicts the loss development, those violations are usually don't get caught for many years. Even if a insurance company is forced to return premiums back to its customers, the insurance company can still benefit from the extra investment income from the extra loss reserves for few years. In another word, a robust statistical loss reserving method which accurately estimates the loss reserves for all lines of businesses can protect insurance companies from under or over reserve, and it can protect consumers from being over charged as well.

The objective of my doctoral thesis is to develop a Non-Parametric statistical procedure to estimate the IBNR reserve needed which is robust to outliers and unstable loss development patterns. The advantages for this proposed procedure are as follows:

1. The estimation is asymptotically distribution free.

2. It is robust to outliers and robust to unstable loss development patterns.

3. It does not require a large data set in order to produce reliable results.

4. It provides statistical inference for testing hypotheses. This means actuaries who use this procedure could determine if the estimates are statistically acceptable.

5. It calculates the R^2, an indicator of goodness of fit.

6. It approximates the standard error for IBNR and hence, a confidence interval for IBNR can be calculated.

7. If the estimates are tested to be statistically acceptable, they can be used in further actuarial prediction. Estimates can also be used in other loss reserving methods such as chain-ladder method or Bonheutter-Ferguson method.

The fundamental assumption for this proposed procedure is that the amount of claims incurred in a particular development year and a particular accident year is the product of two unknown factors. The two factors are the total amount of claims incurred for that particular accident year and the loss development factor for that particular development year. Let Y_{ij} be the incurred loss for the ith accident year and the jth development year. Let X_i be the total incurred loss for the ith accident year and P_j be the loss development factor for the jth development year. Then

$$Y_{ij} = X_i P_j \ . \qquad (1)$$

The unknown quantities of X_i and P_j are estimated using rank-based estimation. As it is stated above, the rank-based estimation is robust to outliers and robust to unstable loss development patterns. In addition, the rank-based estimation calculates the dispersion and scale values which can be used for testing hypotheses, goodness of fit, and estimating confidence interval for IBNR. A few hypotheses testings are suggested to determine the reliability of the estimates.

The results of this rank-based estimation will be compared with the results estimated by using the Chain Ladder, DVLS methods, and other methods for estimating IBNR reserves. The objectives of the comparison are:

1. The robust estimation of IBNR reserves will be very close to the estimations calculated by the Chain Ladder and DVLS methods if the data set has stable loss development pattern.

2. The estimation calculated by this robust procedure will be far superior compared to the estimations of other classical methods if there is outlier within the data set.

The performance of these loss reserving methods are measured based on the stability of the estimation for various types of data sets, data sets with stable loss development pattern as well as data sets contain outlier and contaminated loss incurred data.

Casualty Actuarial Society sponsors a fall conference each year called Loss Reserve Seminar. Actuaries throughout the North America will get together to discuss any issues related to loss reserving. The theme for this year's conference is Measuring the Performance of Reserving Methods. This thesis has been submitted to the conference committee and was chosen to be presented at the conference in September.

1.2 Loss Reserve Analysis

One of the major responsibilities for a casualty actuary is to estimate the provisions of financial liabilities for an insurance company to its policyholders. All insurance companies are required to set up adequate reserves to pay for claims which have been incurred and reported as well as those which have been incurred but not yet been reported. The main purpose of these reserves is to ensure the protection of insured so that if a claim is filed to the insurer, there will be sufficient funds to pay the claimants despite of a first or third party claims.

The insurance commissioner of each state requires that each insurance company which does business in that state to provide a financial statements each year to prove that the insurance company has adequate reserves. The financial statement which is filed with the state insurance commissioner's office usually need to be supported by a certified external auditor to ensure the insurance company has adequate reserves, not under or over reserves.

It is neither difficult nor is it the actuary's responsibility to estimate the financial liability of a claim incurred and reported to the insurance company. If a claim is reported to the insurance company, a claim adjuster will estimate the total liability for that claim when the claim is closed. A claim can not be closed until all financial responsibilities have been fulfilled. That is why some workers' compensation claims or product liability claims remain open for many years. The summation of the estimates for all claims incurred in one year (Accident Year) and are reported is called *total case reserves*. Since case reserves are estimated case by case and the facts for each individual claim are known to the claim adjuster, it is not difficult to determine a relative adequate reserves for all claims reported. The difficult part and this is the responsibility of an actuary to estimate the adequate reserves for those claims which have been Incurred But Not Reported (IBNR). Since the IBNR reserves are not estimated case by case, it is also referred to as the non-case reserves.

Loss reserving is the term used to denote the actuarial procedures of estimating the amount of case and non-case loss reserves.

1.3 Importance of Accurate Loss Reserves

It was mentioned above that all insurance companies are required to show evidence of adequate loss reserves, not over or under reserves. If an insurance company is found to be under reserve, that means the financial strength of this company is questionable and it may not have enough reserves to pay for all the claims reported and the claims incurred but not reported. Insolvent insurance companies are not be allowed to continue to sell insurance policies to the public. Because it does not have the financial strength to keep its contractual obligations to its policyholders.

In a simplified term, the profit of an insurance company is calculated as the total income minus the total expenses. Total expenses includes the loss reserves. Under estimating the loss reserves leads to under estimating the expenses. This means its profit earning could possibly be over stated. Insurance companies share their profits by paying dividends to their shareholders if they are stock owned companies, or to their policyholders if they are Mutual companies (A mutual insurance company is owned by its policyholders). An under reserve insurance company tends to over pay its shareholders or policyholders with dividends. Over paying of dividends would further threaten the solvency of the insurer.

On the other hand, if an insurance company is found to be over reserved, its profit earnings are possibly under stated. This means its shareholder or policyholders are under paid with dividends. One of the major factor that pricing actuaries use to determine the insurance premium is the profit earnings for that particular line of business. If the profit earning for that particular business is under stated, the insurance premium for that business could possibly be over charged. That is the reason why an insurance company that wants to file for rate increase at the state insurance commissioner's office, has to submit along with the loss reserves information for that line of business.

A major part of income for an insurance company comes from investment income. Premium received from policyholders is not simply put in bank waiting to pay losses. Portion of the premium received is reserved for operating expenses and claims reported. The remaining premium received will be placed in different types of investments. Some investments are short term and some are long term. Some high return investments, such as real estate, do not allow assets to be liquidated for a long period of time. Investment department for an insurance company depends heavily on a loss pay out schedule estimated by the actuarial department. Under or over estimating the ultimate loss reserves can distort the potential investment income for an insurance company. In order for the investment department to manage the company's financial portfolio, actuary needs to predict the loss paid out schedule for each line of business accurately.

1.4 IBNR Reserves

Suppose a surgeon performed an procedure for his patient in June of 1990. His patient died in July of 1993 and it was determined that the cause of death was the surgeon's negligence in that particular procedure. This accident was reported to the insurer in September of 1993. Though this claim was unknown to the insurer at year end 1990, according to the principle of actuarial accounting, the financial liability of this claim should have been recognized in the 1990 financial statement, the year that this accident incurred. The estimated amount of liability for all these incurred claims which are not reported to insurer as of the date of financial statement is called the IBNR reserves.

Additional to estimating the non-case reserves, actuaries also need to estimate the change of case reserves. Some reported claims could possibly be over reserved due to unexpected early claim closing, salvages, or subrogation. On the other hand, some reported claims could also be under reserved due to unforeseeable legal liabilities. To better illustrate the robust procedures for estimating IBNR reserves, throughout this thesis, IBNR reserves means non-case reserves plus the change of case reserves.

As the simple illustration above has shown, it is not easy to determine the adequate IBNR reserves because the facts and figures for those claims are unknown to the insurer. Various methods have been used by actuaries to estimate the IBNR reserves. In general

IBNR reserves are estimated based on historical claims paid out amounts and their paid out pattern. The most commonly used method is called Chain Ladder Method which is based on averages of loss development factors. Other methods are based on statistical analyses such as regression analyses, time series, credibility theory, and compound Poisson distributions.

IBNR reserves are estimated by different accounting periods such as accident year, report year, calendar year, policy year, or fiscal year. Most insurance companies analyze their loss reserves annually, but some would analyze their reserves quarterly, or semiannually. Loss data can be categorized into 3 types: direct, net, and ceded losses. Losses are analyzed either gross or net of salvages and subrogation. Actuary group loss data of different lines of businesses with similar loss development pattern into a larger data set to increase the credibility of the analysis. These groups are typically referred to as: (a) Medical Malpractice, (b) Professional Liability, (c) Workers' Compensation, (d) Bonds, (e) Personal Liability, (f) Commercial Liability, (g) Ocean Marine, (h) Inland Marine, (i) Property Damage, (j) Excess, and (k) Reinsurance.

Table 1 displays a typical claims data set which represents the liability claims incurred over a five-year period, 1990-1994 in units of thousands of dollars. Similar data sets are used by actuaries to estimate the IBNR reserves. Throughout this thesis, the data set of Table 1 will be used to illustrate the various methods presented.

Table 1 indicates that $250,000 incurred losses (paid losses plus case reserves) corresponding to accidents incurred and reported between 1/1/1990 and 12/31/1990.

Table 1

Cumulative Incurred Losses

Loss Development Year	Accident Year				
	1990	1991	1992	1993	1994
1	250	267	298	289	300
2	550	582	642	601	
3	667	702	766		
4	717	757			
5	733				

$550,000 incurred losses for the second loss development year for accident year 1990 represented the cumulative paid losses plus case reserves for accidents incurred during 1990 and reported before year end 1991. As the data set above indicates, loss development gradually decreases in the fourth and fifth years. For some insurance liabilities such as auto liability or property damage, the loss development would last for only 3 or 4 years. However, other insurance liabilities could possibly have further loss development incurred beyond the fifth loss development year. In fact, the loss developments for some insurance liabilities such as workers' compensation, product liability, and professional liability could extend well beyond 10 or 15 years. It is very important to have data sets which cover the entire loss development for that particular line of business.

The triangular shape of the data set indicated that the loss incurred data was collected at year end 1994 and it was assumed there were no further loss information was available beyond 1994. Therefore, for accident year 1994, only one, the first loss development year incurred, is recorded. And for accident year 1993, 2 loss development years, the first loss development year which is 1993 and the second loss development year which is 1994, are recorded. Since the recorded loss incurred represents the total loss development for each accident year, the loss incurred data increased along with the loss development year. Without loss of generality, it is assumed there are only five years of loss development for this particular insurance liability.

Table 2 displays the incremental loss incurred for the data set displayed on Table 1. Loss data for each cell represents the losses paid plus the change of case reserves within that particular loss development year for one particular accident year. It is assumed that the error for each of the incremental loss incurred data is identically and independently distributed.

The actuarial work of estimating IBNR is to estimate each of the lower triangular cells accurately based on the recorded loss incurred activities. For accident year 1994, there is only one loss incurred data recorded which makes the predicted IBNR reserves for the following loss development years very sensitive to this loss amount. It is assumed that the cumulative or incremental development pattern for paid or incurred losses are stable across the accident years. It is also assumed that the paid and incurred losses grow in a stable pattern along the accident years due to inflation and the growth of business. This means that each of the lower triangular cells, for example, the second loss development year for accident year 1994, can be estimated based on the previous accident year's loss development patterns and the growth experience for this business along the accident years.

Table 2

Incremental Incurred Losses

Loss Development Year	Accident Year				
	1990	1991	1992	1993	1994
1	250	267	298	289	300
2	300	315	344	312	
3	117	120	124		
4	50	55			
5	16				

CHAPTER II

TRADITIONAL METHODS OF IBNR ESTIMATION

2.1 The Chain Ladder Method

Many actuarial methods have been published for estimating IBNR reserves. Among those methods, the chain ladder method is the most commonly used in the actuarial profession. It is based on the assumption that loss development patterns for all accident years are stable. Age-to-age loss development factor for i to i+1 development years is the ratio of the cumulative loss incurred for the i+1 loss development year to the cumulative loss incurred for the i development year. An age-to-age loss development factor for each i to i+1 development years are selected to predict future loss incurred development. The selected age-to-age loss development factors are picked based on the averages or weighted averages of loss development factors across the accident years. Predicted incurred are calculated by multiplying the selected age-to-age loss development factors to the latest loss incurred for each of the accident year.

2.2 Numerical Illustration for the Chain Ladder Method

Chain ladder method can be easily illustrated by using the data set displayed in Chapter I.

Table 3

Cumulative Loss Incurred

Loss Development Year	Accident Year				
	1990	1991	1992	1993	1994
1	250	267	298	289	300
2	550	582	642	601	
3	667	702	766		
4	717	757			
5	733				

Table 3 displays the losses incurred between accident years 1990 and 1994. Age-to-age loss development factors are then calculated based on the data set displayed above. The age-to-age loss development factor for loss development year 1 to loss development year 2 for accident year 1990 for instance is calculated at 2.20 which is the ratio of 550 to 250. Age-to-age loss development factors for the other accident years are calculated in the same manner.

Table 4 displays age-to-age loss development factors for the data set displayed on Table 3. Table 5 displays the averages and the weighted averages for age-to-age loss development factors across the accident years. Selected age-to-age loss development factors are then picked to calculate the future incurred losses for each accident year. The most recent incurred loss amount for each accident year is used to predict the future incurred loss by multiplying it to the selected age-to-age loss development factors.

Table 4

Age-to-age Loss Development Factors

Loss Development Year	Accident Year				
	1990	1991	1992	1993	1994
1 - 2	2.200	2.180	2.154	2.080	
2 - 3	1.213	1.206	1.193		
3 - 4	1.075	1.078			
4 - 5	1.022				

Table 5

Selected Age-to-age Loss Development Factors

Loss Development Year	Loss Development Factors			
	Average	Weighted Average		Selected
1 - 2	2.153	2.151		2.152
2 - 3	1.204	1.203		1.204
3 - 4	1.077	1.077		1.077
4 - 5	1.022	1.022		1.022

As the selected age-to-age loss development factors show, picked with reference to the average and weighted average of recorded incurred loss development, it is believed that the incurred loss grows 115.2% from the development year 1 to development year 2. And continues to grow 20.4% from development year 2 to development year 3. For accident year 1994, the current incurred loss recorded is $300,000. If the current incurred loss applies to the selected age-to-age loss development factor, the predicted incurred loss for development year 2 for accident year 1994 calculates as 300 x 2.152 = 645.6. The predicted incurred loss for development year 3 calculates as 645.6 x 1.204 = 777.3. The same calculation process continued will reflect an 856 incurred loss for development year 5 for accident year 1994. It is assumed that there is no more development after the 5th loss development year. This implies the predicted IBNR reserve for accident year 1994 is 856 - 300 = 556. Predicted IBNR reserves are estimated with the same process for all other accident years prior to 1994.

Table 6

Predicted Cumulative Loss Incurred

Loss Development Year	Accident Year					
	1990	1991	1992	1993	1994	Total
1					300	
2				601	646	
3			766	724	777	
4		757	825	779	837	
5	733	774	843	796	856	
Total IBNR Reserves	0	17	77	195	556	845

Table 6 displays the predicted loss incurred calculated based on chain ladder method. Table 6 also displays the predicted IBNR reserves for all accident years. IBNR reserves may also be calculated for individual development years of each accident year to estimate the IBNR reserves needed for future calendar years.

Table 7 displays the predicted IBNR reserves for each particular accident year which is then used to determine if this insurance company is adequately reserved. The total IBNR reserves estimated using the chain ladder method for the 5 accident years is 845. Table 7 also displays the IBNR emerge schedule for each future calendar year. As the section Importance of Accurate Loss Reserve in Chapter I stated, investment department requires information similar to that displayed in Table 7 in order to make proper investment decisions. Therefore, an accurately predicted Table 7 can improve the potential income of an insurance company.

Table 7

Calendar Year IBNR Reserves

Calendar Year	Accident Year					
	1990	1991	1992	1993	1994	Total
1995	0	17	59	123	346	545
1996	0	0	18	55	131	204
1997	0	0	0	17	60	77
1998	0	0	0	0	19	19
Total IBNR Reserves	0	17	77	195	556	845

As the illustration has shown, the basic objective of IBNR reserve estimation is to predict the lower triangular loss incurred data based on the upper triangular loss incurred data.

2.3 Deficiencies of the Chain Ladder Method

The primary reason that the chain ladder method is the most commonly used method for IBNR reserve estimation is its relatively simple calculation method and easy comprehension. It simply predicts IBNR reserves based on ratios of the incurred loss for one development year to the incurred loss from a previous development year. Age-to-age loss development factors which are used to predict future incurred losses are selected from either the average or weighted average of the loss incurred ratios from one development year to its previous year. This method does not require sophisticated computer software to do the estimation. Nor does it need any complicated statistical procedures to predict IBNR reserves. Actuaries can usually accomplish IBNR estimation by using a personal computer equipped with spreadsheet software.

Despite its simplicity, the chain ladder method has two major problems. The first is the age-to-age loss development factors that are used to predict future incurred losses are selected based on average or weighted average of loss development ratios. This means if there is an outlier, a very large loss year for one accident year, it would generate an extremely large loss development ratio. The average loss development ratio for that development year could possibly be extraordinarily large. If an exceptional large age-to-age loss development factor is selected, the predicted incurred losses for that particular development year for all accident years will be over estimated.

Breakdown point is defined to be the maximum proportion, ϵ, of gross outlier contamination that a data set can tolerate without breaking down the estimating procedure. Since the chain ladder method heavily depends on averages or weighted averages, it takes only one outlier, a data point close to infinity to break down the method. This implies that the breakdown point for the chain ladder method is equal to 0.

The second deficiency this method has is it does not have any statistical properties. It does not measure the variance for the loss development ratios between accident years. This means it is not possible to perform any statistical tests in evaluating the reliability of the predictions using this method.

In addition, selected age-to-age loss development factors are picked based on averages, it does not reflect any loss incurred trends along the accident years. For instance, due to advanced technical claim handling using latest electronic aides such as portable personal computers, claims incurred in recent accident years can be paid out or closed earlier. This would lead to larger losses incurred for the first and second development years and smaller loss development for the cumulative incurred in later development years. This means loss development ratios for the first and second development years should decrease following the accident years. If the average loss development ratios are selected as the age-to-age loss development factors for predicting the future loss incurred, they would be too large in relation to the loss development ratio trends. Application of over exaggerated loss development factors to the initial loss incurred for the most recent accident year, which is already larger than prior accident years' due to the increasing trends, the predicted IBNR reserve for the most recent accident year will consequently be excessively too large.

2.4 De Vylder's Least Squares Method

The fundamental assumption of the De Vylder's Least Squares Method (DVLS), Estimation of IBNR Claims by Least Squares (1978), is that the amount of claims incurred in a particular development year in a particular accident year is the product of two unknown factors. The two factors are the total amount of claims incurred for that particular accident year and the loss development factor for that particular development year. Let Y_{ij} be the incurred loss for the ith accident year and jth development year. Let X_i be the total incurred loss for the ith accident year and P_j be the loss development factor for the jth development year.

$$Y_{ij} = X_i P_j \tag{2}$$

The unknown quantities of X_i and P_j are estimated from the solution to the following argument:

$$\textit{Minimize} \sum_{i=1}^{n} \sum_{j=1}^{m-i+1} (Y_{ij} - X_i P_j)^2 \tag{3}$$

where n is the number of accident years and m is the number of development years.

If X_i and P_j are the solution for (3), then

$$X_i' = cX_i,$$
$$P_j' = \frac{P_j}{c}, \quad (c>0) \tag{4}$$

are also solutions to (3), since $X_i P_j = X_i' P_j'$. This means that the argument (3) is indeterminate. The indetermination of argument (3) can be eliminated by introducing a constraint such as

$$\sum_{j=1}^{m} P_j = 1. \tag{5}$$

This constraint will force (3) to have one solution

$$Let\ c = \sum_{j=1}^{m} p_j,$$
$$X_i' = cX_i, \tag{6}$$
$$P_j' = \frac{P_j}{c}$$

2.5 Iterative Solution for the Minimization

The partial derivatives in X_i and P_j for equation (3) lead to the following equations:

$$X_i = \sum_{j=1}^{m} \frac{Y_{ij} P_j}{\sum_{j=1}^{m} p_j^2} \tag{7}$$

$$P_j = \sum_{i=1}^{n} \frac{Y_{ij} X_i}{\sum_{i=1}^{n} X_i^2} . \tag{8}$$

These equations can be solved iteratively. It is convenient to start with

$$P_j = \frac{1}{m} \tag{9}$$

as the initial values for P_j's in (7) to obtain the values for X_i's. Using the calculated values of X_i's to obtain the values of P_j in (8) and then recalculate X_i's and P_j's again until the solutions converge.

2.6 Model With Inflation Effects

In the model with inflation effects, the Y_{ij}'s are approximated by expressions X_i, P_j, and U where U is the appropriate incurred loss inflation index. The unknown quantities of X_i, P_j, and U can be estimated by solving the following argument:

$$\textit{Minimize} \sum_{i=1}^{n} \sum_{j=1}^{m-i+1} (Y_{ij} - X_i P_j U^{i+j})^2 \tag{10}$$

In practice, the incurred loss inflation index U is usually pre-determined. In order to solve argument (10), we need to first solve argument (3) by using the iteration method mentioned above. Let X_i' and P_j' be the solution of (10) and taking into account of indetermination,

$$c = \sum_{j=1}^{m} P_j$$

$$X_i = \frac{c X_i'}{U^i} \qquad (11)$$

$$P_j = \frac{P_j'}{c U^j}$$

Then X_i, P_j, and U^{i+j} are the solutions of argument (10).

2.7 Numerical Illustration for De Vylder's LS Method

Using the same data displayed in Table 3 to estimate the ultimate loss incurred, X_i's, and the loss development factors, P_j's, with the assumption that the inflation index for those accident years are 1.

Table 8 displays the incremental loss data for the loss data displayed in Table 3. The solutions calculated by iteration method are as shown in Table 9.

With the estimated value of X_i and P_j, future incurred losses can be estimated as shown in Table 10.

The total IBNR reserves estimated for all accident years using the DVLS method is 645 which is significantly lower than the total IBNR reserves estimated using the chain ladder method.

Table 8

Incremental Loss Incurred

Loss Development Year	Accident Year				
	1990	1991	1992	1993	1994
1	250	267	298	289	300
2	300	315	344	312	
3	117	120	124		
4	50	55			
5	16				

Table 9

Estimated X's and P's Using DVLS

Accident Year	1990	1991	1992	1993	1994
Ultimate Incurred	697	701	733	654	671

Development Year	1	2	3	4	5
Portion Paid	0.368	0.404	0.146	0.063	0.019

Age-to-age loss development factors for this data set can be calculated with the estimated P_j's.

$$F_k = \frac{\sum_{j=1}^{k+1} P_j}{\sum_{j=1}^{k} P_j}, \qquad where\ k = 1,\ldots,m-1 . \tag{12}$$

F_k is the age-to-age loss development factor for development year k to k+1. The calculated age-to-age loss development factors for this data set are shown in Table 11.

Table 10

Predicted IBNR Reserves

Loss Development Year	Accident Year				
	1990	1991	1992	1993	1994
2					271
3				95	98
4			46	41	42
5		13	14	12	13
Total IBNR Reserves	0	13	60	148	424

Table 11

Loss Development Factors Comparison

Development Year	Loss Development Factors	
	DVLS	Chain-Ladder
1 - 2	2.098	2.157
2 - 3	1.189	1.204
3 - 4	1.069	1.077
4 - 5	1.019	1.022

The calculated age-to-age loss development factors by the DVLS methods are very close to the age-to-age loss development factors selected based on average on age-to-age loss development ratios in the chain ladder method. However, the selected age-to-age loss development factors for the chain ladder method are consistently larger than the calculated development factors for the DVLS method. This is the reason the estimated IBNR reserves for the chain ladder method is significantly larger than the estimated IBNR reserves using the DVLS method.

2.8 Deficiencies of De Vylder's Least Squares Method

De Vylder's least squares method estimates the X_i's, loss incurred ultimates, and the P_j's, incremental loss development factors, that minimizes the argument

$$\sum_{i=1}^{n}\sum_{j=1}^{m}(Y_{ij} - X_i P_j)^2. \tag{13}$$

If there is one outlier in the data set, the whole estimation will be changed. If the outlier is significantly larger than the other loss data, the estimated incurred ultimates and so are the estimated incremental loss development factors will be distorted so much that could generates unreasonable predicted future incurred losses. This implies the breakdown point for this method is equal to zero.

Table 12 displayed a contaminated data set that one loss incurred data in Table 8 is replaced with an unusual large number. The incremental loss incurred for the 2nd loss development year for accident year 1992 has been changed from 344 to 1000. Future incurred losses and IBNR reserves were then estimated with De Vylder's least squares method.

Table 12

Contaminated Data Set

Loss Development Year	Accident Year				
	1990	1991	1992	1993	1994
1	250	267	298	289	300
2	300	315	1000	312	
3	117	120	124		
4	50	55			
5	16				

Estimating the IBNR reserves for the data above with De Vylder's Least Squares Method will generate the following results shown in Table 13.

Table 13

Distorted DVLS Estimates

Accident Year		1990	1991	1992	1993	1994
Ultimate Incurred	W/O Outliers	697	701	733	654	671
	W Outliers	610	612	1466	553	943

Development Year		1	2	3	4	5
Portion Paid	W/O Outlier	0.368	0.404	0.146	0.063	0.019
	W Outlier	0.262	0.550	0.095	0.072	0.022

With the estimated value of X_i and P_j, future incurred losses can be estimated as in Table 14.

Table 14

Distorted DVLS Predicted IBNR Reserves

Loss Development Year	Accident Year				
	1990	1991	1992	1993	1994
2					519
3				52	89
4			106	40	68
5		13	32	12	20
Total IBNR Reserves	0	13	138	104	696

The total IBNR is estimated as 951. As the tables above have shown, just one outlier significantly changes the estimated IBNR reserve needed. In Addition, the estimated loss development factors were also changed tremendously due to that one outlier as shown in Table 15.

This simple example demonstrates that the De Vylder's Least Squares method is not robust to outliers.

Compared to the incurred ultimates estimated by chain ladder method, the results estimated by this method are not sensitive to the accident year incurred trend. By looking at the data triangle, Table 16, one could easy see that there is an increasing trend for the incurred losses along the accident year. However, the ultimate incurred losses for accident years 1993 and 1994 estimated by the De Vylder's least squares method, as Table 17 shown, are smaller than the one for accident year 1990 which is inconsistent with what the data triangle shows.

Table 15

Estimated Loss Development Factors Comparison

Development Year	Estimated Loss Development Factors	
	Without Outlier	With Outlier
1 - 2	2.098	3.012
2 - 3	1.189	1.117
3 - 4	1.069	1.080
4 - 5	1.019	1.022

Table 16

Incremental Loss Incurred

Loss Development Year	Accident Year				
	1990	1991	1992	1993	1994
1	250	267	298	289	300
2	300	315	344	312	
3	117	120	124		
4	50	55			
5	16				

The last but rather important deficiency for this method is it does not provide any statistical inferences for testing hypotheses. Actuaries using this method would not be able to determine if the estimated IBNR reserves and the calculated loss development factors are statistical acceptable.

Table 17

Estimated Loss Incurred Ultimates

Accident Year	Estimated Incurred Ultimates				
	1990	1991	1992	1993	1994
Chain Ladder	733	774	843	796	856
De Vylder's	697	701	733	654	671

CHAPTER III

ROBUSTIFICATION OF DE VYLDER'S LEAST SQUARES METHOD

3.1 Norm

A Norm is a non-negative function, $\|.\|$, defined on R^n with the following properties:

1. $\|\mathbf{y}\| \geq 0$ for all $\mathbf{y}$

   $\|\mathbf{y}\| = 0$ if and only if $\mathbf{y} = 0$

2. $\|a\mathbf{y}\| = |a|\,\|\mathbf{y}\|$ for all real a

3. $\|\mathbf{y} + \mathbf{z}\| \leq \|\mathbf{y}\| + \|\mathbf{z}\|$

The distance between two vectors is $d(\mathbf{z},\mathbf{y}) = \|\mathbf{z} - \mathbf{y}\|$.

Given a linear model,

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{e}, \tag{14}$$

where $\mathbf{Y}' = (Y_1,\ldots,Y_n)$ is an $N \times 1$ observation vector, $\mathbf{X}'$ is an $n \times p$ matrix whose columns $x_1,\ldots,x_p$ are linearly independent, $\boldsymbol{\beta}' = (\beta_1,\ldots,\beta_p)$ is a $p \times 1$ vector of unknown regression parameters, and $\mathbf{e}' = (e_1,\ldots,e_n)$ is a $N \times 1$ vector of iid errors from some absolutely continuous distribution F and with density f. With a specified norm, $\|.\|$, $\hat{\beta}$ can then be estimated as

$$\hat{\beta} = argmin \| \boldsymbol{Y} - \boldsymbol{X}\beta \| . \tag{15}$$

The dispersion function induced by the norm is given by,

$$D(\beta) = \| \boldsymbol{Y} - \boldsymbol{X}\beta \| . \tag{16}$$

$D(\beta)$ is a convex and continuous function of β which is differentiable almost everywhere. $\hat{\beta}$ can be estimated by minimizing $D(\beta)$. The gradient process is defined by the function

$$S(\beta) = -\frac{d}{d\beta} D(\beta) \tag{17}$$

where $S(\beta)$ is a nonincreasing function. Its discontinuities are the points where $D(\beta)$ is nondifferentiable. The minimizing value is a value where $S(\beta) \doteq 0$. Hence, estimating the $\hat{\beta}$ can then be expressed as solving the equation

$$S(\hat{\beta}) \doteq 0. \tag{18}$$

3.2 L_2-Norm

The square of the L_2-*Norm* is defined as

$$\| x \|_2^2 = \sum_{i=1}^{n} x_i^2 . \tag{19}$$

The dispersion function for the linear model (14) induced by the L_2-*Norm* is

$$D(\beta) = \sum_{i=1}^{n} \left(y_i - \sum_{j=1}^{p} x_{ij}\beta_j \right)^2 , \tag{20}$$

and the gradient function induced by the L_2-norm is

$$S(\beta) = -2X'(Y - X\beta) . \tag{21}$$

Solving $S(\beta) = 0$ is hence equivalent to

$$X'Y = X'X\beta . \tag{22}$$

Since X is full rank and so does $X'X$, hence $X'X$ is invertible. The solution to the equation (22) is then

$$\hat{\beta} = (X'X)^{-1}X'Y . \tag{23}$$

The estimation derived from the L_2-*Norm* is commonly known as the least square estimation. When it is further assumed that the errors are normally distributed

$$e \sim N_n(0,\sigma^2 I_n) , \tag{24}$$

then

1. $\hat{\beta}$ is also the Maximum Likelihood Estimator of β.
2. $\hat{\beta} \sim N_{p+1}(\beta, \sigma^2 (X'X)^{-1})$
3. $\frac{(n-p-1)s^2}{\sigma^2} \sim \chi^2_{n-p-1}$

4. The standard error of $\hat{\beta}_i$ is estimated by $s[(X'X)^{-1}]_{ii}^{1/2}$,

where

$$s^2 = \frac{1}{n-p} \| Y - X\hat{\beta} \|_2^2 . \tag{25}$$

It can be shown that the estimation of De Vylder's least squares method is derived from the L_2-*Norm*. Let Y_{ij} be the incurred loss for accident year i and development year j. Let X_i be the ultimate incurred loss for accident year i and let P_j be the proportion of loss being paid for development year j. Then

$$\begin{aligned} Y &= XP \\ and \quad X'Y &= X'XP \end{aligned} \tag{26}$$

This implies the P's can be estimated by solving the equation

$$\hat{P} = (X'X)^{-1}X'Y \tag{27}$$

which is equivalent to the estimation of P's in De Vylder's least square method. The same implication applies to the estimation of X's.

In the De Vylder's least square method, P's are estimated by the equation

$$\hat{P}_j = \frac{\sum_{i=1}^{n} Y_{ij}X_i}{\sum_{i=1}^{n} X_i^2} . \tag{28}$$

As it was mentioned in Chapter II, if there is a incurred loss in the jth development year that is an outlier, an extremely large loss relative to other incurred losses, the predicted P_j will be over estimated. Then when the predicted P's are used to estimate the X's,

$$\hat{X}_i = \frac{\sum_{j=1}^{m} Y_{ij} P_j}{\sum_{j=1}^{m} P_j^2} , \tag{29}$$

the predicted X's will also be distorted. This means if there is one extremely large outlier, this estimation will break down. Using the definition of breakdown point stated in Chapter II, the breakdown point for the L_2*-Norm* is 0.

3.3 L_1-Norm

The L_1*-Norm* is defined as

$$\|X\|_1 = \sum_{i=1}^{n} |x_i| . \tag{30}$$

Given a linear model

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{e} \tag{31}$$

where **Y** is an $N \times 1$ observation vector, **X** is an $N \times p$ matrix of known regression constants, $\boldsymbol{\beta}$ is a $p \times 1$ vector of unknown regression parameters, and **e** is an $N \times 1$ vector of iid errors with distribution function F and density function f which is symmetric about 0. Assume the scalar intercept for this linear model is 0. The dispersion function induced by the L_1-*Norm* is

$$D(\beta) = \sum_{i=1}^{n} \left| Y_i - \sum_{j=1}^{p} X_{ij}\beta_j \right| , \tag{32}$$

and the gradient function for each β_j is given by

$$S(\beta_j) = \sum_{i=1}^{n} X_i sgn\left(Y_i - \sum_{j=1}^{p} X_{ij}\beta_j \right) , \tag{33}$$

where sgn(u) = 1 if u > 0,

sgn(u) = 0 if u = 0, and

sgn(u) = -1 if u < 0.

β_j can then be estimated by solving the equation

$$S(\hat{\beta}_j) \doteq 0. \tag{34}$$

Since S(β) is a step function, more than one β estimates could possibly exist.

With the same assumption of the De Vylder's method

$$Y_{ij} = X_i P_j , \tag{35}$$

the dispersion functions for estimating P_j and X_i are

$$D(P_j) = \sum_{i=1}^{n} \left| Y_{ij} - X_i P_j \right| \tag{36}$$

and

$$D(X_i) = \sum_{j=1}^{m} \left| Y_{ij} - X_i P_j \right| \tag{37}$$

respectively. The corresponding gradient functions are

$$\begin{aligned} S(P_j) &= -\frac{\partial D(P_j)}{\partial P_j} = \sum_{i=1}^{n} X_i sgn(Y_{ij} - X_i P_j) \\ and \quad S(X_i) &= -\frac{\partial D(X_i)}{\partial X_i} = \sum_{j=1}^{m} P_j sgn(Y_{ij} - X_i P_j) \ . \end{aligned} \tag{38}$$

This implies that X's and P's can be estimated by solving the following equations

$$\begin{aligned} S(\hat{P}_j) &\doteq 0 \\ and \quad S(\hat{X}_i) &\doteq 0 \ . \end{aligned} \tag{39}$$

In the process of estimating IBNR reserves, X's and P's can then be estimated by first let

$$P_j = \frac{1}{m}, \quad j = 1,...m \tag{40}$$

as the initial values of P. Then solve for $\hat{X}_i$ such that

$$0 \doteq \sum_{j=1}^{m} P_j sgn(Y_{ij} - \hat{X}_i P_j) \ . \tag{41}$$

Take the estimated value of X's and solve for $\hat{P}_j$ such that

$$0 \doteq \sum_{i=1}^{n} X_i sgn(Y_{ij} - X_i \hat{P}_j) \ . \tag{42}$$

Recalculate X_i's and P_j's again until the solutions converge.

3.4 Numerical Illustration for L_1 Estimation

As it was stated above, the gradient function for the L_1-*Norm* is a non-increasing step function. The values of $\hat{X}_i$ and $\hat{P}_j$ cannot be calculated directly from the 2 equations (41) and (42). A numerical method is needed to solve for $\hat{X}_i$ and $\hat{P}_j$. For this numerical illustration, the bisection algorithm is used.

For comparison, the same data set is used to estimate the X_i, total incurred loss for accident year i, and P_j, the proportion of incurred losses for loss development year j, with the L_1-*Norm* estimation. Table 18 displays the estimated value for X_i's and P_j's.

Table 18

Estimated X's and P's L_1-*Norm* Estimation

Accident Year	1990	1991	1992	1990	1994
Ultimate Incurred	699	734	801	727	839

Development Year	1	2	3	4	5
Portion Paid	0.341	0.409	0.159	0.068	0.023

With the estimated values of X_i's and P_j's, future incurred losses can be estimated as shown in Table 19.

Table 19

Predicted IBNR Reserves

Loss Development Year	Accident Year				
	1990	1991	1992	1993	1994
2					343
3				116	134
4			55	50	57
5		17	18	17	19
Total IBNR Reserves	0	17	73	182	553

Estimated total IBNR reserves for the 5 accident years is 825. This is very close to the total IBNR reserves estimated using the chain ladder method.

3.5 Weighted L_1-Norm

Let $R(|x_i|)$ be the rank of $|x_i|$ among $|x_1|,\ldots,|x_n|$. Then consider the function

$$\|x\|_3 = \sum_{i=1}^{n} R(|x_i|)\,|x_i| \ . \qquad (43)$$

Hettmansperger and McKean (1983) showed that $\|x\|_3$ is a norm on R^n. Given the same linear model of (31), let

$$e_i = y_i - x_i\beta \ . \tag{44}$$

Then the dispersion function induces by this norm is

$$D(\beta) = \sum_{i=1}^{n} R(|e_i|)\,|e_i| \ , \tag{45}$$

and the gradient function for this norm is

$$S(\beta) = \sum_{i=1}^{n} x_i R(|e_i|) sgn(e_i) \ . \tag{46}$$

Given the $\|x\|_3$ and the assumption stated in (35), $\hat{X}_i$ and $\hat{P}_j$ can be estimated by solving the following equations

$$\begin{aligned} 0 &\doteq \sum_{j=1}^{m} P_j R(|Y_{ij} - \hat{X}_i P_j|) sgn(Y_{ij} - \hat{X}_i P_j) \\ \text{and} \quad 0 &\doteq \sum_{i=1}^{n} X_i R(|Y_{ij} - X_i \hat{P}_j|) sgn(Y_{ij} - X_i \hat{P}_j) \end{aligned} \tag{47}$$

respectively.

Table 20 displays the estimated value of X_i's and P_j's for the same data set using the $\|x\|_3$ norm. With the estimated values of X_i's and P_j's, future incurred losses can be estimated as shown in Table 21.

Estimated total IBNR reserves for the 5 accident years is 849. Again, this is very closed to the total IBNR reserves estimated using the L1-Norm and chain ladder method.

Table 20

Estimated X's and P's *Weighted L_1-Norm* Estimation

Accident Year	1990	1991	1992	1993	1994
Ultimate Incurred	732	772	846	796	860

Development Year	1	2	3	4	5
Portion Paid	0.349	0.406	0.153	0.070	0.022

Table 21

Predicted IBNR Reserves

Loss Development Year	Accident Year				
		1991	1990	1993	1994
2					350
3				121	131
4			59	56	60
5		17	19	17	19
Total IBNR Reserves	0	17	78	194	560

3.6 Pseudo-Norm

An operator $\|.\|_*$ is called a pseudo-norm if it satisfies the following conditions:

1. $\|\mathbf{y}\|_* \geq 0$ for all $\mathbf{y} \in R^n$

2. $\|a\mathbf{y}\|_* = |a|\,\|\mathbf{y}\|_*$ for all real $a \in R$, $\mathbf{y} \in R^n$

3. $\|\mathbf{y} + \mathbf{z}\|_* \le \|\mathbf{y}\|_* + \|\mathbf{z}\|_*$ for all $\mathbf{y},\mathbf{z} \in R^n$

4. $\|\mathbf{y}\|_* = 0$ if and only if $y_1 = \cdots = y_n$

Note that the regular norm satisfies all the properties above except the fourth property, the norm of a vector is 0 if and only if the vector is **0**.

McKean and Schrader (1981) showed that the function

$$\|\boldsymbol{u}\|_* = \sum_{i=1}^{n} a(R(u_i))u_i \ , \tag{48}$$

where *a(i)* are scores such that $a(i) \le \cdots \le a(n)$ and $\sum a(I) = 0$, is a pseudo-norm.

In addition, for general scores of the form

$$a_\varphi(i) = \varphi\left(\frac{i}{n+1}\right) , \tag{49}$$

φ(u) satisfies the following assumptions:

1. $\varphi(u)$ is a nondecreasing function defined on the interval (0,1),
2. $\int_0^1 \varphi(u)du = 0$, and
3. $\int_0^1 \varphi^2(u)du = 1$.

The corresponding pseudo-norm for scores generated by $\varphi(u)$ is denoted as

$$\|\boldsymbol{u}\|_\varphi = \sum_{i=1}^{n} a_\varphi(R(u_i))u_i \ . \tag{50}$$

For example, Wilcoxon pseudo-norm is generated by the linear score function

$$\varphi_R(u) = \sqrt{12}(u - \tfrac{1}{2}) \tag{51}$$

and the sign pseudo-norm is generated by the score function

$$\varphi_S(u) = sgn(2u - 1). \tag{52}$$

3.7 R-Estimate

Let

$$Y = [\mathbf{1} \;\; X]\begin{pmatrix}\alpha \\ \beta\end{pmatrix} + e \tag{53}$$

where Y is a $N \times 1$ observation vector, **1** is an $N \times 1$ column vector of ones, **X** is an $N \times p$ matrix of known regression constants, α is the scalar intercept parameter, β is a $p \times 1$ vector or unknown regression parameters, and **e** is an $N \times 1$ vector of iid errors with distribution function F and density function f. Note that the density function of **e** is not necessary to be symmetric about 0. Hence the median of the distribution of Y_i is

$$\alpha + x_i'\beta \tag{54}$$

where $\mathbf{x}_i'$ is the ith row of **X**.

Let $\mathbf{X}_c$ denote the centered X-matrix, i.e.

$$X_c = X - \mathbf{1}(\bar{x}_1,\cdots,\bar{x}_p), \tag{55}$$

where $\bar{x}_i$ is the mean of the *i*th column of **X**. Then (53) can be written in the form

$$Y = [\mathbf{1} \quad X_c]\begin{pmatrix} \alpha^* \\ \beta \end{pmatrix} + e \, . \tag{56}$$

Definition 1. Let $D(\bullet)$ be a measure of variability that satisfies two properties:

1. $D(\mathbf{Y} + \mathbf{1}a) = D(\mathbf{Y})$
2. $D(-\mathbf{Y}) = D(\mathbf{Y})$

for every $N \times 1$ vector **Y** and scalar a. Then $D(\bullet)$ is called an even, location free measure of dispersion.

If $D(\bullet)$ is an even and location free dispersion function, then $D(\mathbf{Y} - \mathbf{1}\alpha - \mathbf{X}\beta) = D(\mathbf{Y} - \mathbf{1}\alpha^* - \mathbf{X}_c\beta) = D(\mathbf{Y} - \mathbf{X}_c\beta) = D(\mathbf{Y} - \mathbf{X}\beta)$ where $\alpha^* = \alpha + \bar{x}'\beta$ and $\bar{x}' = (\bar{x}_1,...,\bar{x}_p)$. Hence, when estimating $\hat{\beta}$ with $D(\bullet)$, the results will not be changed if either **X** or $\mathbf{X}_c$ is used.

Definition 2. A rank estimate (R estimate) of β is the value $\hat{\beta}$ which minimizes

$$D(Y - X\beta) = \sum a[R(Y_i - x_i'\beta)](Y_i - x_i'\beta) \tag{57}$$

where $R(Y_i - x_i'\beta)$ is the rank of $Y_i - x_i'\beta$ among $Y_1 - x_1'\beta,...,Y_N - x_N'\beta$.

Unlike the least square, this even, location free measure is a linear, rather than quadratic, function of the residuals. Since the influence of outlier enters in a linear rather than quadratic fashion, it is hoped that the estimates generated by the R estimate will be more robust than least-squares estimates.

As the properties of the pseudo-norm indicate, this function $D(\boldsymbol{Y} - \boldsymbol{X}\beta)$ is a nonnegative, continuous, and convex function of $\boldsymbol{\beta}$.

The gradient function of $D(\boldsymbol{Y} - \boldsymbol{X}\beta)$ is given by

$$\begin{aligned}\frac{\partial}{\partial\beta_j}D(\boldsymbol{Y} - \boldsymbol{X}\beta) &= \sum_{i=1}^{N} a[R(Y_i - \boldsymbol{x}_i'\beta)(-x_{ij}) \\ &= -\sum_{i=1}^{N}(x_{ij} - \bar{x}_j)a[R(Y_i - \boldsymbol{x}_i'\beta)]\end{aligned} \tag{58}$$

for j=1,...,p. Let

$$\begin{aligned}S_j(\boldsymbol{Y} - \boldsymbol{X}\beta) &= -\frac{\partial}{\partial\beta_j}D(\boldsymbol{Y} - \boldsymbol{X}\beta) \\ &= \sum_{i=1}^{N}(x_{ij} - \bar{x}_j)a[R(Y_i - \boldsymbol{x}_i'\beta)]\end{aligned} \tag{59}$$

and let

$$\boldsymbol{S}(\boldsymbol{Y} - \boldsymbol{X}\beta) = \begin{pmatrix} S_1(\boldsymbol{Y} - \boldsymbol{X}\beta) \\ \vdots \\ S_p(\boldsymbol{Y} - \boldsymbol{X}\beta)\end{pmatrix}. \tag{60}$$

Then estimating $\hat{\beta}$ is to solve the equation

$$\boldsymbol{S}(\boldsymbol{Y} - \boldsymbol{X}\beta) \doteq 0\,. \tag{61}$$

The Wilcoxon score function where

$$a(i) = \phi\left(\frac{i}{N+1}\right)$$
$$\text{and} \quad \phi(u) = \sqrt{12}\left(u - \frac{1}{2}\right) , \tag{62}$$

will be used as the score function for the R-estimate throughout the rest of this thesis. Note that the Wilcoxon score function has the following two properties:

$$\int_0^1 \phi(u)du = 0$$
$$\text{and} \quad \int_0^1 \phi^2(u)du = 1 .$$

Though the $\hat{\beta}$ is estimated independent to the intercept parameter α, it is estimated with the assumption that the intercept parameter exist. If the distribution of **e** is assumed to be symmetric and $\hat{\beta}$ is estimated using the Wilcoxon scores, the intercept parameter α can be estimated as the Walsh averages of the residuals; i.e. let

$$r_i = Y_i - \mathbf{x}_i'\hat{\beta} , \quad where\ i = 1,\ldots,N , \tag{64}$$

be the residuals. Then

$$\hat{\alpha} = med_{i \le j}\left(\frac{r_i + r_j}{2}\right) . \tag{65}$$

If the distribution of **e** is not assumed to be symmetric and $\hat{\beta}$ is estimated with the Wilcoxon scores, the intercept parameter α can be estimated as the median of the residuals,

$$\hat{\alpha} = med(r_i) \ . \tag{66}$$

3.8 Estimating IBNR Reserves With R-estimates

With the same assumption in (35) for incurred losses, $\hat{X}_i$ can be estimated by solving the equation

$$0 \doteq \sum_{j=1}^{m} (P_j - \bar{P}) a[R(Y_{ij} - \hat{X}_i P_j)] \ , \quad \text{where } \bar{P} = \frac{\sum_{j=1}^{m} P_j}{m} \tag{67}$$

with

$$P_j = \frac{1}{m} \ , \quad \text{for } j = 1, \ldots, m \tag{68}$$

as the initial value of P_j. Then $\hat{P}_j$ can be estimated by solving the equation

$$0 \doteq \sum_{i=1}^{n} (X_i - \bar{X}) a[R(Y_{ij} - X_i \hat{P}_j)] \quad \text{where } \bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} \ . \tag{69}$$

Recalculating X_i's and P_j's again until the solutions converge.

One problem does exist with using the general R-estimate method to estimate IBNR reserves. According to the linear model in (35), there are no intercept parameters for estimating X_i's and P_j's. However, R-estimate is based on the assumption that an intercept does exist while estimating the slope. Consequently, estimating X_i's and P_j's with iteration without estimating intercepts while using R-estimate will not converge.

Dixon and McKean (1996) provided a simple solution for this problem. It was suggested to estimate a dummy intercept parameter with the estimation stated above after the slope was estimated with R-estimate. Using the predicted intercept and slope parameters to calculate the predicted values. Then project the predicted value to the plane which has no intercept. Refer to Appendix A for the algorithm for estimating IBNR reserves with R-estimate.

Table 22 displays the estimated value of X_i's and P_j's for the same data set using the $\|x\|_3$ norm.

Table 22

Estimated X's and P's With *R-estimate*

Accident Year	1990	1991	1992	1993	1994
Ultimate Incurred	697	714	729	654	675

Development Year	1	2	3	4	5
Portion Paid	0.370	0.403	0.145	0.063	0.019

With the estimated values of X_i's and P_j's, future incurred losses can be estimated as shown in Table 23. Estimated total IBNR reserves for the 5 accident years is 647 which is very close to the total IBNR reserves estimated by the De Vylder's least square method.

Table 23

Predicted IBNR Reserves

Loss Development Year	Accident Year				
	1990	1991	1992	1993	1994
2					272
3				95	98
4			46	41	42
5		14	14	12	13
Total IBNR Reserves	0	14	60	148	425

3.9 Iteration Procedures

Two iteration procedures are needed to estimate IBNR reserves with the non-parametric methods described above. One is needed to solve the gradient function S(β). Since S(β) is a non-increasing step function, a iterative algorithm is needed to estimate the $\hat{\beta}_j$. The gradient function S(β) does not necessary have one unique solution. In fact, it could have infinitely many solutions. Since S(β) is a step function it is not easy to show that other iterative algorithms will perform better than the binary-search method because. Other numerical methods such as Newton's method have been tested against the binary-search method to solve for S(β) and found did not perform better than the binary-search method. For its simplicity, binary-search is used to solve for S(β) for the non-parametric methods described above.

In order to use the binary-search method to solve the gradient function, two initial values, a and b such that S(a) × S(b) < 0, are needed. Let a < b, then S(a) > S(b) for S(β) is a non-increasing function. For the linear model without intercept parameter

$$\boldsymbol{Y} = \boldsymbol{X}\beta \ , \tag{70}$$

it is safe to assume that

$$\min\left\{\frac{Y_i}{X_i}\right\} \le \hat{\beta} \le \max\left\{\frac{Y_i}{X_i}\right\} . \tag{71}$$

Thus let

$$a = \min\left\{\frac{Y_i}{X_i}\right\} \text{ and } b = \max\left\{\frac{Y_i}{X_i}\right\} \tag{72}$$

be the two initial values for solving S(β). Note that S(a) > 0 and S(b) < 0 imply that S(a) × S(b) < 0 which means a < $\hat{\beta}$ < b.

Two stopping procedures were tested for the binary-search method. Assume the value of tolerance ε >0 has been determined. $\beta_1, \ldots, \beta_n$ are generated until the following 2 conditions are met:

$$|\beta_n - \beta_{n-1}| < \varepsilon \tag{73}$$

or

$$\left| \frac{\beta_n - \beta_{n-1}}{\beta_n} \right| < \varepsilon \ , \quad \text{where } \beta_n > 0 \ . \tag{74}$$

Without additional knowledge about the gradient function or β, the inequality (74) is the best stopping criteria to apply because it measures the relative error. The stopping criteria of (74) is used for the iteration methods described above.

One would want to know the number of iteration, n, that is needed for the estimated parameter to converge. Let β_0 be the true value of β. Then

$$|\beta_n - \beta_0| \leq \frac{b - a}{2^n} \ , \quad n \geq 1 \ . \tag{75}$$

One can easily determine the minimum number of iteration, n, as follow:

$$n > \frac{-\log\left(\frac{\varepsilon}{b-a}\right)}{\log 2} \ . \tag{76}$$

3.10 Convergence of the Estimation Procedure

The second iteration method is needed to calculate the X_i's and P_j's. X_i's are estimated first with the initial values of P_j's as the independent variables. Then P_j's are estimated with the estimated X_i's as the independent variables. X_i's and P_j's again are estimated until they converge. Each iteration is intend to minimize the dispersion function in terms of X_i's and in terms of P_j's. If X_i's and P_j's are the solution of the original model

$$Y_{ij} = X_i P_j \,, \tag{77}$$

then

$$X_i' = cX_i \,, \quad P_j' = \frac{P_j}{c} \,, \quad \text{where } c > 0 \tag{78}$$

is also a solution because $X_i'P_j' = X_iP_j$. Thus, the solutions generated by the iteration is indeterminate. As it was mentioned in Chapter II, the indeterminate can be eliminated by a constraint such as

$$\sum_{j=1}^{m} p_j = 1. \tag{79}$$

Different solutions can be generated by the iteration method depending on the initial values of the P_j's. However, all these different solutions are connected by the constraint (79). Since the X_i's and P_j's do not have unique solutions, it is not necessary to iterate the estimation until both X_i's and P_j's converge. The estimates obtained after four or five iterations can be considered as limit solutions from the practical point of view.

Numerous data sets have been tested with these non-parametric iteration methods. After a few iterations, the X_i's tend to jump between 2 sets of numbers. Though the 2 sets of numbers differ very little, they do not necessarily converge. On the other hand the P_j's tend to converge very well after a few iterations. Using the stopping criteria (74) to measure the estimated P_j's to determine the convergence of the iteration does generate consistent results.

3.11 Deficiencies of Rank-Based Iteration Method

The non-parametric methods described above are robust to outliers. However, as McKean and Hettmansperger (1996) has shown, the influence function of $\hat{\beta}_R$ is given by

$$IF(x,y;\hat{\beta}_R) = \tau\Sigma^{-1}\varphi(F(y))\mathbf{x} \ , \tag{80}$$

where Σ is the variance-covariance matrix of $\mathbf{x}$, τ is the scale parameter, and $\varphi(\bullet)$ is the score function corresponding to the R-estimate. This indicates that the influence function of $\hat{\beta}_R$ is bounded in the **y**-space but not in the **x**-space. Figure 1 of Appendix B displays a simple example that the R-estimates are robust to outlier in the **y**-space. Figure 2 of Appendix B displays a different example which shows that the R-estimates are not robust to outlier in the **x**-space.

Since the R-estimate is robust in the **y**-space but not in the **x**-space, estimating the X_i's and P_j's with iteration by this method is not robust to outliers. If there is a significantly large incurred loss in the data set, an outlier to the **y**-space, one of the initial estimated X_i will be much larger than the others. Then when the P_j's are being estimated with the estimated X_i's as the independent variable, the unusually large X's will become an outlier in the **x**-space which will cause unreasonable estimation for the P_j's. While estimating the X_i's again with the distorted estimated P_j's as the independent variable, outlier in the **x**-space again, the estimated X_i's will be distorted. Refer to Appendix B for a numerical illustration.

In addition to not being robust to outliers, the non-parametric methods introduced above do not provide statistical inferences for hypotheses testing. This means that analysts who choose to use these methods to estimate IBNR reserves will not be able to determine if the results are statistically acceptable.

A one-step method which estimating IBNR reserves with general R-estimate will be introduced in Chapter IV that is robust to outliers. This method also provides statistical inferences for hypotheses testing.

CHAPTER IV

THE LOG-MULTIPLICATIVE MODEL

4.1 Introduction

A log-multiplicative model was introduced by Kremer (1982) which takes De Vylder's multiplicative model and add a multiplicative lognormal random error,

$$Y_{ij} = X_i P_j E_{ij} \tag{81}$$

where X_i is a column effect for accident year i and P_j is a row effect for loss development year j. The product of X_i and P_j will correspond to the amount of claims occurred during accident year i incurred on development year j. The random error, E_{ij}, are independent, identically distributed lognormal random, $LN(0,\sigma^2)$, and with density function

$$f(t;\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma t} e^{-\frac{1}{2}\left(\frac{\log t}{\sigma}\right)^2}, \quad t > 0 . \tag{82}$$

Taking the logarithm on both sides of (81), the multiplicative model can be changed into a additive linear model

$$\log Y_{ij} = \log X_i + \log P_j + \log E_{ij} . \tag{83}$$

Let N be the number of incremental incurred losses, Y_{ij}, and **Z** be a N x 1 vector that

$$\boldsymbol{Z} = \begin{bmatrix} \log Y_{11} \\ \vdots \\ \log Y_{1m} \\ \log Y_{21} \\ \vdots \\ \log Y_{2m} \\ \vdots \\ \log Y_{nm} \end{bmatrix}. \tag{84}$$

Let p be the number of parameters, p = m+n-1, then **β** be a p x 1 vector,

$$\boldsymbol{\beta} = \begin{bmatrix} \log X_1 \\ \vdots \\ \log X_n \\ \log P_2 \\ \vdots \\ \log P_m \end{bmatrix}. \tag{85}$$

Note that the P_1 is set equal to 1 (or log $P_1 = 0$) in order to ensure the regression matrix not to be singular. Let **W** be a N x p design matrix,

$$W = \begin{bmatrix} 1\ 0\ \ldots\ 0\ 0\ 0\ 0\ \ldots\ 0 \\ 1\ 0\ \ldots\ 0\ 1\ 0\ 0\ \ldots\ 0 \\ 1\ 0\ \ldots\ 0\ 0\ 1\ 0\ \ldots\ 0 \\ \vdots \\ 1\ 0\ \ldots\ 0\ 0\ 0\ 0\ \ldots\ 1 \\ 0\ 1\ \ldots\ 0\ 0\ 0\ 0\ \ldots\ 0 \\ 0\ 1\ \ldots\ 0\ 1\ 0\ 0\ \ldots\ 0 \\ 0\ 1\ \ldots\ 0\ 0\ 1\ 0\ \ldots\ 0 \\ \vdots \\ 0\ 1\ \ldots\ 0\ 0\ 0\ 0\ \ldots 1 \\ \vdots \\ \vdots \\ 0\ 0\ \ldots\ 1\ 0\ 0\ 0\ \ldots\ 1 \end{bmatrix} . \qquad (86)$$

Let $\epsilon_{ij} = \log E_{ij}$. Then ϵ_{ij} are iid $N(0,\sigma^2)$ random variables. Let **e** be a N x 1 vector of ϵ_{ij}. The model (83) can then be expressed as a linear model,

$$Z = W\beta + e . \qquad (87)$$

4.2 Least Squares Estimation

One can easily verify that the maximum likelihood estimator of β for the linear model (87), which is the same as the least square estimator, is

$$\hat{\beta} = (W'W)^{-1} W'Z .$$

The unbiased estimator of σ^2 is

$$\hat{\sigma}^2 = \frac{1}{N-p}(\boldsymbol{Z}-\boldsymbol{W}\hat{\beta})'(\boldsymbol{Z}-\boldsymbol{W}\hat{\beta}) = \frac{SS}{N-p}$$

where SS is the sum of square residuals. It is also well know that:

1. $\hat{\beta}$ Has a multivariate normal distribution $MN(\beta, \sigma^2(W'W)^{-1})$,
2. $(N-p)\dfrac{\hat{\sigma}^2}{\sigma^2}$ has a χ^2 distribution with (N-p) degrees of freedom,
3. $\hat{\beta}$ and $\hat{\sigma}^2$ are independent,
4. ($\hat{\beta}$,SS) are jointly complete and sufficient statistics for (β, σ^2).

With the estimated β, the estimated X_i's and P_j's can be calculated as follow:

$$\begin{aligned} \tilde{X}_i &= \exp[\hat{\beta}_i] \ , \quad \text{for } i = 2,\ldots,n \ , \\ \tilde{P}_1 &= 1 \ , \\ \tilde{P}_j &= \exp[\hat{\beta}_{j+n-2}] \ , \quad \text{for } j = 2,\ldots,m \ . \end{aligned} \tag{90}$$

and

$$\begin{aligned} \hat{X}_i &= C\tilde{X}_i \ , \\ \hat{P}_j &= C\tilde{P}_j \\ \text{where } C &= \sum_{j=1}^{m} \tilde{P}_j \ . \end{aligned} \tag{91}$$

IBNR reserves needed is estimated as

$$\sum_{i=2}^{n} X_i \sum_{j=1}^{i-1} P_{m-j+1} \ . \tag{92}$$

The same data set which is used in the previous chapters to estimate the X_i, total incurred loss for accident year i, and P_j, the proportion of incurred losses for loss development year j, with the log-multiplicative model. Table 24 displays the estimated value for X_i's and P_j's.

Table 24

Estimated X's and P's *LS* Estimation

Accident Year	1990	1991	1992	1993	1994
Ultimate Incurred	734	778	835	798	856

Development Year	1	2	3	4	5
Portion Paid	0.351	0.404	0.154	0.069	0.022

With the estimated values of X_i's and P_j's, future incurred losses can be estimated as displayed in Table 25.

The total IBNR estimated by this log-multiplicative model is 844 which is very close to the one estimated by chain ladder method. This method seems to provide good estimation as long as the data set being analyzed has stable loss distribution and is free of outliers. On the other hand, as it is shown in Table 25 and Table 26, this method generates unreasonable estimation if there is outlier in the data set. As it was done in the previous chapters to test this method, a incremental loss incurred has been replaced with an outlier. X_i's and P_j's are then estimated as shown in Table 27.

Table 25

Least Squares Predicted IBNR Reserves

Loss Development Year	Accident Year				
	1990	1991	1992	1993	1994
2					346
3				123	132
4			58	55	57
5		17	18	18	19
Total IBNR Reserves		17	76	0	555

Table 26

Contaminated Incurred Data

Loss Development Year	Accident Year				
	1990	1991	1992	1993	1994
1	250	267	298	289	300
2	300	315	1000	312	
3	117	120	124		
4	50	55			
5	16				

Additional to the estimated total accident year incurred losses, X's, has been distorted, the loss development year incurred portions, P's, are also distorted. Note that there is only one outlier which is not so extraordinary large was contaminating the data set. However, all the estimated accident year total incurred losses along with all the estimated loss development year incurred portions have been distorted.

Table 27

Estimated X's and P's With Outlier

Accident Year	1990	1991	1992	1993	1994
Ultimate Incurred	766	812	1244	785	962

Development Year	1	2	3	4	5
Portion Paid	0.312	0.470	0.131	0.067	0.021

With the estimated values of X_i's and P_j's, future incurred losses are estimated as displayed in Table 28. The estimated total IBNR reserves for this data set is 960 which is much larger than the one estimated without outlier. Further sensitivity studies have been done to compare this least square method with the Rank-based method which is introduced in the following section. The results of the sensitivity studies are shown in the next chapter.

4.3 Rank-based Linear Model

Consider the linear model

$$Z = \alpha \mathbf{1} + V\beta + e \,, \tag{93}$$

where $\mathbf{Z}' = (Z_1,\dots,Z_N)$ are observations, $\mathbf{1}' = (1,\dots,1)$, V is an N × p centered, full rank design matrix, $\boldsymbol{\beta}' = (\beta_1,\ldots,\beta_p)$ and $\mathbf{e}' = (e_1,\ldots,e_N)$ are iid errors from some absolutely continuous distribution F with density f and median 0. The unknown parameters are α and $\boldsymbol{\beta}$. It is convenient to keep the intercept parameter α and the vector of regression parameters $\boldsymbol{\beta}$ separate.

Table 28

Predicted IBNR Reserves

Loss Development Year	Accident Year				
	1990	1991	1992	1993	1994
2					452
3				103	126
4			83	53	64
5		17	26	16	20
Total IBNR Reserves	0	17	109	172	662

Consider the general R pseudo-norm discussed in Chapter III which is given as

$$\|v\|_\varphi = \sum_{i=1}^{n} a(R(v_i))\, v_i \tag{94}$$

where $a(1) \le a(2) \le \ldots \le a(n)$ is a set of scores generated as

$$a(i) = \varphi\left(\frac{i}{n+1}\right)$$

for some nondecreasing score function $\varphi(u)$ defined on the interval $(0,1)$ and standardized such that

$$\int \varphi(u)du = 0$$

$$\text{and} \quad \int \varphi^2(u)du = 1 \ .$$

Recall that the Wilcoxon scores are generated by the linear score function $\varphi(u) = \sqrt{12}(u - \frac{1}{2})$ and throughout this chapter, the Wilcoxon score function will be used for the R-estimates.

The dispersion function for this linear model can be defined as

$$\begin{aligned} D_R(\beta) &= \|Z - V\beta\|_R \\ &= \sum_{i=1}^{n} a(R(Z_i - v_i'\beta))(Z_i - v_i'\beta) \ , \end{aligned} \tag{97}$$

where $\mathbf{v}_i'$ is the ith row of $\mathbf{V}$. Note that the matrix $\mathbf{V}$ is centered by subtracting the column means to get $V_c = V - \mathbf{1}(\bar{v}_1,...,\bar{v}_N)$ where $\bar{v}_i$ is the mean of the ith column of $\mathbf{V}$. As it was stated in Chapter III, either $\mathbf{V}$ or $\mathbf{V}_c$ can be used without altering the estimates while working with $D_R(\bullet)$. Throughout this chapter, $\mathbf{V}$ is written as the centered matrix.

Jaeckel (1972) showed that $D_R(\beta)$ is continuous and convex in β. Hence, the estimate of β is a value that minimizes $D_R(\beta)$. Note again that the pseudo-norm is invariant to α, the intercept, which can be estimated by a location estimate of the residuals as described in the previous chapter.

The dispersion function $D_R(\beta)$ is differentiable almost everywhere. Let $S_R(\beta)$ be the gradient function of $D_R(\beta)$ such that

$$S(\beta) = -\frac{dD_R(\beta)}{d\beta} = \sum_{i=1}^{n} a(R(Z_i - v_i'\beta))v_i . \tag{98}$$

The estimate of β is a vector of $\hat{\beta}$ that satisfies the equation

$$S(\hat{\beta}_R) \doteq 0 . \tag{99}$$

Let Σ be a positive definite matrix that

$$\Sigma = \lim_{n\to\infty} n^{-1} V'V , \tag{100}$$

τ be the scale parameter, and $\varphi(u)$ be a score function. Then the influence function for the R-estimate for this linear model is given by

$$IF(v,z) = \tau\Sigma^{-1}\varphi(F(z - v\beta))v . \tag{101}$$

This indicates that the influence function for $\hat{\beta}_R$ is bounded in the **y**-space but not in the **x**-space. Since the influence function is unbounded in the **x**-space, the R-estimates are susceptible to a large distortion when there is contamination in **x**. This implies that the breakdown point of the R-estimates in regression is 0. However, this will not be a problem for this linear model because the variable **V** is a designed matrix which is not a realization of a random vector.

Let β_0 be the true parameter vector. The asymptotic distribution of the R-estimates is

$$\sqrt{n}(\hat{\beta}_R - \beta_0) \xrightarrow{D} N_p(\mathbf{0},\tau^2\Sigma^{-1}) \tag{102}$$

which provides the basis for estimation and standard error of the estimators. The scale parameter τ is defined by

$$\tau^{-1} = \int \varphi(u)\left(-\frac{f'(F^{-1}(u))}{f(F^{-1}(u))}\right) du \ . \tag{103}$$

In the case where Wilcoxon scores are used, τ becomes

$$\tau = \frac{1}{\sqrt{12}\int f^2(u)du} \ . \tag{104}$$

Note that the distribution of the least square estimate is

$$\sqrt{n}(\hat{\beta}_L - \beta_0) \xrightarrow{D} N(\mathbf{0},\sigma^2\Sigma^{-1}) \ . \tag{105}$$

This implies the asymptotic relative efficiency (ARE) for the R-estimate relative to the least square estimate is

$$e = \frac{\sigma^2}{\tau^2} \ . \tag{106}$$

In the case where Wilcoxon scores are used, the ARE of the R-estimate becomes

$$e = 12\sigma^2\left(\int f^2(u)du\right)^2 . \qquad (107)$$

If the error distribution is normal, the ARE for the R-estimate is equal to 0.955 which means less than 5% of efficiency relative to the optimal estimate is lost. However, if the distribution for the error is long tailed, the ARE for the R-estimate relative to the classical analysis will be substantially larger than 1.

For the linear model stated above, the intercept parameter estimate $\hat{\alpha}$ denote the location R-estimate computed from the residuals $\hat{e}_i = Z_i - \nu'\hat{\beta}_R$. Assuming that the error distribution is symmetric about 0, the R-estimate of the intercept parameter, α, is the value $\hat{\alpha}$ which minimizes the dispersion function

$$D^*(\alpha) = \sum_{i=1}^{n} a^*(R^*(|\hat{e}_i - \alpha|)\,|\hat{e}_i - \alpha| . \qquad (108)$$

Minimizing $D^*(\alpha)$ is equivalent to solving the equation

$$S^*(\alpha) = \sum_{i=1}^{n} a^*(R^*(|\hat{e}_i - \alpha|) sgn(\hat{e}_i - \alpha) \doteq 0 , \qquad (109)$$

where $R^*(\bullet)$ is the rank among the absolute value and $a^*(i) = \varphi^*(\frac{i}{n+1})$. $\varphi^*(u)$ is nonnegative and nondecreasing on (0,1). Then $\hat{\alpha}$ is a Hodges & Lehmann (1963) type of estimate of the intercept. If the Wilcoxon scores are used then

$$\hat{\alpha} = med\left(\frac{\hat{e}_i + \hat{e}_j}{2}\right), \quad \text{where } i \le j \ . \tag{110}$$

This is the Walsh average of the residuals. McKean and Hettmansperger (1978) showed that

$$\sqrt{n}\begin{pmatrix}\hat{\alpha}\\ \hat{\beta}_R\end{pmatrix} \overset{D}{\rightarrow} N_{p+1}(0, \tau^2\Sigma^{-1}) \ . \tag{111}$$

Note that $\hat{\alpha}$ and $\hat{\beta}_R$ are asymptotically independent because of the centered design described above.

Since the distribution theory for $\hat{\beta}_R$ does not require the error distribution to be symmetric, the intercept parameter α can be estimated as

$$\hat{\alpha}_R = med(\hat{e}_i) \ . \tag{112}$$

This identifies the intercept as the median of the residuals. If the intercept parameter α is estimated as the median of the residuals, $\tau^2\Sigma^{-1}$ in (111) is replaced by

$$\begin{bmatrix}\tau_s^2 & \mathbf{0}'\\ \mathbf{0} & \tau^2\Sigma^{-1}\end{bmatrix}, \tag{113}$$

where

$$\tau_S^2 = \frac{1}{4f^2(0)} \ . \tag{114}$$

As it was stated above, the asymptotic distribution for $\hat{\alpha}$ is normal with mean 0 and variance τ_s^2. This normal distribution can be used to approximate probabilities concerning the sample median. When the underlying form of the distribution is unknown, this asymptotic variance must be estimated. Theorem below provides the key to the estimation of the asymptotic variance.

Theorem 1. Suppose S(θ) is a Pitman Regular estimating function with efficacy c. Let L be the length of the corresponding confidence interval. Then

$$\frac{\sqrt{N}L}{2z_{\alpha/2}} \xrightarrow{P} \frac{1}{c}, \tag{115}$$

where θ is the unique median.

Refer to McKean and Hettmansperger (1996) for the proof of this theorem. Based on this theorem, the asymptotic standard error, the square root of the asymptotic variance, for the sample median can be estimated. McKean and Schrader (1984) suggested an estimator of τ_s on a normal approximation to the length of a 100(1 - α)% distribution-free confidence interval for θ. The resulting estimator of the asymptotic standard error of the residuals is given by

$$\tau_s = \frac{\left(e_{(N-k+1)} - e_k\right)}{2Z_{\alpha/2}} \tag{116}$$

where

$$k = \frac{N+1}{2} - Z_{\alpha/2}\sqrt{\frac{N}{4}} \tag{117}$$

and $Z_{\alpha/2}$ is the 100(1 - α/2) percentile of the standard normal distribution. Note that (e_{N-k+1}, e_k) is the asymptotic confidence interval for the $\hat{\alpha}_R$.

As a simple example, take α = 0.05, $Z_{\alpha/2}$ = 2, then k = $\frac{N+1}{2} - \sqrt{N}$. The asymptotic confidence interval for the $\hat{\alpha}_R$ will be

$$\left(e_{(\frac{N+1}{2}+\sqrt{N})}, e_{(\frac{N+1}{2}-\sqrt{N})}\right) \tag{118}$$

and the estimate of the asymptotic standard error for the $\hat{\alpha}_R$ is given by

$$\tau_s = \frac{\left(e_{(\frac{N+1}{2}+\sqrt{N})} - e_{(\frac{N+1}{2}-\sqrt{N})}\right)}{4} . \tag{119}$$

Sheather and McKean (1987) discuss different approaches to the estimation of this standard error.

Based on the asymptotic distribution for the $\hat{\beta}$ stated above, the asymptotic confidence interval for a linear function **h'**β is given by

$$\boldsymbol{h}'\hat{\boldsymbol{\beta}}_\varphi \pm t_{(\alpha/2, N-p-1)}\hat{\tau}_\varphi\sqrt{\boldsymbol{h}'(V'V)^{-1}\boldsymbol{h}} , \tag{120}$$

4.4 Estimates of the Scale Parameter τ

Let e_i be the residuals of the estimate that

$$e_i = Z_i - \hat{\alpha} - v_i \hat{\beta}_R .$$

Let

$$\begin{aligned} H(t) &= P(|e_1 - e_2| \leq t) \\ &= \int [F(e_1 + t) - F(e_2 - t)] dF(e_2) , \end{aligned} \tag{122}$$

where $t > 0$. $H(t) = 0$ if $t \leq 0$. Let $h(t)$ denote the density of $H(t)$. Then upon differentiating under the integral sign in (122), it follows that

$$h(0) = 2 \int f^2(e_2) de_2 . \tag{123}$$

From (104),

$$\tau = \frac{1}{\sqrt{3}\left(2 \int f^2(u) du\right)} = \frac{1}{\sqrt{3}\, h(0)} . \tag{124}$$

So to estimate τ, $h(0)$ needs to be estimated. As estimate of $H(t)$, consider the following empirical distribution function

$$H_N(t) = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{i=1}^{N} I(|e_i - e_j| \leq t) , \tag{125}$$

where $I(\bullet)$ is the indication function. An estimate of $h(0)$ is an estimate of the form

$$\hat{h}(0) = \frac{\hat{H}_N(t_N)}{t_N} . \tag{126}$$

With the bias corrected empirical derivative,

$$\hat{\tau} = \sqrt{\frac{N}{N-p-1}} \frac{t_N}{\sqrt{3}H_N(t_N)} . \tag{127}$$

For moderate sample sizes that the ratio of N/p exceeds 5, it is recommended $\delta = .80$ where

$$t_N = \frac{H_N^{-1}(\delta)}{\sqrt{N}} . \tag{128}$$

For a ratio less than 5, larger values of δ, around .90 is recommended.

4.5 Hypotheses Testing

Consider the hypotheses

$$H_0\text{: } \boldsymbol{M\beta} = \mathbf{0} \quad \textit{versus} \quad H_A\text{: } \boldsymbol{M\beta} \neq \mathbf{0}, \tag{129}$$

where **M** is a full row rank $q \times p$ matrix which determines the linear combination of the components of β under test. Let D_R(full) denote the minimum value of the dispersion function (98) when the full model is fit. Let D_R(Red) denote the minimum value of the dispersion function (98) when the reduced model is fit. Model (93) based on H_0 is the reduced model.

The F test of the hypotheses for the least square estimation is based on the standardized reduction of the sums of squares between fitting the reduced model and the full model. The F test of the hypotheses for the rank-based estimation is similar to the least square. Decision is made based on the reduction in dispersion given by

$$RD = D_R(Red) - D_R(Full). \tag{130}$$

McKean and Hettmansperger (1976) showed that under H_0, and the same regularity conditions used to establish the asymptotic theory for $\hat{\boldsymbol{\beta}}_R$,

$$F_R = \frac{\frac{RD}{q}}{\frac{\hat{\tau}}{2}} \xrightarrow{D} \frac{\chi_q^2}{q} \tag{131}$$

distribution. McKean and Sheather (1991) showed that the use of F critical values with q and $N - p - 1$ degree of freedom instead of χ^2 critical values provide stable significance levels for a broad range of underlying error distributions and moderate sample sizes. Hence the test based on F_R can be summarized in a robust analysis of variance table, Table 29, which is similar to the classical ANOVA table. For an asymptotic level α test, reject H_0 if $F_R \geq F(\alpha,q,N-p-1)$.

McKean and Hettmansperger (1976) showed that the test based on F_R is consistent under the same regularity conditions. Under a sequence of contiguous alternatives, F_R converges in distribution to a noncentral χ^2 with the same noncentrality parameter as the least squares F test with τ^2 replacing σ^2. Hence the asymptotic relative efficiency of the F_R test statistic to the least squares F test statistic is

$$e = \frac{\sigma^2}{\tau^2} . \tag{132}$$

Thus F_R test statistic possesses both robustness of validity and robustness of efficiency.

Table 29

Rank-based ANOVA Table for H_0: $\mathbf{M\beta} = 0$

Source	Reduction in Dispersion	df	Mean Reduction in Dispersion	F_R
Regression	RD=D_R(Red)-D_R(Full)	q	RD/q	$\frac{RD/q}{\hat{\tau}/2}$
Error		N-p-1	$\hat{\tau}/2$	

This reduction in the dispersion test discussed above is similar to the likelihood ratio test in classical inference. There are two other tests usually discussed in classical inference. They are the Wald test and the Rao-scores test. The rank-based analogues of the Wald test is a quadratic form in full model estimates and is given by

$$F_{R,Q} = \frac{(M\hat{\beta})'[M(V'V)^{-1}M']^{-1}(M\hat{\beta})/q}{\hat{\tau}^2} . \tag{133}$$

Provided $\hat{\tau}$ is a consistent estimate of τ. It follows from the asymptotic distribution of

$\hat{\beta}_R$, (111), that under H_0 $qF_{R.Q}$ has an asymptotic χ^2 distribution. Hence the test statistics F_R and $F_{R.Q}$ have the same null asymptotic distributions and the difference of the test statistics converges to zero. Similar to reduction in dispersion test, reject H_0 if $F_{RQ} \geq F(\alpha, q, N-p-1)$ for an asymptotic level α test.

Refer to Hettmansperger and McKean (1983) for the discussion of the rank-based analogues of the Rao-scores test and the geometries of the three rank-based tests.

4.6 Goodness of Fit Determination

If a linear model has been fitted, summary statistics are needed to determine how good the fit was. One of the most widely used summary statistics for fits based on least square estimates is R^2, the proportion of the variance accounted for in fitting the predictors **V**.

$$R^2 = \frac{SS(0) - SS(\hat{\beta})}{SS(0)}, \tag{134}$$

where SS(0) is the sum of square error of **Z** when no predictors are fit and $SS(\hat{\beta})$ is the sum of square error when $\hat{\beta}$ is fitted.

The analogue for the R-estimates is the proportion of dispersion accounted for in fitting the predictors V. Hence the dispersion accounted for by fitting **V** is given by

$$R_{R,1}^2 = \frac{D_R(0) - D_R(\hat{\beta})}{D_R(0)}, \tag{135}$$

where $D_R(0)$ is the dispersion of **Z** when no predictors are fitted. This summary statistic satisfies many of the properties that R^2 has such as its value is between 0 and 1. It has the value of 1 for a perfect fit and it increases if a supermodel, a model which contains the predictors **V** as a proper subset, is fit. It seems that this $R^2_{R,1}$ is an intuitively pleasing summary statistic. However, as Witt, Naranjo, and McKean (1994) showed, $R^2_{R,1}$ is not robust. The denominator of the $R^2_{R,1}$ has an unbounded influence function.

The numerator, though, is the numerator of the test statistic

$$F_R = \frac{\frac{D_R(0) - D_R(\hat{\beta})}{p}}{\frac{\hat{\tau}}{2}} , \qquad (136)$$

to test the hypotheses

$$H_0\text{: } \beta = 0 \quad versus \quad H_A\text{: } \beta \neq 0 . \qquad (137)$$

Witt, Naranjo, and McKean (1995) showed that the influence function for F_R is bounded in the **y**-space but is unbounded in the **x**-space. Note that for least square,

$$F_{LS} = \frac{\frac{SS(0) - SS(\hat{\beta})}{p}}{\frac{SS(\hat{\beta})}{n-p-1}} , \qquad (138)$$

for testing the hypotheses (137). Algebraically, it can be shown that

$$R^2 = \frac{F_{LS}}{F_{LS} + \frac{n-p-1}{p}} \qquad (139)$$

4.7 IBNR Reserve Estimation

Let Y_{ij} be non-negative random variables which represent the incremental incurred losses paid by development year j for claims occurred in accident year i where i = 1,...,n and j = 1,...,m. A random error E_{ij} is added to the multiplicative model used by the De Vylder linear regression method such that

$$Y_{ij} = X_i P_j E_{ij} . \tag{140}$$

This multiplicative model can be changed into an additive model by taking a logarithm of both sides of the equation.

$$\log Y_{ij} = \log X_i + \log P_j + \log E_{ij} . \tag{141}$$

With the logarithmic transformation, this model can be expressed as a linear model introduced above. Let

$$\boldsymbol{Z} = \mathbf{1}\alpha + \boldsymbol{V}\beta + \boldsymbol{e} \tag{142}$$

where

$$Z = \begin{bmatrix} \log Y_{11} \\ \vdots \\ \log Y_{1m} \\ \log Y_{21} \\ \vdots \\ \log Y_{2m} \\ \vdots \\ \log Y_{nm} \end{bmatrix} . \tag{143}$$

Let N be the number of observations in **Z**. In order to ensure the regression matrix not to be singular, there is one necessary constrain that the P_1 is set equal to 1.

As it was mentioned in the previous chapter that for rank-based estimation, β is being estimated with the assumption that intercept exists. The linear model (87) does not have an intercept. As it was described in Chapter III, one way to solve this problem is to estimate a dummy intercept after β is estimated. Then project the predicted values, $\hat{Z}$, to the plain without intercept. This problem can also be taken care of by setting $\alpha = \log X_1$ and then let

$$\beta_k = \log(\frac{X_{k+1}}{X_1}) , \quad \text{for } k = 1,\ldots,n-1$$
$$\text{and} \quad \beta_{l+n-1} = \log P_{l+1} , \qquad \text{for } l = 1,\ldots,m-1 . \tag{144}$$

In other word, let the first accident year incurred loss ultimate as the intercept and let the other accident year effects as the factors of the first accident year incurred loss ultimate. Then

$$\beta = \begin{bmatrix} \log X_2 - \alpha \\ \vdots \\ \log X_n - \alpha \\ \log P_2 \\ \vdots \\ \log P_m \end{bmatrix} . \tag{145}$$

Let p be the number of parameters in β. Then **V** is a N × p design matrix where

$$V = \begin{bmatrix} 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 1 & 0 & \dots & 0 \\ & & & & \vdots & & & & \\ 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 1 \\ 1 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 & 1 & 0 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 & 0 & 1 & 0 & \dots & 0 \\ & & & & \vdots & & & & \\ 1 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 1 \\ 0 & 1 & \dots & 0 & 0 & 0 & 0 & \dots & 0 \\ & & & & \vdots & & & & \\ & & & & \vdots & & & & \\ 0 & 0 & \dots & 1 & 0 & 0 & 0 & \dots & 1 \end{bmatrix} . \tag{146}$$

Then this log-multiplicative model becomes the linear model (93). Solve the equation (98) to get the value of $\hat{\beta}$. Then calculate the residuals **e** with the estimated β such that

$$\hat{e}_i = Z_i - v'\hat{\beta} , \tag{147}$$

and the intercept parameter α can be estimated as the median or the Walsh average of the residuals.

With the estimated α and β, the estimated X and P can be calculated as follow:

$$\begin{aligned} \tilde{X}_1 &= \exp[\hat{\alpha}] \ , \\ \tilde{X}_i &= \tilde{X}_1 \exp[\hat{\beta}_{i-1}] \ , \quad \text{for } i = 2,\ldots,n \ , \\ \tilde{P}_1 &= 1 \ , \\ \tilde{P}_j &= \exp[\hat{\beta}_{j+n-2}] \ , \quad \text{for } j = 2,\ldots,m \ . \end{aligned} \tag{148}$$

Record the constrain that is needed for P_j estimate that

$$\sum_{j=1}^{m} P_j = 1 \ . \tag{149}$$

Let

$$C = \sum_{j=1}^{m} \tilde{P}_j \ , \tag{150}$$

then

$$\hat{X}_i = C\tilde{X}_i \quad \text{and} \quad \hat{P}_j = \frac{\tilde{P}_j}{C} \ . \tag{151}$$

The IBNR reserves needed can be estimated as

$$IB\hat{N}R = \sum_{i=2}^{n} \hat{X}_i \sum_{j=1}^{i-1} \hat{P}_{m-j+1} .$$

4.8 Confidence Intervals

Let

$$\boldsymbol{K} = \mathbf{1} \cdot \boldsymbol{V}$$

$$\text{and} \quad \boldsymbol{\beta}_\varphi = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} . \tag{153}$$

Then the model (142) becomes

$$\boldsymbol{Z} = \boldsymbol{K}\boldsymbol{\beta}_\varphi + \boldsymbol{e} . \tag{154}$$

It was stated in the previous sections that for the rank-based estimation, the asymptotic confidence interval for a linear function $\mathbf{h}'\boldsymbol{\beta}_\varphi$ is given by

$$\boldsymbol{h}'\hat{\boldsymbol{\beta}}_\varphi \pm t_{(\alpha/2, N-p-1)} \hat{\tau}_\varphi \sqrt{\boldsymbol{h}'(\boldsymbol{K}'\boldsymbol{K})^{-1}\boldsymbol{h}} . \tag{155}$$

Let $\boldsymbol{h}_1' = [1\ 0\ \ldots\ 0]$, then $\boldsymbol{h}_1'\hat{\boldsymbol{\beta}}_\varphi$ is the estimated intercept and

$$SE_{\varphi 1} = \hat{\tau}_\varphi \sqrt{\boldsymbol{h}_1'(\boldsymbol{K}'\boldsymbol{K})^{-1}\boldsymbol{h}_1} \tag{156}$$

is the standard error for the estimated intercept.

Theorem 2. If

$$X_n \xrightarrow{D} N(\mu, \sigma_n^2)$$

and if g is differentiable at μ, then $g(X_n)$ is $N(g(\mu),[g'(\mu)]^2\sigma_n^2)$.

An approximation for the variance of $g(X_n)$ is given by Theorem 2 that if

$$Var[X_n] = \sigma_n^2 \text{ and } X_n \to \mu$$

$$\text{then } Var[g(X_n)] \approx (g'(\mu))^2\sigma_n^2 .$$

Recall that the estimated first accident year ultimate

$$\tilde{X}_1 = \exp[\boldsymbol{h}_1'\hat{\beta}_\varphi] . \tag{159}$$

Let $\tilde{X}_1$ be a function of $\boldsymbol{h}_1'\hat{\beta}_\varphi$, then

$$\begin{aligned} \tilde{X}_1 &= g(\boldsymbol{h}_1'\hat{\beta}_\varphi) = \exp[\boldsymbol{h}_1'\hat{\beta}_\varphi] \\ \text{and } g'(\boldsymbol{h}_1'\hat{\beta}_\varphi) &= \exp[\boldsymbol{h}_1'\hat{\beta}_\varphi] . \end{aligned} \tag{160}$$

Based on Theorem 2 the standard error for $\tilde{X}_1$ is

$$g'(\boldsymbol{h}_1'\hat{\beta}_\varphi)SE_{\varphi 1} = \exp[\boldsymbol{h}_1'\hat{\beta}_\varphi]SE_{\varphi 1} . \tag{161}$$

From (151) and Theorem 2, the standard error for the estimated first year accident ultimate, $\hat{X}_1$ is

$$C\exp[\boldsymbol{h}_1'\hat{\beta}_\varphi]SE_{\varphi 1} . \tag{162}$$

and the asymptotic confidence interval for $\hat{X}_1$ can be calculated as

$$h_1'\hat{\beta}_\varphi \pm t_{(\alpha/2,N-p-1)} C \exp[h_1'\hat{\beta}_\varphi'] \hat{\tau}_\varphi \sqrt{h_1'(K'K)^{-1} h_1} \ . \tag{163}$$

Let $h_2' = [1 \;\; 1 \;\; 0 \; \cdots \; 0]$, then

$$\tilde{X}_2 = \exp[h_2'\hat{\beta}_\varphi] \ . \tag{164}$$

Then the asymptotic confidence interval for the estimated second accident year ultimate, $\hat{X}_2$, can be calculated as

$$h_2'\hat{\beta}_\varphi \pm t_{(\alpha/2,N-p-1)} C \exp[h_2'\hat{\beta}_\varphi'] \hat{\tau}_\varphi \sqrt{h_2'(K'K)^{-1} h_2} \ . \tag{165}$$

For properly setting the h vector, the asymptotic confidence intervals for other estimated accident year ultimates and the estimated loss development year paid portions can then be calculated by the same manner as described above.

Note that the estimated IBNR can be written as a function of the estimated parameters,

$$I\hat{BNR} = \sum_{i=1}^{n-1} \exp[\hat{\beta}_1 + \hat{\beta}_{i+1}] \sum_{j=1}^{i} \exp[\hat{\beta}_{p-j+1}] \tag{166}$$

where p is the number of parameters including the intercept. The multi-dimensional Δ method provides an approximation for the variance of IBNR. Let IBNR be a function of β,

$$IBNR = g(\beta_1, \beta_2, \ldots, \beta_p) , \tag{167}$$

then

$$Var[IB\hat{N}R] \approx \sum_{i=1}^{p} Var[\hat{\beta}_i](\frac{\partial g}{\partial \hat{\beta}_i})^2 + \sum_{i<j} Cov[\hat{\beta}_i,\hat{\beta}_j] \frac{\partial g}{\partial \hat{\beta}_i} \cdot \frac{\partial g}{\partial \hat{\beta}_j} . \tag{168}$$

Let **d** be the partial derivative matrix of the parameters such that

$$\boldsymbol{d} = \begin{bmatrix} \frac{\partial g}{\partial \beta_1} \\ \frac{\partial g}{\partial \beta_2} \\ \vdots \\ \frac{\partial g}{\partial \beta_p} \end{bmatrix} ,$$

where $IBNR = g(\beta_1, \beta_2, \ldots, \beta_p)$, then the variance of IBNR is approximated by

$$Var[IBNR] \approx \tau_\varphi^2 \boldsymbol{d}'(\boldsymbol{K}'\boldsymbol{K})^{-1}\boldsymbol{d} .$$

The standard error for IBNR can also then be calculated. The confidence interval for the IBNR will be

$$IB\hat{N}R \pm t_{(\alpha/2,N-p-1)} \hat{\tau}_\varphi \sqrt{\boldsymbol{d}'(\boldsymbol{K}'\boldsymbol{K})^{-1}\boldsymbol{d}} . \tag{171}$$

Sometimes the formula for the standard error of an estimate can not be derived analytically, then an estimate can be obtained by means of the bootstrap method. This means the standard error for IBNR estimate can be obtained by bootstrap method. Let

$$\hat{Z}_i = K_i\hat{\beta} \tag{172}$$

where K_i is the ith row of **K**, and let r_i be the residuals corresponding to the estimated Z_i such that

$$r_i = Z_i - \hat{Z}_i \ . \tag{173}$$

Residuals are randomly picked and added to $\hat{Z}_i$ to form a new dependent matrix **Z***. A new IBNR then be estimated with the **Z***. Repeat the same process for many times, such as 1000 times. The standard deviation of the new estimated IBNR's will be the bootstrap estimate of standard error for the IBNR.

4.9 Numerical Example

The same data set displayed on the previous chapter will be used to estimate the X_i and P_j. Note that a computer software titled "RGLM", a Robust General Linear Model package developed by Kapenga, McKean, and Vidmar was used to estimate the value of α and β for this linear model. In addition to calculating the R-estimate, RGLM also provides results estimated by least square for comparison. Table 30 and Table 31 show the rank-based and least square estimates for α and β calculated by the RGLM. The estimated X's and P's are then calculated using the estimated values of α and β.

Table 30

Comparison of R-based and Least Square Estimates

Parameter	R-estimate	Least Squares
α	5.555316	5.550020
β_1	0.057406	0.058802
β_2	0.130695	0.129618
β_3	0.083298	0.083722
β_4	0.150622	0.153760
β_5	0.142007	0.141943
β_6	-0.823074	-0.822846
β_7	-1.617730	-1.619740
β_8	-2.780570	-2.777430

Table 31

Estimated X's and P's Comparison

	Rank-based		Least Square	
	X	P	X	P
1	735.97	0.3506	733.62	0.3507
2	779.45	0.4041	778.04	0.4141
3	838.72	0.1540	835.14	0.1540
4	799.90	0.0695	797.68	0.0694
5	855.60	0.0217	855.55	0.0218

Table 31 indicates that the estimated X's and P's using rank-based and least square method are almost identical for this data set. Figure 1 displays the q-q plot for the least

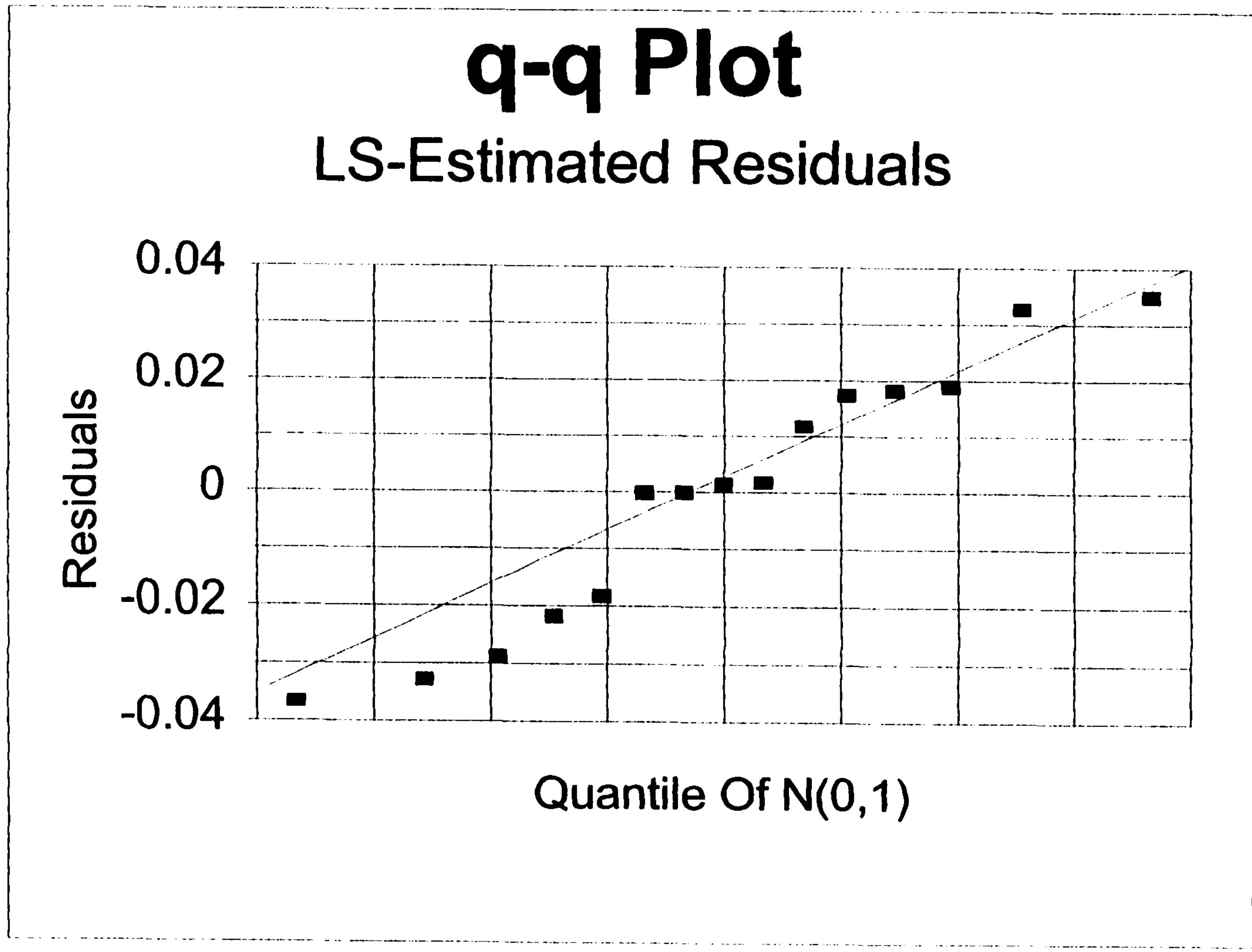

Figure 1. q-q Plot: LS-Estimated Residuals.

square estimated residuals. Figure 2 displays the q-q plot for the rank-based estimated residuals. Both figures indicate the errors of this data set are normally distributed.

Using the estimated X's and P's above, IBNR reserves needed can be calculated as shown in Table 32:

Table 32

Rank-based Estimated IBNR Reserves Needed

Accident Year	1991	1992	1993	1994	Total
R-Estimate	17	77	196	556	845
Least Square	17	76	196	556	844

The rank-based estimated total reserves needed, 845, is identical to the one estimated by the chain ladder method and is very close to the one estimated by the least square estimation. This indicates that least square and rank-based estimations both generate good results when the errors have normal distribution.

The calculated R^2 for this estimation is 98.71% which indicates this log-multiplicative model is a good fit for this data set.

The estimated scale, $\hat{\tau}_\varphi$, for this data set is 0.075. The standard error for the estimated parameters are shown in Table 33. With the estimated scale, the 95% confidence intervals for the estimated accident year ultimates are calculated as shown in Table 34.

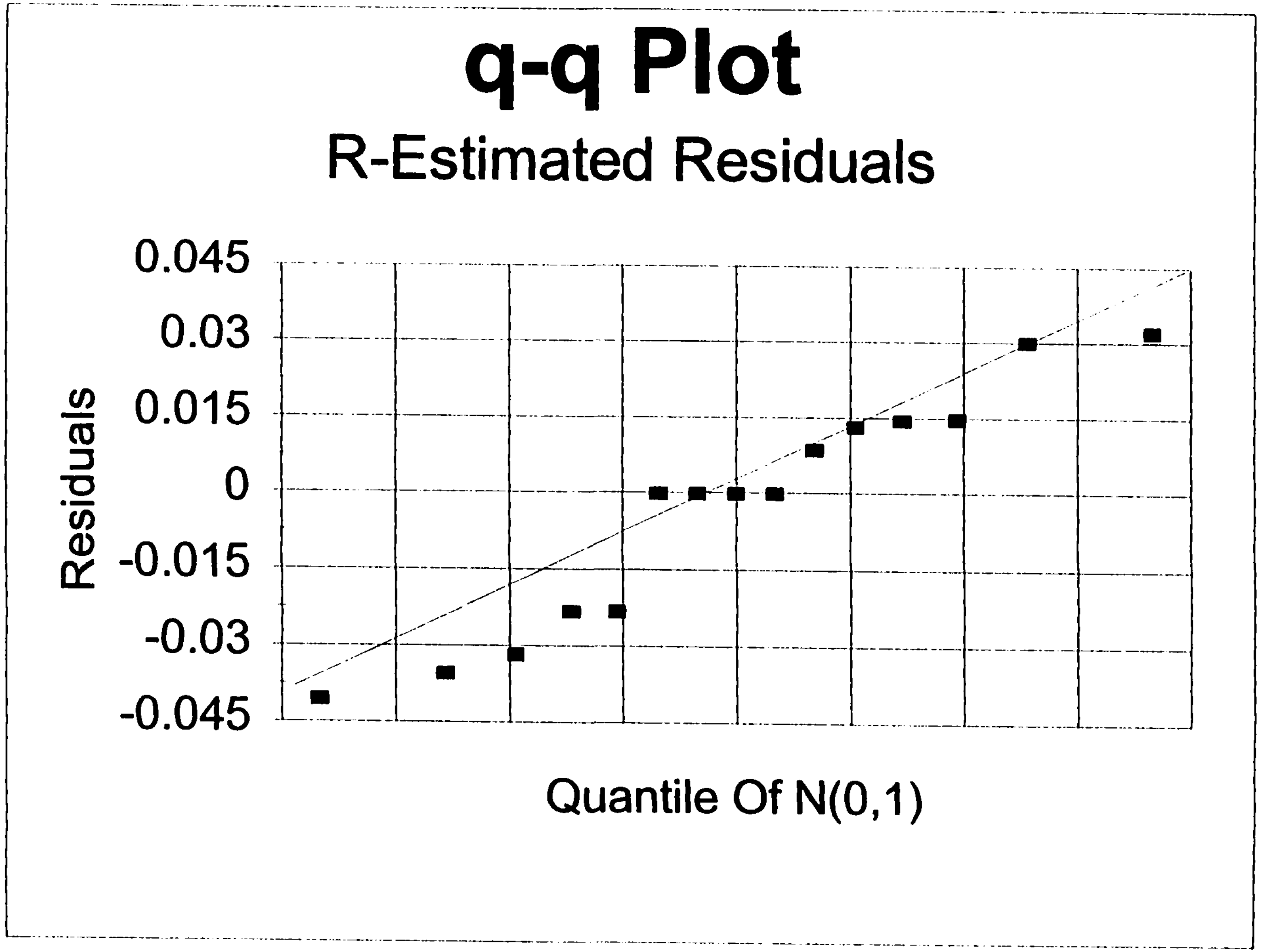

Figure 2. q-q Plot: R-Estimated Residuals.

Table 33

Estimated Standard Error

β_i	Standard Error
1	0.0504
2	0.0530
3	0.0593
4	0.0690
5	0.0910
6	0.0530
7	0.0593
8	0.0690
9	0.0910

Table 34

Confidence Intervals for Estimated Ultimates

Accident Year	Confidence Interval	
	Upper Limit	Lower Limit
1990	801	671
1991	848	711
1992	916	761
1993	881	719
1994	966	746

The standard error for the estimated IBNR approximated based on the Δ method is 35.66. Then the 95% confidence interval for the estimated IBNR is (753,937). The standard error estimated by the bootstrap method is 30.99 and the 95% confidence interval for the estimated IBNR is (765,925).

4.10 Hypotheses Testing

Let

$$Z = K\beta + e \tag{174}$$

where **K** = [**1**,**V**] and **β** includes the intercept parameter. Let

$$M = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{175}$$

and consider the hypotheses

$$H_0 : M\beta = 0 \quad versus \quad H_A : M\beta \neq 0 . \tag{176}$$

Under the null hypothesis, there is no significant accident year effect for this data set. The calculated value of F_R given by the RGLM is 3.618. The $F_{(.05;5,6)} = 4.39$ and $F_{(.1;5,6)} = 3.11$ which indicates that there is significant accident year effect at 0.1 level of significance but not significant at 0.05 level of significance.

Consider another hypotheses that

$$H_0 : \boldsymbol{M\beta} = 0 \quad versus \quad H_A : \boldsymbol{M\beta} \neq 0 \text{ ,} \tag{177}$$

where

$$\boldsymbol{M} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{178}$$

Under this null hypothesis, there is no loss development effect for this data set. The calculated F_R given by RGLM is 106.771. The $F_{(.05;4,6)} = 4.53$ and $F_{(.1;4,6)} = 3.18$ which indicates that the loss development effect for data set is significant.

For this example, the results of the hypotheses suggesting that the loss development effect strongly predicting the amounts of loss incurred at different loss development period. However, the accident year effect are statistically significant in predicting the loss incurred amount only at 0.10 or larger level of significance.

One can use the same procedure of hypotheses testing to evaluate all parameters together, each individual parameter, or any types of combination by setting a proper design matrix **M**.

The F values calculate by least square method given by the RGLM is 25562.9 for the accident year effects and 1695.73 for the loss development effects. This indicates that based on least square method, both accident year and loss development effects are significantly predicting the loss incurred amounts for this data set.

4.11 Rank-based Estimation vs Least Squares Estimation

If the estimates and the hypotheses testing results for the rank-based method and the classical least square method are so close, why would anyone choose to use the rank-based method rather than the classical least square method? As it was mentioned above, both least square estimation and rank-based estimation generate acceptable results as long as the errors for the linear model distributed fairly normal and free of outlier. However, if the error distribution is skewed toward right or left as well as error with long tail, or if there is one or more significantly large outlier, the results estimated by least square will be distorted. As it will be shown in the following chapter with greater detail, rank-based estimation is robust to outlier and is relatively more efficient when the error is long tailed comparing to least square estimation.

As it was mentioned in the previous chapter, for rank-based estimation, the influence of outlier enters in a linear rather than quadratic fashion, the estimates generated by the rank-based estimation are more robust than the least square estimates. Unlike the least square estimation, the predicted total IBNR reserve estimated by rank-based estimation for a data set contaminated with outlier change very little compare to the one estimated without outlier. In addition, the rank-based estimated IBNR reserve for each accident year is not distorted despite there is an outlier in the data set. It is rather more important that the incurred portion for each loss development year has not been changed due to the outlier. This means the loss development distribution which is used to predict the loss incurred for future accident years will not be changed due to significantly large outlier.

Chapter V discusses the robustness for the rank-based method in greater detail and also presents a sensitivity study for this method relative to least square estimation.

CHAPTER V

RANK-BASED ESTIMATION EVALUATION

5.1 Influence Function

Consider the linear model from the last two chapters

$$\boldsymbol{Z} = \alpha \mathbf{1} + \boldsymbol{V}\beta + \boldsymbol{e}. \tag{179}$$

As it was stated in Chapter IV that the influence function of the R-estimate is given by

$$IF(v, z; \hat{\beta}_R) = \tau \Sigma^{-1} \varphi(F(z)) v \ . \tag{180}$$

This indicates that the influence function for $\hat{\beta}_R$ is bounded in the **y**-space but not in the **x**-space. Since the independent variable **V** for the linear model (179) is a design matrix, it should not have contamination in the **x**-space, so there should not be any distortions on the estimates.

A sensitivity analysis has been done to test this rank-based method with the same data set that was tested with other methods. As the same way De Vylder's least squares method was tested, as shown in Table 35, a loss incurred was substituted with a significantly large number to see if this contamination would distort the results. The incremental loss incurred for the 2nd loss development year for accident year 1992 was changed from 344 to 1000.

Table 35

Contaminated Data Set

Loss Development Year	Accident Year				
	1990	1991	1992	1993	1994
1	250	267	298	289	300
2	300	315	1000	312	
3	117	120	124		
4	50	55			
5	16				

Estimating the IBNR reserves for the data above with rank-based method generates the results in Table 36.

Table 36

Distorted Estimates

Accident Year	1990	1991	1992	1993	1994
Ultimate Incurred	729	778	861	772	866

Development Year	1	2	3	4	5
Portion Paid	0.346	0.409	0.154	0.069	0.022

With the estimated value of X_i and P_j, future incurred losses can be estimated as shown in Table 37.

The results displayed above shows that a large contaminated data that would distort the estimates for the chain ladder method, De Vylder's method, and other non-parametric iteration methods do not distort the rank-based estimates. The total reserve estimate for all accident years 851 is not significantly different from the one estimated without the contamination 845. It is more important that the estimated loss incurred portion for the loss development years are not changed due to the contamination.

Table 37

Distorted Predicted IBNR Reserves

Loss Development Year	Accident Year					
	1990	1991	1992	1993	1994	Total
2					354	
3				119	133	
4			60	53	60	
5		17	19	17	19	
Total IBNR Reserves	0	17	79	189	566	851

In order to further test the sensitivity of the rank-based method, the contaminated data are changed from 1,000 to 5,000 and 10,000. Table 38 and Table 39 below display the results of the two contaminated data sets estimated with the rank-based method.

Tables 38 and 39 showed that contaminations did not distort the results estimated by the rank-based method which confirmed that the influence function for the R-estimates is bounded in the y-space. And they also indicate that the magnitude of the contamination does not affect the results of the R-estimates.

Table 38

Estimated Total Incurred Comparison

	Accident Year				
	1990	1991	1992	1993	1994
Contamination	Estimated Total Incurred				
No	736	779	839	800	856
1000	729	778	861	772	866
5000	728	776	861	781	867
10000	728	776	861	775	867

Table 39

Estimated Portion Paid Comparison

	Accident Year				
	1990	1991	1992	1993	1994
Contamination	Estimated Total Incurred				
No	0.351	0.404	0.154	0.070	0.022
1000	0.346	0.409	0.154	0.069	0.022
5000	0.346	0.409	0.154	0.069	0.022
10000	0.346	0.409	0.154	0.069	0.022

This is very important for loss reserve analysis because the loss development pattern should not be changed just because there is one extraordinary large paid incurred or loss reserve set aside due to an unforeseeable catastrophe. It is also important that the loss development patterns and the total reserve needed for all accident years should not be distorted by the magnitude of the unusual large loss happen in one particular loss development year for one particular accident year.

Using the estimated incurred portion for loss development years, P, age-to-age loss development can then be calculated as shown in Table 40.

Table 40

Estimated Loss Development Factors Comparison

Loss Development Year	Loss Development Factors			
	Chain Ladder	De Vylder's LS	Least Square	Rank-Based
1 - 2	2.152	2.098	2.151	2.151
2 - 3	1.204	1.190	1.204	1.204
3 - 4	1.077	1.069	1.076	1.077
4 - 5	1.022	1.019	1.022	1.022

As Table 39 shown, the rank-based estimated loss development factors are identical to the selected loss development factors for the chain ladder method. On the other hand, the loss development factors estimated by De Vylder's method are different from both chain ladder and rank-based methods.

Since the rank-based method generates consistent estimation for the loss development pattern, the loss development factors estimated by rank-based method can be used for other loss reserve methods such as chain ladder method or Bonheutter-Ferguson method. These methods require actuaries to input the loss development factors in order to predict the IBNR. One can use the loss development factors estimated by the rank-based method along with chain ladder or Bonheutter-Ferguson methods to predict the IBNR reserves and compare them with the IBNR reserves estimated by the rank-based method to ensure the accuracy of the estimation.

5.2 Breakdown Study

It was shown that the influence function for the rank-based method was bounded in the y-based. However, it is not guarantee that a contaminated loss data will not distort the loss data or the loss incurred prediction. Keep in mind that loss incurred data sets are triangular matrixes with accident years as the columns and with loss development years as the rows. Because of the shape of the loss data triangles, it is not practical to make a general statement concerning the breakdown point for the rank-based method. It is not just the portion of the contaminated data contain in the data set but also the location of the contaminated data within the loss data triangle that will affect the outcome of the estimation. In addition, each contaminated data will affect both the row effect as well as the column effect.

A simulation study has been done to evaluate the results of the rank-based estimates when one contaminated data is placed at all possible location of the data set. Numerous sizes of data triangles have been tested to confirm the validity of the results for this analysis. As it was showed in the previous section, one contaminated data would not affect the results of the R-estimates. However, if it was placed at the corners of the data triangle the results of the estimation will be distorted. The data set that was used in the previous chapters displayed again in Table 41 to help illustrating the sensitive locations of a data triangle.

Table 41

Incremental Loss Incurred

Loss Development Year	Accident Year				
	1990	1991	1992	1993	1994
1	250	267	298	289	300
2	300	315	344	312	
3	117	120	124		
4	50	55			
5	16				

If there is a contaminated data at the upper right corner of the data triangle, the estimated P_j's will not be distorted. However, the X_i's will be estimated as shown in Table 42.

Table 42

Distorted Ultimate Incurred Estimates

Accident Year	1990	1991	1992	1993	1994
Ultimate Incurred	735	779	839	800	2852

Note that the ultimate incurred for accident year 1994 is much larger than the ultimates incurred for the prior accident years. It is because apply the exaggerated incurred loss, 1000, which is the only incurred loss data for that accident year to the estimated loss development factors will result with a extraordinary large ultimate incurred. In a practical point of view, it can not simply conclude that it is a bad estimation. Other factors could possibly cause the extraordinary large incurred loss for that accident year that leads to much larger ultimate incurred loss. However, it is still a good indicator for further investigation.

If a contaminated incurred loss, 1000, is placed at the lower left corner, both the estimated X_i's and the estimated P_j's are significantly distorted. The paid out porting for each loss development year P_j and the ultimate incurred loss for each accident year are estimated as shown in Table 43.

Table 43

Distorted P_j and X_i Estimates

Accident Year	1990	1991	1992	1993	1994
Ultimate Incurred	1720	1822	1960	1869	2000

Development Year	1	2	3	4	5
Portion Paid	0.150	0.173	0.066	0.030	0.581

Since there is only one data for the 5th loss development year, one contaminated data will shift the estimated P_j's enough to distort the estimated X_i's.

Numerous Simulations have been done to determine the largest portion of contaminated data within one row or one column without breaking down this rank-based method. Different numbers of outliers were randomly put within one row or one column of a data set and a check was performed to see if the estimated results were significantly distorted. The same process was done on data set with different sizes. The results of the simulations have consistently shown that if 50% or more data within one row or one column are contaminated, the results of the rank-based estimation will significantly be distorted. This means if there is one contaminated data, there will be 6 places within the triangular data set that can breakdown this rank-based method. Table 44 shows the 6 places that can breakdown the method.

Table 44

Places That Break Down R-estimation With 1 Outlier

Loss Development Year	Accident Year				
	1990	1991	1992	1993	1994
1				XXX	XXX
2				XXX	
3					
4	XXX	XXX			
5	XXX				

Note that the same 6 places apply to other sizes of data triangles if there is one contaminated data. For a 5x5 data triangle, 40% (6/15) of the time results of the estimation should be distorted if there is one outlier. For a 6x6 data triangle, 28.6% (6/21) of the time results of the estimation should be distorted if there is one outlier. The same calculation can apply to data set with different sizes. Let C1 be the percentage of time that the results of the R-estimation is distorted if there is one contaminated data in a nxn data triangle.

$$CI = \frac{6}{\frac{n(n+1)}{2}} \tag{181}$$

Equation (181) indicates that as the size of the data triangle increase, the probability that the rank-based method would breakdown with one outlier decreases.

Simulations have been done to confirm the calculation above. Total IBNR reserve was estimated with one contaminated data randomly put in the same 5x5 data set which has been used for illustration and repeating this process for 1,000 times. 41.5% of the results were distorted significantly.

Let C2 be the probability that a nxn data triangle would breakdown with two outlier.

$$C2 = 1 - \frac{\binom{ND-6}{2} - 2\times\left[\binom{3}{2} + \binom{4}{2}\right]}{\binom{ND}{2}} \tag{182}$$

where

$$ND = \frac{n(n+1)}{2} . \tag{183}$$

For a 5x5 data triangle, C2 is calculated as 77.2%. Two outliers were randomly put in the 5x5 data triangle to estimate the total IBNR reserve with rank-based method and repeating this process for 1000 times. 76.2% of the time the estimated results were significantly distorted. Note that when the data triangle is relative small, such as 5x5, the probability that the results are being distorted by just two outliers are very high, over 3/4 of the time. However, when the size of the data triangle increases, the probability that the results will be distorted by two outlier decreases significantly. For example, if there are two outlier within a 10x10 data triangle, the probability that this rank-based method breaking down is only 22%.

Note that the intercept parameter is estimated as the median of the Walsh averages of the residuals after the slope parameters are estimated and the breakdown point for the Walsh averages estimation is known to be around 0.29. This means the intercept estimation for the rank-based method will breakdown if there is more than 29% of contaminated data within the data triangle. Despite the location of the contaminated data, as long as there are 29% or more contaminated data within the data triangle, this rank-based method for sure will breakdown.

5.3 Confidence Interval

A simulation study has been done to evaluate the standard error approximated by the delta method. Known X_i's and P_j's are set to create a loss incurred data triangle which means the true IBNR is known also. Taking a logarithm transformation for the created **Y**, $Y_{ij} = X_i P_j$, to change the multiplicative model to a additive model

$$Z = \log Y . \quad (184)$$

Randomly pick errors, **e**, with normal distribution **N(0,** σ^2 **)** and added to **Z** to create **Z***
where

$$Z^* = \mathbf{1}\alpha + V\beta + e . \quad (185)$$

An IBNR is estimated based on **Z*** by the rank-based estimation and a 90% confidence interval for the estimated IBNR is constructed also. Then check if the true IBNR falls within this constructed IBNR confidence interval. Repeat this process for 1000 times to count the number of estimated IBNR confidence intervals contain the true IBNR. After 1000 simulations have been done, there were 896 out of 1000, 89.6%, estimated IBNR confidence intervals contain the true IBNR which indicated the standard error approximated by the delta method is quite accurate.

5.4 Comparison of Least Squares and Rank-based Estimates

Chapter IV showed that the results of the least square and rank-based estimates are very close if there is no extraordinary large loss data within the loss triangle. They both generate reasonable results compared to other loss reserving methods. However, if there is an outlier in the data triangle, the results estimated by the least square will be distorted. The second advantage of rank-based estimates over the least square estimates is the magnitude of the outlier does not affecting the results of the rank-based estimation. This estimation is very similar to how an actuary estimating IBNR reserves uses the chain ladder method. If there is a extraordinary large loss incurred in one particular loss development year for one particular accident year which generates an extraordinary aga-to-age loss development factor, the loss reserve analyst would usually down weight that particular age-to-age loss development factor while picking the selected loss development factors. This down weighting process will prevent the predicted IBNR reserves being over projected by that extraordinary loss incurred.

A contaminated loss data was placed in the same data set, Table 45, that has been tested with to compare the results. The incurred loss for the third loss development year for accident year 1992 has been replaced with a much larger loss incurred of 500.

Table 45

Contaminated Data Set

Loss Development Year	Accident Year				
	1990	1991	1992	1993	1994
1	250	267	298	289	300
2	300	315	344	312	
3	117	120	500		
4	50	55			
5	16				

Estimating the total IBNR reserves for the data set above with least square and rank-based methods. Then increase the outlier to 1,000, 5,000, and 10,000.

Table 46 displays the results estimated by the rank-based method and the least squares method. Note that the results estimated by the rank-based method were not distorted by the magnitude of the contaminated loss data. On the other hand, the results estimated by least square method were distorted significantly by the outlier, and the distortion increased as the magnitude of the outlier increased.

Table 46

R-estimates Vs LS-estimates

Contaminated data	Estimated Total IBNR Reserves	
	Rank-based	Least-square
No	845	844
500	862	1,065
1,000	861	1,216
5,000	865	1,724
10,000	862	2,037

5.5 Testing With Real Loss Data Sets

The rank-based method was tested with real loss data of various lines of businesses from various sizes of insurance companies and from large corporations which are self-insured or which have very high self-retention. Figures 3 through 7 display the predicted IBNR reserves estimated by the rank-based method from five different lines of insurance businesses along with the results of the chain ladder method for comparison. Note that the loss matrix under the chain ladder estimates is a cumulative loss data while the loss data matrix under the rank-estimates is an incremental loss data. Also the bold printed data are the predicted loss data and the light print are the historical loss data. Under the age-to-age loss development factors triangle, the Walsh average and median of the age-to-age loss development factors for the corresponding column are displayed along with the average.

Figure 3 displays the loss data as well as the estimations for Homeowner's Liability which loss development is very stable and has a relatively short tail. The rank-based estimation shows a very good fit, the R^2 of the estimation is over 98%. The 90% confidence interval for the IBNR reserves is between \$8.2 and \$10.8 million. Note that the total IBNR reserves estimated by the rank-based estimation is very close to the one estimated by the chain ladder method.

Figure 4 displays the loss data from Workers' Compensation which loss development is very stable but has a long tail. Again, the rank-based estimation has a very good fit, the R^2 of the estimation is over 96%. The results of the rank-based estimation are very close to the results of the chain ladder method.

Figure 5 displays the loss data from Other Liability for occurrence policies. The age 1 to age 2 loss development factors showed possible 1 or 2 outliers. Since there are outliers within the loss development factors, the Walsh average and the median do provide additional information that helps picking the selected factors. Despite of the outliers, the rank-based estimation still has a good fit, the R^2 of the estimation is almost 91%. The results of the rank-based estimation is still relatively close to the results of the chain ladder method.

Figure 6 displays the loss data from Medical Malpractice for occurrence policies.

Chain-Ladder Estimates: Homeowners

Age-To-Age Loss Development Factors:

	DY 1-2	DY 2-3	DY 3-4	DY 4-5	DY 5-6	DY 6-7	DY 7-8	DY 8-9	DY 9-10
1986	1.307	1.048	1.028	1.017	1.010	1.005	1.003	1.002	1.001
1987	1.379	1.035	1.028	1.016	1.010	1.006	1.003	1.002	
1988	1.346	1.044	1.024	1.012	1.009	1.005	1.003		
1989	1.386	1.040	1.027	1.008	1.008	1.004			
1990	1.339	1.044	1.023	1.014	1.008				
1991	1.316	1.049	1.024	1.013					
1992	1.269	1.034	1.015						
1993	1.286	1.059							
1994	1.258								
1995									
Average	1.321	1.044	1.024	1.013	1.009	1.005	1.003	1.002	1.001
Walsh Avg	1.322	1.044	1.025	1.013	1.009	1.005	1.003	1.002	1.001
Median	1.316	1.044	1.024	1.013	1.009	1.005	1.003	1.002	1.001
Select	1.300	1.045	1.024	1.013	1.009	1.005	1.003	1.002	1.001

											Est Ultimate	Est Reserves
1986	6,529,434	8,533,151	8,941,104	9,191,703	9,344,357	9,437,389	9,485,615	9,511,897	9,526,711	9,535,137	9,535,137	0
1987	6,539,819	9,016,637	9,332,038	9,588,922	9,739,104	9,837,336	9,895,435	9,923,895	9,939,756	**9,949,696**	9,949,696	9,940
1988	7,376,354	9,930,508	10,363,322	10,608,862	10,731,040	10,827,308	10,880,901	10,912,392	**10,934,217**	**10,945,151**	10,945,151	32,759
1989	9,154,123	12,686,923	13,195,481	13,553,875	13,664,771	13,772,370	13,821,182	**13,862,646**	**13,890,371**	**13,904,261**	13,904,261	83,079
1990	9,203,424	12,319,372	12,856,452	13,151,721	13,332,277	13,437,803	**13,504,992**	**13,545,507**	**13,572,598**	**13,586,171**	13,586,171	148,368
1991	10,626,497	13,982,638	14,664,517	15,017,120	15,207,705	**15,344,574**	**15,421,297**	**15,467,561**	**15,498,496**	**15,513,995**	15,513,995	306,290
1992	17,414,216	22,103,002	22,860,400	23,208,015	**23,509,719**	**23,721,307**	**23,839,913**	**23,911,433**	**23,959,256**	**23,983,215**	23,983,215	775,200
1993	11,302,805	14,533,920	15,388,681	**15,758,009**	**15,962,863**	**16,106,529**	**16,187,062**	**16,235,623**	**16,268,094**	**16,284,362**	16,284,362	895,681
1994	13,203,660	16,611,020	**17,358,516**	**17,775,120**	**18,006,197**	**18,168,253**	**18,259,094**	**18,313,871**	**18,350,499**	**18,368,849**	18,368,849	1,757,829
1995	12,787,316	**16,623,511**	**17,371,569**	**17,788,486**	**18,019,737**	**18,181,914**	**18,272,824**	**18,327,642**	**18,364,298**	**18,382,662**	18,382,662	5,595,346
											150,453,499	9,604,492

Rank-Estimates:

	DY 1	DY 2	DY 3	DY 4	DY 5	DY 6	DY 7	DY 8	DY 9	DY 10	Est Ultimate	Est Reserves
1986	6,529,434	2,003,717	407,953	250,599	152,654	93,032	48,226	26,282	14,814	8,426	9,535,137	0
1987	6,539,819	2,476,818	315,401	256,884	150,182	98,232	58,099	28,460	15,861	**9,100**	9,948,856	9,100
1988	7,376,354	2,554,154	432,814	245,540	122,178	96,268	53,593	31,491	**15,887**	**9,036**	10,937,315	24,923
1989	9,154,123	3,532,800	508,558	358,394	110,896	107,599	48,812	**33,223**	**18,376**	**10,452**	13,883,233	62,051
1990	9,203,424	3,115,948	537,080	295,269	180,556	105,526	**63,820**	**35,086**	**19,406**	**11,038**	13,567,154	129,351
1991	10,626,497	3,356,141	681,879	352,603	190,585	**133,719**	**73,689**	**40,511**	**22,407**	**12,745**	15,490,775	283,070
1992	17,414,216	4,688,786	757,398	347,615	**252,063**	**173,703**	**95,723**	**52,625**	**29,107**	**16,556**	23,827,792	619,777
1993	11,302,805	3,231,115	854,761	**379,387**	**208,786**	**143,880**	**79,288**	**43,590**	**24,110**	**13,713**	16,281,434	892,753
1994	13,203,660	3,407,360	**697,748**	**396,739**	**218,335**	**150,460**	**82,914**	**45,583**	**25,212**	**14,341**	18,242,353	1,631,333
1995	12,787,316	**4,113,925**	**746,207**	**424,292**	**233,498**	**160,910**	**88,673**	**48,749**	**26,963**	**15,336**	18,645,869	5,858,553
	0.68579	0.22064	0.04002	0.02276	0.01252	0.00863	0.00476	0.00261	0.00145	0.00082	150,359,917	9,510,910

90% Confidence Interval: 8,239,121 10,782,699

R-square: 98.41%

Figure 3. Chain Ladder Vs Rank-based Estimates, Homeowners.

Chain-Ladder Estimates: Workers' Comp

		DY 1-2	DY 2-3	DY 3-4	DY 4-5	DY 5-6	DY 6-7	DY 7-8	DY 8-9	DY 9-10		
1986		2.451	1.379	1.160	1.078	1.044	1.039	1.005	1.022	1.022		
1987		2.607	1.375	1.182	1.115	1.058	1.010	1.024	1.022			
1988		2.742	1.399	1.177	1.086	1.057	1.033	1.023				
1989		2.931	1.368	1.151	1.089	1.039	1.032					
1990		3.047	1.342	1.170	1.071	1.054						
1991		2.964	1.373	1.154	1.089							
1992		2.963	1.310	1.147								
1993		2.514	1.327									
1994		2.426										
1995												
	Average	2.738	1.359	1.163	1.088	1.050	1.029	1.017	1.022	1.022		
	Walsh Avg	2.738	1.360	1.163	1.088	1.050	1.032	1.019	1.022	1.022		
	Median	2.742	1.370	1.160	1.088	1.054	1.033	1.023	1.022	1.022		
	Select	2.740	1.360	1.163	1.088	1.050	1.032	1.023	1.022	1.022	Est Ultimate	Est Reserves
1986	21,371	52,388	72,254	83,825	90,335	94,307	97,994	98,454	100,635	102,887	102,887	0
1987	21,268	55,447	76,212	90,082	100,398	106,204	107,243	109,831	112,222	114,691	114,691	2,469
1988	22,190	60,834	85,104	100,151	108,812	114,967	118,790	121,558	124,232	126,965	126,965	5,407
1989	26,542	77,798	106,407	122,422	133,359	138,599	143,029	146,319	149,538	152,828	152,828	9,799
1990	32,977	100,494	134,886	157,758	168,991	178,065	183,763	187,990	192,125	196,352	196,352	18,287
1991	38,604	114,428	157,103	181,322	197,411	207,282	213,915	218,835	223,649	228,569	228,569	31,158
1992	42,466	125,820	164,776	189,045	205,681	215,965	222,876	228,002	233,018	238,144	238,144	49,099
1993	46,447	116,764	154,897	180,145	195,998	205,798	212,383	217,268	222,048	226,933	226,933	72,036
1994	41,368	100,344	136,468	158,712	172,679	181,313	187,115	191,418	195,630	199,933	199,933	99,589
1995	35,719	97,870	133,103	154,799	168,421	176,843	182,501	186,699	190,806	195,004	195,004	159,285
											1,782,307	447,130

Rank-Estimates:

	DY 1	DY 2	DY 3	DY 4	DY 5	DY 6	DY 7	DY 8	DY 9	DY 10	Est Ultimate	Est Reserves
1986	21,371	31,017	19,866	11,571	6,510	3,972	3,687	460	2,181	2,252	102,887	0
1987	21,268	34,179	20,765	13,870	10,316	5,806	1,039	2,588	2,391	2,657	114,879	2,657
1988	22,190	38,644	24,270	15,047	8,661	6,155	3,823	2,768	2,708	2,879	127,145	5,587
1989	26,542	51,256	28,609	16,015	10,937	5,240	4,430	3,058	3,130	3,327	152,544	9,515
1990	32,977	67,517	34,392	22,872	11,233	9,074	5,725	3,952	4,044	4,299	196,085	18,020
1991	38,604	75,824	42,675	24,219	16,089	10,085	6,655	4,594	4,702	4,997	228,443	31,032
1992	42,466	83,354	38,956	24,269	15,872	10,222	6,745	4,656	4,766	5,065	236,371	47,326
1993	46,447	70,317	38,133	25,598	15,819	10,188	6,723	4,641	4,750	5,049	227,665	72,768
1994	41,368	58,976	37,454	23,145	14,303	9,212	6,079	4,196	4,295	4,565	203,593	103,249
1995	35,719	61,904	35,014	21,637	13,372	8,612	5,683	3,923	4,015	4,267	194,144	158,425
	0.18398	0.31885	0.18035	0.11145	0.06887	0.04436	0.02927	0.02021	0.02068	0.02198	1,783,756	448,579
								90% Confidence Interval:			397,892	499,266
								R-square:		96.19%		

Figure 4. Chain Ladder Vs Rank-based Estimates, Workers' Comp.

Chain-Ladder Estimates: Other Liability Occurrence

		DY 1-2	DY 2-3	DY 3-4	DY 4-5	DY 5-6	DY 6-7	DY 7-8	DY 8-9	DY 9-10		
1986		3.367	1.988	1.331	1.244	1.162	1.046	1.036	1.021	1.024		
1987		2.836	2.160	1.482	1.270	1.115	1.041	1.051	1.033			
1988		6.599	2.015	1.748	1.148	1.097	1.034	1.026				
1989		4.614	2.120	1.396	1.227	1.140	1.041					
1990		8.138	2.156	1.473	1.123	1.091						
1991		13.773	2.074	1.411	1.133							
1992		7.217	1.909	1.361								
1993		9.989	1.821									
1994		6.689										
1995												
	Average	7.025	2.030	1.457	1.191	1.121	1.041	1.038	1.027	1.024		
	Walsh Avg	6.689	2.034	1.428	1.189	1.119	1.041	1.038	1.027	1.024		
	Median	6.689	2.044	1.411	1.187	1.115	1.041	1.036	1.027	1.024		
	Select	6.700	2.030	1.420	1.190	1.120	1.041	1.038	1.027	1.024	Est Ultimate	Est Reserves
1986	5,151	17,345	34,475	45,896	57,102	66,353	69,434	71,959	73,492	75,277	75,277	0
1987	7,095	20,122	43,472	64,431	81,853	91,277	94,996	99,852	103,185	**105,661**	105,661	2,476
1988	3,988	26,317	53,030	92,675	106,375	116,717	120,656	123,816	**127,159**	**130,211**	130,211	6,395
1989	6,160	28,424	60,266	84,159	103,224	117,670	122,518	**127,174**	**130,607**	**133,742**	133,742	11,224
1990	3,891	31,666	68,283	100,591	112,960	123,245	**128,298**	**133,173**	**136,769**	**140,052**	140,052	16,807
1991	3,355	46,209	95,820	135,184	153,166	**171,546**	**178,579**	**185,365**	**190,370**	**194,939**	194,939	41,773
1992	6,945	50,124	95,684	130,189	**154,925**	**173,516**	**180,630**	**187,494**	**192,556**	**197,178**	197,178	66,989
1993	8,127	81,180	147,839	**209,931**	**249,818**	**279,797**	**291,268**	**302,336**	**310,499**	**317,951**	317,951	170,112
1994	9,773	65,367	**132,695**	**188,427**	**224,228**	**251,135**	**261,432**	**271,366**	**278,693**	**285,382**	285,382	220,015
1995	11,156	**74,745**	**151,733**	**215,461**	**256,398**	**287,166**	**298,940**	**310,299**	**318,677**	**326,326**	326,326	315,170
											1,906,718	850,960

Rank-Estimates:

	DY 1	DY 2	DY 3	DY 4	DY 5	DY 6	DY 7	DY 8	DY 9	DY 10	Est Ultimate	Est Reserves
1986	5,151	12,194	17,130	11,421	11,206	9,251	3,081	2,525	1,533	1,785	75,277	0
1987	7,095	13,027	23,350	20,959	17,422	9,424	3,719	4,856	3,333	**2,679**	105,864	2,679
1988	3,988	22,329	26,713	39,645	13,700	10,342	3,939	3,160	**2,611**	**2,619**	129,046	5,230
1989	6,160	22,264	31,842	23,893	19,065	14,446	4,848	**4,508**	**3,177**	**3,187**	133,389	10,871
1990	3,891	27,775	36,617	32,308	12,369	10,285	**4,364**	**4,084**	**2,878**	**2,887**	137,458	14,213
1991	3,355	42,854	49,611	39,364	17,982	**16,015**	**6,082**	**5,693**	**4,012**	**4,024**	188,992	35,826
1992	6,945	43,179	45,560	34,505	**23,524**	**17,896**	**6,797**	**6,362**	**4,483**	**4,497**	193,748	63,559
1993	8,127	73,053	66,659	**56,039**	**34,416**	**26,182**	**9,944**	**9,307**	**6,559**	**6,579**	296,865	149,026
1994	9,773	55,594	**65,313**	**54,907**	**33,720**	**25,653**	**9,743**	**9,119**	**6,426**	**6,446**	276,695	211,328
1995	11,156	**61,619**	**73,831**	**62,069**	**38,119**	**28,999**	**11,014**	**10,308**	**7,264**	**7,287**	311,667	300,511
	0.03579	0.19771	0.23689	0.19915	0.12231	0.09305	0.03534	0.03307	0.02331	0.02338	1,849,001	793,243
								90% Confidence Interval:			600,444	986,042
								R-square:		90.95%		

Figure 5. Chain Ladder Vs Rank-based Estimates, Other Liability.

Chain-Ladder Estimates: Medical Malpractice Occurrence

		DY 1-2	DY 2-3	DY 3-4	DY 4-5	DY 5-6	DY 6-7	DY 7-8	DY 8-9	DY 9-10	Est Ultimate	Est Reserves
1986		5.963	2.828	1.955	1.574	1.211	1.206	1.117	1.051	1.047		
1987		6.501	3.026	2.213	1.545	1.362	1.164	1.115	1.065			
1988		6.155	2.820	1.992	1.611	1.324	1.172	1.094				
1989		4.771	3.402	2.234	1.476	1.257	1.169					
1990		7.142	3.668	1.740	1.412	1.254						
1991		10.248	2.273	1.833	1.456							
1992		1.722	1.897	1.670								
1993		8.289	3.672									
1994		7.511										
1995												
	Average	6.478	2.948	1.948	1.512	1.282	1.178	1.109	1.058	1.047		
	Walsh Avg	6.553	2.949	1.954	1.512	1.287	1.172	1.110	1.058	1.047		
	Median	6.501	2.927	1.955	1.511	1.257	1.171	1.115	1.058	1.047		
	Select	6.500	2.945	1.950	1.510	1.260	1.175	1.110	1.058	1.047	Est Ultimate	Est Reserves
1986	13,385	79,820	225,739	441,367	694,877	841,646	1,015,415	1,134,116	1,191,474	1,247,811	1,247,811	0
1987	7,517	48,869	147,866	327,229	505,672	688,742	801,977	894,190	952,153	996,904	996,904	44,751
1988	8,989	55,323	155,989	310,726	500,539	662,881	777,113	849,915	899,210	941,473	941,473	91,558
1989	9,422	44,957	152,943	341,743	504,502	634,219	741,632	823,212	870,958	911,893	911,893	170,261
1990	8,178	58,407	214,262	372,753	526,381	659,840	775,312	860,596	910,511	953,305	953,305	293,465
1991	10,504	107,642	244,617	448,295	652,720	822,427	966,352	1,072,651	1,134,864	1,188,203	1,188,203	535,483
1992	89,178	153,564	291,277	486,519	734,644	925,651	1,087,640	1,207,280	1,277,303	1,337,336	1,337,336	850,817
1993	7,172	59,450	218,299	425,683	642,781	809,905	951,638	1,056,318	1,117,584	1,170,111	1,170,111	951,812
1994	11,823	88,800	261,516	509,956	770,034	970,243	1,140,035	1,265,439	1,338,834	1,401,760	1,401,760	1,312,960
1995	11,877	77,201	227,355	443,343	669,448	843,505	991,118	1,100,141	1,163,949	1,218,655	1,218,655	1,206,778
											11,367,450	5,457,884

Rank-Estimates:

	DY 1	DY 2	DY 3	DY 4	DY 5	DY 6	DY 7	DY 8	DY 9	DY 10	Est Ultimate	Est Reserves
1986	13,385	66,435	145,919	215,628	253,510	146,769	173,769	118,701	57,358	56,337	1,247,811	0
1987	7,517	41,352	98,997	179,363	178,443	183,070	113,235	92,213	57,963	41,763	993,916	41,763
1988	8,989	46,334	100,666	154,737	189,813	162,342	114,232	72,802	45,924	40,427	936,266	86,351
1989	9,422	35,535	107,986	188,800	162,759	129,717	107,413	79,513	44,952	39,571	905,668	164,036
1990	8,178	50,229	155,855	158,491	153,628	133,459	112,454	81,780	46,234	40,700	941,007	281,167
1991	10,504	97,138	136,975	203,678	204,425	174,492	144,444	105,044	59,386	52,278	1,188,364	535,644
1992	89,178	64,386	137,713	195,242	223,802	181,617	150,342	109,333	61,811	54,412	1,267,837	781,318
1993	7,172	52,278	158,849	174,587	184,311	149,570	123,813	90,041	50,904	44,811	1,036,335	818,036
1994	11,823	76,977	156,866	233,208	246,197	199,791	165,386	120,274	67,996	59,857	1,338,377	1,249,577
1995	11,877	67,747	152,185	226,249	238,850	193,829	160,450	116,685	65,967	58,071	1,291,910	1,280,033
	0.00919	0.05244	0.11780	0.17513	0.18488	0.15003	0.12420	0.09032	0.05106	0.04495	11,147,491	5,237,925
									90% Confidence Interval:		4,661,524	5,814,326
									R-square:	94.74%		

Figure 6. Chain Ladder Vs Rank-based Estimates, Medical Malpractice.

The observed loss data does show some unstable loss development and outliers. However, the rank-based estimation still has a good fit, the R^2 of the estimation is almost 95%.

Figure 7 displays the occurrence loss data from Product Liability. The magnitude of the loss data is very small. The loss development pattern is unstable and has a relatively long tail. The R^2 of the estimation is only 69% which indicates this is not a very good fit. However, the total IBNR reserves estimated by the rank-based method is not far off from the result estimated by the chain ladder method.

The rest of the estimations, which are not displayed in this paper, show similar results as displayed in Figures 3 through 7. They all seem to have reasonably good results and the estimated IBNR reserves are close to the ones estimated by the chain ladder method.

Chain-Ladder Estimates: Product Liability Occurrence

		DY 1-2	DY 2-3	DY 3-4	DY 4-5	DY 5-6	DY 6-7	DY 7-8	DY 8-9	DY 9-10		
1986		1.818	1.727	1.942	1.167	1.105	1.397	1.012	1.021	1.062		
1987		2.036	1.400	2.372	1.018	1.113	1.094	1.059	1.026			
1988		1.913	1.373	1.181	1.493	1.077	1.191	1.028				
1989		2.261	1.917	1.205	1.063	1.052	1.018					
1990		2.815	1.132	1.198	1.573	1.389						
1991		6.316	1.108	1.038	1.246							
1992		1.283	1.712	1.059								
1993		1.042	1.121									
1994		3.353										
1995												
	Average	2.537	1.436	1.428	1.260	1.147	1.175	1.033	1.023	1.062		
	Walsh Avg	2.197	1.420	1.203	1.255	1.105	1.167	1.032	1.023	1.062		
	Median	2.036	1.386	1.198	1.207	1.105	1.143	1.028	1.023	1.062		
	Select	2.200	1.450	1.200	1.260	1.150	1.110	1.050	1.025	1.020	Est Ultimate	Est Reserves
1986	329	598	1,033	2,006	2,342	2,587	3,615	3,658	3,733	3,966	3,966	0
1987	247	503	704	1,670	1,700	1,892	2,070	2,192	2,248	**2,293**	2,293	45
1988	286	547	751	887	1,324	1,426	1,699	1,746	**1,790**	**1,825**	1,825	79
1989	134	303	581	700	744	783	797	**837**	**858**	**875**	875	78
1990	27	76	86	103	162	225	**250**	**262**	**269**	**274**	274	49
1991	19	120	133	138	172	**198**	**220**	**231**	**236**	**241**	241	69
1992	46	59	101	107	**135**	**155**	**172**	**181**	**185**	**189**	189	82
1993	119	124	139	**167**	**210**	**242**	**268**	**282**	**289**	**295**	295	156
1994	17	57	**83**	**99**	**125**	**144**	**160**	**167**	**172**	**175**	175	118
1995	50	**110**	**160**	**191**	**241**	**277**	**308**	**323**	**331**	**338**	338	288
											10,471	964

Rank-Estimates:

	DY 1	DY 2	DY 3	DY 4	DY 5	DY 6	DY 7	DY 8	DY 9	DY 10	Est Ultimate	Est Reserves
1986	329	269	435	973	336	245	1,028	43	75	233	3,966	0
1987	247	256	201	966	30	192	178	122	56	**160**	2,408	160
1988	286	261	204	136	437	102	273	47	**52**	**149**	1,946	200
1989	134	169	278	119	44	39	14	**16**	**17**	**49**	879	82
1990	27	49	10	17	59	63	**23**	**6**	**6**	**19**	279	54
1991	19	101	13	5	34	**12**	**14**	**3**	**4**	**11**	216	44
1992	46	13	42	6	**15**	**10**	**13**	**3**	**4**	**10**	161	54
1993	119	5	15	**12**	**15**	**11**	**13**	**3**	**4**	**10**	207	68
1994	17	40	**18**	**13**	**17**	**12**	**15**	**4**	**4**	**12**	151	94
1995	50	**46**	**40**	**30**	**38**	**27**	**33**	**8**	**9**	**26**	306	256
	0.16328	0.14923	0.13078	0.09687	0.12291	0.08878	0.10623	0.02687	0.02981	0.08523	10,519	1,012
									90% Confidence Interval:		59	1,965
									R-square:	68.99%		

Figure 7. Chain Ladder Vs Rank-based Estimates, Product Liability.

CHAPTER VI

CONCLUSION

The intention of this paper is to apply one of the practical uses of Robust Statistics to the insurance industry. Loss Reserve Analysis has been a repetitive task in the actuarial profession. Insurance companies rely on actuaries to estimate the IBNR reserves needed to determine its financial strength. This is why it is critical to accurately estimate the IBNR reserves needed routinely, such as quarterly, semi-annually, or annually. With the speed and capacity of today's computer, a tedious and repetitive loss reserve analysis can be accurately done with a statistical routine. However, classical statistical methods do not consistently deliver accurate results due to the following reasons:

1. Classical statistical methods are sensitive to outliers. Some lines of casualty property insurance tend to have outliers and unstable loss incurred pattern which will distort the classical statistical estimates.

2. Classical statistical methods need to have relatively large data sets to generate consistent results. For smaller or newly formed insurance companies, large data sets are not always available.

3. Classical statistical methods assume the distribution of the error to be normal. However, it is known that the distribution of IBNR reserves tend to have long tails.

As the results of this paper show, estimating IBNR reserves with the rank-based method did show positive results. This rank-based method generates consistent results despite outlier and contaminated data within the loss triangle. In addition, it also provides statistical inference for hypotheses testing. The R^2 value provided by this rank-based method is also a valid indicator for showing the goodness of fit. This rank-based method also estimates the standard error for the estimated IBNR from which the 100(1-α)% confidence interval for IBNR can be calculated.

The positive results for this rank-based method is a good start, but there is so much more studies that remain to be done. There are so many existing lines of insurance businesses that have different types of loss distributions. It is good that this rank-based method is distribution free and is able to be applied to different types of loss distributions. Though this rank-based method shows promising results with the tested data sets, it still needs to be tested with more actual loss data to determine the reliability of this method.

A few questions were raised by this research and need to be addressed by further research work.

1. What if the loss paid out model is not multiplicative but additive. A model such as

$$Y_{ij} = \alpha + \beta_i + \gamma_j + e_{ij} \qquad (186)$$

where α is the intercept parameter, β_i is the accident year effect, and γ_j is the loss development year effect.

2. What if there are negative numbers among the loss data. It is possible to have a negative number as the loss incurred data. A negative loss incurred number is indicating

that during that particular loss development period, the total amount loss incurred paid out is less than the total amount of salvage and subrogation received by the insurance company for accidents which occurred in that accident period. Since a claim has to completely pay out to claimants before the subrogation process starts, it is quite possible that for later loss development years, the total subrogation received is larger than the total loss paid out for one particular accident year.

If there are negative numbers among the loss data, this multiplicative model introduced in Chapter IV will not work because the logarithm of a negative number is undefined. In this case the additive model rather than the multiplicative model may be more appropriate.

3. As it was shown in the previous chapters, the iterative non-parametric methods were not robust to outliers. They did not generate reasonable results if there was contaminated data within the loss incurred triangle. A high breakdown rank-based estimation can be used rather than those introduced in Chapter III.

4. What if other score functions are used to make the iterative methods more robust to outliers. Some score functions are designed for error distributions with heavy right tails such as the Positively Skewed Winsorized Wilcoxon score function. They consist of a linear piece followed by a flat piece; hence, the resulting analysis is less sensitive to outliers on the right. Using some of these special score functions, the iterative methods can possibly be robust to the outliers in the **x**-space.

It would be very beneficial to the IBNR reserves estimation if the non-parametric iterative methods are robust to outliers because they are not restricted to estimate with only positive loss data.

There is a robust estimation routine called Least Median Square (LMS) which influence function is bounded in the **x**-space.

Consider the following linear model:

$$Y_i = X_i\beta + e_i \ , \tag{187}$$

and let

$$\hat{e}_i = Y_i - X_i\hat{\beta} \ . \tag{188}$$

LMS estimates $\hat{\beta}$ such that for all $-\infty < \beta < \infty$, $\hat{\beta}$ satisfies the following argument:

$$Minimize \ Med(\hat{e}^2) \ . \tag{189}$$

For the results of the preliminary study, when applied to the iterative method, this routine does generate promising results for estimating IBNR reserves. This estimating method is also robust to outliers, even outlier in the **x**-space. However, this LMS routine does have its disadvantages. There are no simple numerical routine that helps locate the $\hat{\beta}$ which generates the least median of the residual square. The process of finding that unique $\hat{\beta}$ can be tedious and time consuming. However, again, with the tremendous speed and capacity of today's computer, the problem concerning this LMS routine could be secondary. Further study of this robust routine could be very beneficial to loss reserve analysis. It provides another dimension in estimating the IBNR reserves.

Appendix A

Algorithm for Estimating IBNR Reserves With Iterative R-Estimate

APPENDIX A

Algorithm for Estimating IBNR Reserves With Iterative R-Estimate

Without loss of generality assume Wilcoxon score function φ(u) is used such that

$$a(i) = \varphi(\frac{i}{n+1})$$
$$\text{and} \quad \varphi(u) = \sqrt{12}(u - \tfrac{1}{2}) \; . \qquad (190)$$

$R(u_i)$ returns with the rank of u_i amoung all u's.

Let ε > 0 be the value to determine the convergency of the iteration.

Step 1: Let

$$\hat{P}_j = \frac{1}{m} \quad , j = 1,\ldots,m \qquad (191)$$

as the initial values of P.

Step 2: Let $P_j = \hat{P}_j$ be the independent variable and

$$\bar{P} = \frac{1}{m}\sum_{j=1}^{m} P_j \; .$$

For i = 1,...,n

Step 2a: Solve for $\tilde{X}_i$ such that

$$0 \doteq \sum_{j=1}^{m} (P_j - \bar{P}) a[R(Y_{ij} - \tilde{X}_i P_j)] \ . \tag{193}$$

Step 2b: Let e_i be the residuals

$$e_j = Y_{ij} - P_j \tilde{X}_i \ , \tag{194}$$

for j = 1,...,m.

Solve for a dummy intercept parameter

$$\hat{\alpha} = med(e_j) \ . \tag{195}$$

Step 2c: Calculate the projected value of Y_{ij} such that

$$\hat{Y}_{ij} = \hat{\alpha} + P_j \tilde{X}_i \ , \tag{196}$$

for j = 1,...,m.

Estimate X_i again with the following equation

$$\hat{X}_i = \frac{\sum_{j=1}^{m} \hat{Y}_{ij} P_j}{\sum_{j=1}^{m} P_j^2} \ . \tag{197}$$

Step 3: Let $X_i = \hat{X}_i$ be the independent variable

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \ . \tag{198}$$

For j = 1,...,m

Step 3a: Solve for $\tilde{P}_j$ such that

$$0 \doteq \sum_{j=1}^{m} (X_i - \bar{X}) a[R(Y_{ij} - X_i \tilde{P}_j)] \ . \tag{199}$$

Step 3b: Let e_i be the residuals

$$e_i = Y_{ij} - \tilde{P}_j X_i \ , \tag{200}$$

for i = 1,...,n.

Solve for a dummy intercept parameter

$$\hat{\alpha} = med(e_i) \ . \tag{201}$$

Step 3c: Calculate the projected value of Y_{ij} such that

$$\hat{Y}_{ij} = \hat{\alpha} + \tilde{P}_j X_i \ , \tag{202}$$

for i = 1,...,n.

Estimate P_i again with the following equation

$$\hat{P}_j = \frac{\sum_{i=1}^{n} \hat{Y}_{ij} X_i}{\sum_{i=1}^{n} X_i^2} \ . \tag{203}$$

Step 4: Check if

$$\max_{1 \le i \le m} \left\{ \hat{P}_j - P_j \right\} < \varepsilon \ . \tag{204}$$

If true then stop.

If false then go back to step 2.

Appendix B

Deficiency of Non-parametric Iterative Methods: An Numerical Illustration

APPENDIX B

Deficiency of Non-parametric Iterative Methods
An Numerical Illustration

Consider the same data set being tested in Chapter II:

Table 47

Contaminated Data Set

Loss Development Year	Accident Year				
	1990	1991	1992	1993	300
1	250	267	298	289	
2	300	315	1000	312	
3	117	120	124		
4	50	55			
5	16				

Incurred loss for the second development year for accident year 1992 has been changed from 344 to 1000. With the $P_j = 0.2$ as the initial values, X_i's were estimated with the weighted L-1 norm estimation as shown in Table 48.

Table 48

Distorted X_i Estimates

Accident Year	1990	1991	1992	1993	1994
Ultimate Incurred	750	968	2810	1597	1500

Then P_j's were estimated with the estimated X_i's as the independent variables. The estimated X for accident year 1992 is 2810 which is an outlier in the **x**-space. Graph 1 displayed the actual data along with the predicted line. Note that the predicted line was pulled by the outlier in the **x**-space. Thus the predicted P was smaller than it was supposed to be. The P_j's were estimated as shown in Table 49.

Table 49

Distorted P_j Estimates

Development Year	1	2	3	4	5
Portion Paid	0.154	0.348	0.065	0.061	0.021

By the iteration process, the X_i then are estimated again with the distorted estimated values of P_j as the independent variables. This means there are outlier in the **x**-space again and the estimated results of P_j will be distorted.

Figure 8 demonstrates another linear data set which has no outlier. The rank-based estimates and the least squares estimates are almost identical. However, if there are outliers in the y-space, the rank-based estimates will be different from the least squares estimates. Figure 9 shows the results estimated by rank-based and least squares when there is one outlier in the y-space.

As Figure 9 shown, the rank-based estimations are robust to outlier in the y-space. Unlike the least square estimation, the outlier in this data set did not pulled the slope estimated by the rank-based method larger than it is supposed to be.

Figure 10 shows the results estimated by rank-based and least squares when there is one outlier in the **x**-space. As Figure 10 shown, the rank-based estimations are not robust to outlier in the **x**-space. Like the least square estimation, the outlier in this data set pulled the slope estimated by the rank-based method larger than it is supposed to be. This confirms that the influence function for the R-estimates is bounded in the **y**-space but is not bounded in the **x**-space.

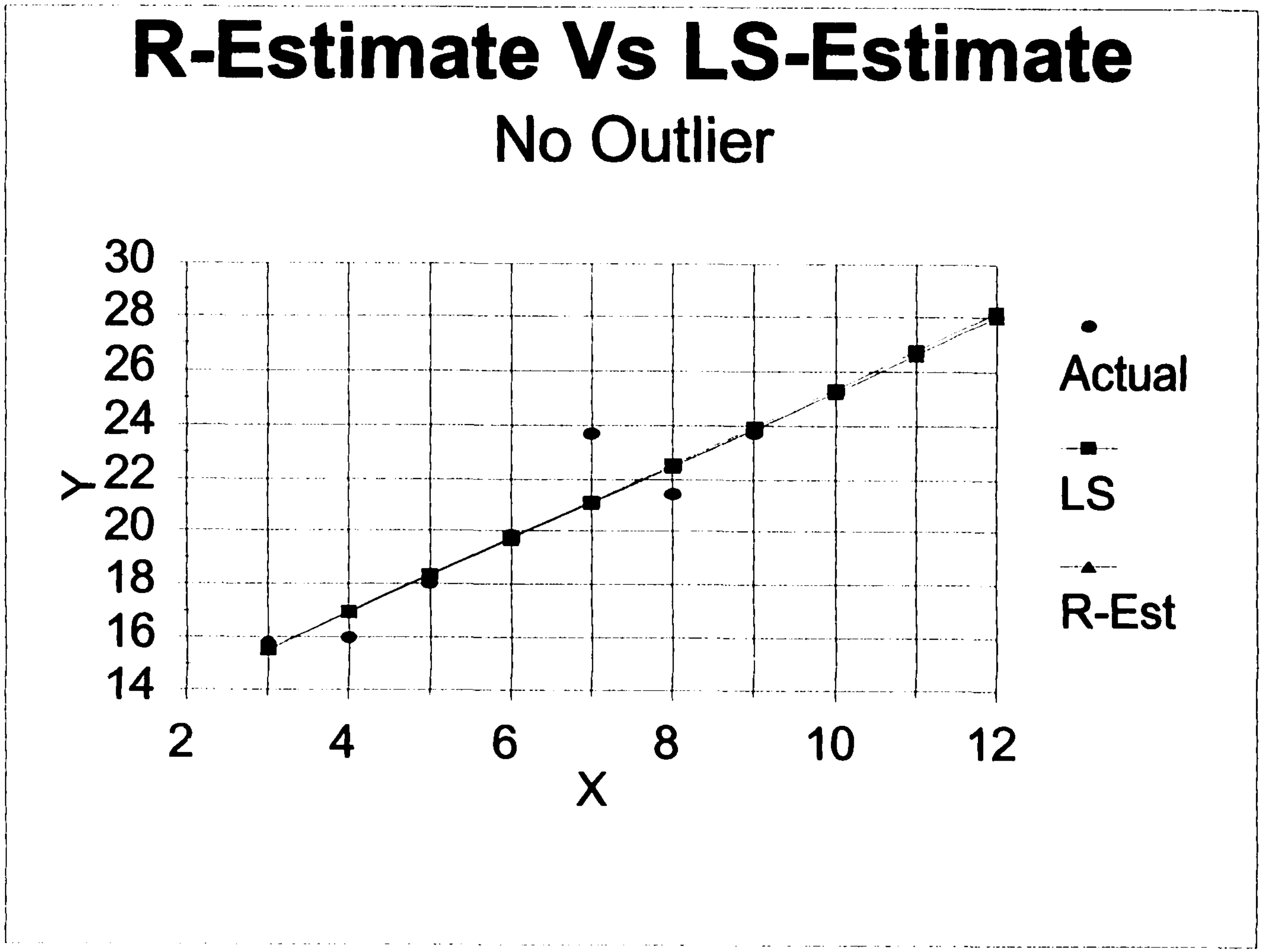

Figure 8. R-estimate Vs LS-estimate, No Outlier.

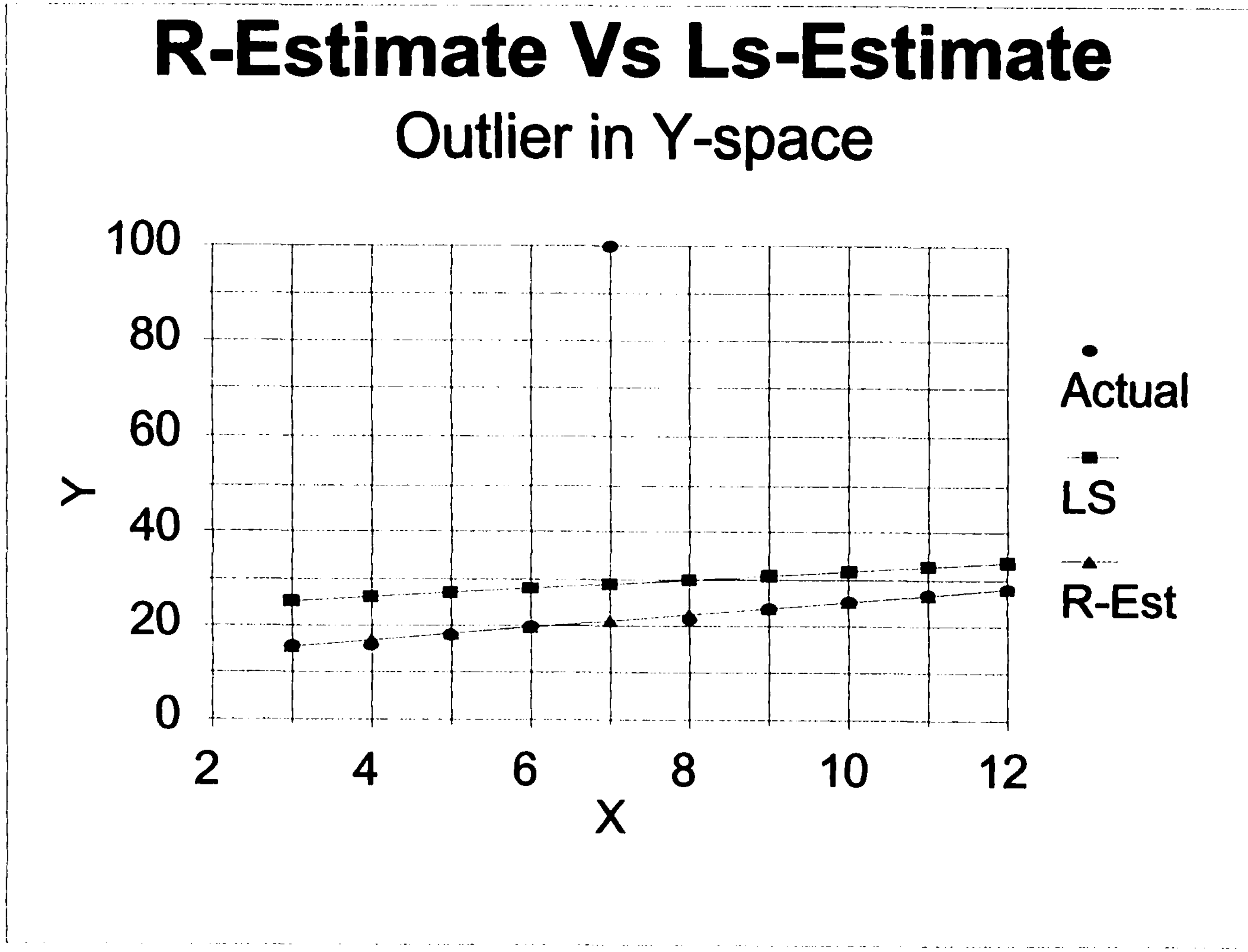

Figure 9. R-estimate Vs LS-estimate, Outlier in Y-space.

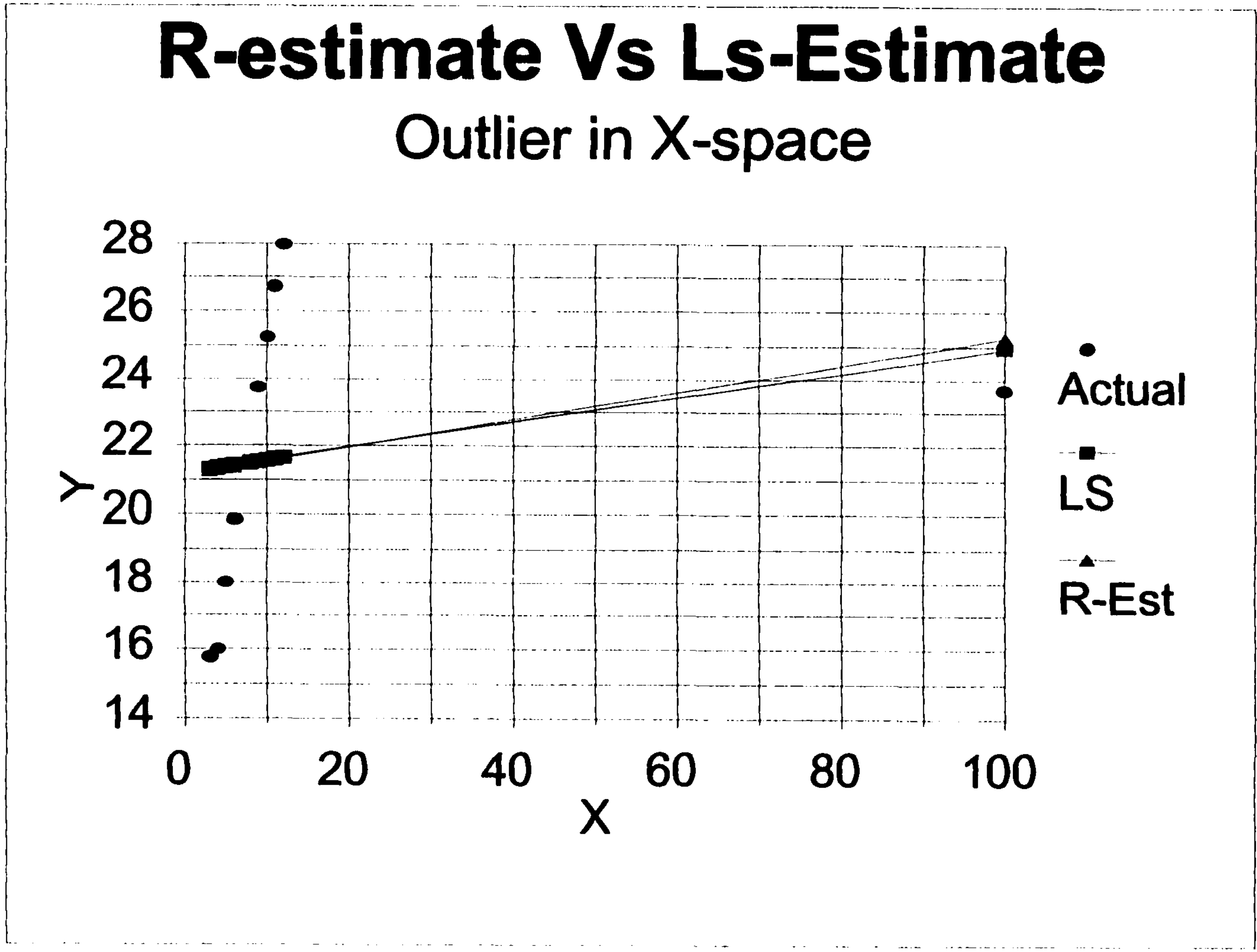

Figure 10. R-estimate Vs LS-estimate, Outlier in X-space.

Appendix C

Assumptions and Theory

APPENDIX C

Assumptions and Theory

C.1 Assumptions for Asymptotic Theory

Certain assumptions on the distribution of the errors, the design matrix, and the scores are needed for the development of the asymptotic theories for the rank-based linear model. The required assumptions may be differ for each of the asymptotic theory. They are all placed together in this section for easy reference.

Let **H** denote the projection matrix where

$$H = V(V'V)^{-1}V' \tag{206}$$

project onto the column space of **V**. The asymptotic theory assumes that the design matrix **V** for the linear model

$$Z = \alpha\mathbf{1} + V\beta + e \tag{207}$$

is imbedded in a sequence of design matrixes which satisfy the following two properties. Note that **Z** is a N × 1 vector of observations. Let h_{iiN} denote the leverage values of the diagonal entries of the N × N projection matrix **H**.

$$(D.2) \quad \lim_{N\to\infty} \max_{1\le i\le N} h_{iiN} = 0 \tag{208}$$

$$(D.3)\quad \lim_{N\to\infty} \frac{V'V}{N} = \Sigma \,, \tag{209}$$

where Σ is a p × p positive definite matrix. The assumption (D.2) is known as Huber's condition. Huber (1973) showed that the assumption (D.2) was a necessary and sufficient design condition for the least squares estimates to have an asymptotic normal distribution provided the error e_i are iid with finite variance. The Huber's condition also implied the following assumption,

$$(N.1)\quad \max_{1\le i\le N} \frac{v_{ij}^2}{\sum_{j=1}^{N} v_{jk}^2} \to 0 \quad \text{for all } k = 1,\dots,p \,. \tag{210}$$

The score function φ(u) which is used by the rank-based estimation needs to satisfy the following assumptions,

(S.1) φ(u) is a nondecreasing, square-integrable, and bounded function which is defined on the interval (0,1) and

$$\int_0^1 \varphi(u)\,du = 0 \,, \qquad \int_0^1 \varphi^2(u)\,du = 1 \,. \tag{211}$$

The following assumptions are needed for estimating the scale parameter τ

$$(S.2)\quad \varphi \text{ is differentiable }. \tag{212}$$

When estimating the intercept parameter based on signed-rank scores, the assumption that the score function is odd about ½ is needed.

$$(S.3)\quad \varphi_f(1-u) = -\varphi_f(u)\ . \tag{213}$$

Fisher information is defined as

$$I(f) = \int_0^1 \varphi_f^2(u)\,du\ , \tag{214}$$

where

$$\varphi_f(u) = -\frac{f'(F^{-1}(u))}{f(F^{-1}(u))}\ . \tag{215}$$

Note that the score function $\varphi_f(u)$ is optimal.

The major assumption on the error density function f for much of the rank-based analyses is

$$(E.1)\quad f \text{ is absolutely continuous, } 0 < I(f) < \infty\ . \tag{216}$$

This assumption also implies that f is uniformly bounded and is uniformly continuous.

Note that the scale parameter is defined as

$$\tau_\varphi = \frac{1}{\int \varphi(u)\,\varphi_f(u)\,du} . \tag{217}$$

Under the assumption (E.1) the scale parameter τ_φ is well defined.

Another assumption is needed for the intercept parameter analyzes.

$$(E.2) \quad f(\theta_e) > 0 , \tag{218}$$

where θ_e denotes the median of the errors distribution.

$$\theta_e = F^{-1}(\tfrac{1}{2}). \tag{219}$$

Recall that

$$\tau_s = \frac{1}{2f(\theta_e)} . \tag{220}$$

Under assumption (E.2) the scale parameter τ_s for the intercept parameter is well defined.

C.2 Theory of Rank-Based Estimates

Consider the linear model given by (207). Let $(\alpha_0, \beta_0)'$ be the vector of the true parameters. Then the true model is

$$Z = \mathbf{1}\alpha_0 + V\beta_0 + e . \tag{221}$$

Recall the gradient function

$$S(\boldsymbol{Z} - \boldsymbol{V}\beta) = \sum_{i=1}^{N} a(R(Z_i - \boldsymbol{v}_i'\beta))\boldsymbol{v}_i . \tag{222}$$

Theorem C.1 Under the model (207),

$$\begin{aligned} E[S(\boldsymbol{Z} - \boldsymbol{V}\beta_0)] &= \boldsymbol{0} \\ Var[S(\boldsymbol{Z} - \boldsymbol{V}\beta_0)] &= \sigma_a^2 \boldsymbol{V}'\boldsymbol{V} , \end{aligned} \tag{223}$$

where

$$\sigma_a^2 = \frac{\sum_{i=1}^{N} a^2(i)}{N - 1} \doteq 1 . \tag{224}$$

Theorem C.2 Let $\boldsymbol{\Sigma} = \mathbf{V'V}$ and under the assumptions of (E.1), (D.2), (D.3), (S.1), and model (207),

$$\frac{S(\boldsymbol{Z} - \boldsymbol{V}\beta_0)}{\sqrt{N}} \overset{D}{\rightarrow} N_p(\boldsymbol{0},\boldsymbol{\Sigma}) . \tag{225}$$

Jureckova (1971) derived an asymptotic linearity result for the process $S(\beta_n)$

$$\frac{1}{\sqrt{n}}S(\beta_n) = \frac{1}{\sqrt{n}}S(\beta_0) - \frac{\sqrt{n}(\beta_n - \beta_0)\Sigma}{\tau} + o_p(1) , \tag{226}$$

uniformly for $\sqrt{n}(\beta_n - \beta_0) = O(1)$ and the scale parameter τ is given by

$$\tau^{-1} = \int \varphi(u)\left(-\frac{f'(F^{-1}(u))}{f(F^{-1}(u))}\right) du \ . \tag{227}$$

Integrating the right hand side of (226) will give a locally smooth approximation of the dispersion function $D(\beta_n)$ which is given by the following quadratic function:

$$Q(\boldsymbol{Z}-\boldsymbol{V}\beta) = \frac{(\beta-\beta_0)'\boldsymbol{V}'\boldsymbol{V}(\beta-\beta_0)}{2\tau} - (\beta-\beta_0)'S(\boldsymbol{Z}-\boldsymbol{V}\beta_0) + D(\boldsymbol{Z}-\boldsymbol{V}\beta_0) \ . \tag{228}$$

Note that Q depends on τ and β_0 so it cannot be used to estimate β. However, the function Q can be used to establish asymptotic properties of the R-estimates and test statistics.

Theorem C.3 Under the model (207) and the assumptions (E.1), (D.1), (D.2), and (S.1), for any $\epsilon > 0$ and $c > 0$,

$$P\left[\max_{\|\beta-\beta_0\| < c/\sqrt{n}} |D(\boldsymbol{Z}-\boldsymbol{V}\beta) - Q(\boldsymbol{Z}-\boldsymbol{V}\beta)| \geq \epsilon\right] \to 0 \ , \tag{229}$$

as $n \to \infty$.

This theorem shows that Q provides a local approximation to D. Result of this theorem can be used to obtain the asymptotic distribution of the R-estimate. Without loss of generality assume the true $\beta_0 = \mathbf{0}$. Then Q, the quadratic function, can be written as

$$Q(\boldsymbol{Z}-\boldsymbol{V}\beta) = \frac{\beta' \boldsymbol{V}' \boldsymbol{V} \beta}{2\tau} - \beta' S(\boldsymbol{Z}) + D(\boldsymbol{Z}) \ . \qquad (230)$$

Since Q is a quadratic function, it follows from differentiation that its minimizing value can be obtained in terms of β. Set

$$\frac{\partial Q}{\partial \beta} = \frac{2}{2\tau}\beta(V'V) - S(\boldsymbol{Z}) = \boldsymbol{0} \qquad (231)$$

implies

$$\tilde{\beta} = \tau \frac{S(\boldsymbol{Z})}{V'V} \ . \qquad (232)$$

Hence, $\tilde{\beta}$ is a linear function of S(**Z**). The asymptotic distribution of $\tilde{\beta}$ can be obtained as the following theorem.

Theorem C.4 Under the model (207) and assumption (E.1), (D.1), (D.2), and (S.1),

$$\sqrt{n}(\tilde{\beta} - \beta_0) \overset{D}{\rightarrow} N_p(\boldsymbol{0}, \tau^2 \Sigma^{-1}) \qquad (233)$$

Since Q is a local approximation to D, it would seem that their minimized values are close also. Jaeckel (1972) showed the proof of the following theorem.

Theorem C.5 Under the model 136 And assumption (E.1), (D.1), (D.2), and (S.1),

$$\sqrt{n}(\hat{\beta} - \tilde{\beta}) \overset{P}{\rightarrow} \boldsymbol{0} \ . \qquad (234)$$

Combining the results of the previous two theorems provides the asymptotic distribution of the R-estimate.

$$\sqrt{n}(\hat{\beta} - \beta_0) \xrightarrow{D} N_p(\mathbf{0}, \tau^2\Sigma^{-1}) \ . \tag{235}$$

Refer to McKean Hettmansperger (1996) for the proofs of the theorems above.

C.3 Theory of Rank-Based Tests

Consider the general linear hypotheses

$$H_0\colon \boldsymbol{M\beta} = \mathbf{0} \quad \text{versus} \quad H_A\colon \boldsymbol{M\beta} \neq \mathbf{0} \ , \tag{236}$$

where **M** is a q × p matrix of full row rank. The following theorems and lemmas develop the asymptotic theory for this test statistic under null hypotheses. Let $\hat{\beta}$ denote the R-estimate of β in the full model of 136 and Q(β) denote the quadratic approximation of the dispersion function D(β).

Lemma C.1 Under assumptions (E.1), (S.1), (D.1), and (D.2),

$$D(\hat{\beta}) - Q(\hat{\beta}) \xrightarrow{P} 0 \ . \tag{237}$$

Lemma C.2 Let $\tilde{\beta}$ be the minimizing value of the quadratic function Q, then under the assumption (E.1), (S.1), (D.1), and (D.2),

$$Q(\tilde{\beta}) - Q(\hat{\beta}) \xrightarrow{P} 0 \ . \tag{238}$$

Let

$$\beta_r = \begin{bmatrix} \beta \\ \mathbf{0} \end{bmatrix}$$

be the reduced model vector of parameters, let $\hat{\beta}_{r,1}$ be the reduced model R-estimate of β_1, and let

$$\hat{\beta}_r = \begin{bmatrix} \hat{\beta}_{r,1} \\ \mathbf{0} \end{bmatrix} . \tag{240}$$

Let RD = D($\hat{\beta}_r$) - D($\hat{\beta}$) denote the drop in dispersion. McKean and Hettmansperger (1976) proved the following theorem.

Theorem C.6 Under the assumptions (E.1), (D.1), (D.2), (S.1), and H_0,

$$\frac{RD}{\tau/2} \overset{D}{\to} \chi^2(q) . \tag{241}$$

The subsequent test statistic that can be used is given by

$$F_R = \frac{RD/q}{\hat{\tau}/2} \tag{242}$$

where $\hat{\tau}$ is the scale parameter estimate. Although the test statistic qF_R has an asymptotic χ^2 distribution, McKean and Sheather (1991) reviewed numerous small sample studies which indicated that it is best to compare the test static with F-critical values with q and n-p-1 degrees of freedom.

Reject H_0: $\mathbf{M\beta} = \mathbf{0}$ in favor or H_A: $\mathbf{M\beta} \neq \mathbf{0}$ if $F_R \geq F(\alpha,q,n\text{-}p\text{-}1)$

where α is the level of significant. These small sample studies have showed that the empirical α level of F_R are close to the nominal values over a variety of designs, sample sizes and error distributions.

The rank analogue of Wald's test is given by

$$F_{R,Q} = \frac{(M\hat{\beta})'[M(V'V)^{-1}M']^{-1}(M\hat{\beta})/q}{\hat{\tau}^2} . \tag{243}$$

From the asymptotic distribution of $\hat{\beta}_R$, theorem (235), that under H_0, $qF_{R,Q}$ has an asymptotic χ^2 distribution. Hence the test statistics F_R and $F_{R,Q}$ have the same null asymptotic distribution. In fact, the difference of the test statistics converges to zero in probability under H_0. However, as Hettmansperger and McKean have shown, they are not algebraically equivalent.

C.4 Asymptotic Relative Efficiency

Let F_{LS} denote the least squares classical F-test which is defined by

$$F_{LS} = \frac{SS/q}{\hat{\sigma}^2} . \tag{244}$$

Under the assumption of the random errors e_i having finite variance σ^2, the null asymptotic distribution of qF_{LS} is a central χ^2_q distribution. Thus both F_R and F_{LS} have the same asymptotic distribution under the null hypothesis.

The sequence of alternative models to the hypothesis H_0: $\beta_2 = 0$ is:

$$Z = 1\alpha + V_1\beta_1 + V_2'(\theta/\sqrt{n}) + e, \tag{245}$$

where θ is a nonzero vector. Let

$$W = \frac{1}{A_2 - BA_1^{-1}B}$$

$$\text{where} \quad X'X = \begin{bmatrix} A_1 & B \\ B' & A_2 \end{bmatrix}. \tag{246}$$

Under the sequence of alternative models and the assumption (E.1), (D.1), (D.2), and (S.1), qF_R has an asymptotic noncentral χ^2-distribution with q degrees of freedom and non-centrality parameter

$$\eta_R = \tau^{-2}\theta W_0^{-1}\theta , \tag{247}$$

Where $\mathbf{W}_0 = \lim_{n\to\infty} n\mathbf{W}$. Under the sequence of alternative models, qF_{LS} has an asymptotic noncentral χ^2-distribution with q degrees of freedom and non-centrality parameter

$$\eta_{LS} = \sigma^{-2}\theta' W_0^{-1}\theta . \tag{248}$$

The asymptotic relative efficiency of F_R and F_{LS} is the ratio of their non-centrality parameters,

$$e(F_R, F_{LS}) = \frac{\eta_R}{\eta_{LS}} = \frac{\sigma^2}{\tau^2}. \tag{249}$$

If the distribution of the errors for a linear model is unknown, Wilcoxon score function is recommended to use for the R-estimate. If the distribution of the errors is normal and Wilcoxon scores are being used, the asymptotic relative efficiency of F_R and F_{LS} is 95%. While the loss in efficiency over the classical analysis at the normal distribution is 5%, the gain in efficiency over the classic analysis for long tailed error distribution can be substantial.

McKean and Sheather (1992) showed that the empirical power of F_{LS} at normal error distributions are slightly better than empirical power of F_R with Wilcoxon scores. However, the empirical power of F_R was much larger than the empirical power of F_{LS} if the error distribution had heavier tail. Note that the non-centrality parameters for the test statistics F_R and F_{LS} differ only in the scale parameters.

Refer to Hettmansperger and McKean (1996) for the proofs of the theories in this appendix.

BIBLIOGRAPHY

Albrecht, P. (1983). "Parametric Multiple Regression Risk Model: Some Connections with IBNR,". *Insurance: Mathematics and Economics*, 2, 69-73.

Burden, R. L., Faires, J. D., and Reynolds, A. C. (1981). *Numerical Analysis*, 2nd Edition, Prindle, Weber & Schmidt.

Cohen, A. C. and Whitten, B. J. (1981). "Estimation of lognormal distributions," *American Journal of Mathematical and Management Sciences*, Vol. 1, No. 2, 139-153.

De Vylder, F. (1976). "Geometrical Credibility," *Scandinavian Actuarial Journal*, 121-149.

De Vylder, F. (1978). "Estimation of IBNR claims by least squares," *Bulletin of the Association of Swiss Actuaries*, 78, 247-254.

De Vylder, F. (1982). "Estimation of IBNR claims by credibility theory," *Insurance: Mathematics and Economics 1*, 35-40.

Dixon, S. L. and McKean, J. W. (1996). "Rank-Based Analysis of the Heteroscedastic Linear Model," *Journal of the American Statistical Association*, 91, 699-712.

Doray, L. G.(1991). "Prediction of IBNR events with generalized regression, time series and compound Poisson models," Ph.D. thesis, University of Waterloo.

Foster, G. T., "The Determination of IBNR Reserves and the Effect of IBNR on the Underwriting of Excess of Loss," *Nederlandse Reassurantie Groep N. V.*, Amsterdam, p. 7.

Hadidi, N. (1985). "A Note on De Vylder's method of estimation of IBNR claims," *Insurance: Mathematics and Economics 4*. 263-266.

Hettmansperger, T. P.(1984). *Statistical inference based on ranks*, New York: John Wiley & Sons.

Hettmansperger, T. P. and McKean, J. W. (1996) *Nonparametrics and Robustness.*

Hettmansperger, T. P., McKean, J. W., and Sheather, S. J. (1996). "Rank-Based Analyses of Linear Models."

Jaeckel, L. A. (1972). "Estimating Regression Coefficients by Minimizing the Dispersion of the Residuals," *The Annals of Mathematical Statistics*, 43, 1449-1458.

Jorgensen, B. (1993). *The Theory of Linear Models*, Chapman & Hall.

Kapenga, J. and McKean, J. W. (1989). "Spline Estimation of the Optimal Score Function for Fitting Linear Models", *American Statistical Association*, Statistical Computing Section, 227-232.

Kapenga, J. A., McKean, J. W., and Vidmar, T. J. (1988). *RGLM: Users Manual*, Amer. Statist. Assoc. Short Course on Robust Statistical Procedures for the Analysis of Linear and Non-linear Models, New Orleans.

Koul, H. L., Sievers, G. L., and McKean, J. W. (1987). "An Estimator of the Scale Parameter for the Rank Analysis of Linear Models under General Score Functions," *Scandinavian Journal of Statistics*, 14, 131-141.

Kremer, E. (1982). "IBNR-Claims and the Two-way Model of ANOVA," *Scandinavian Actuarial Journal*, 47-55.

McKean, J. W. and Hettmansperger, T. P. (1978). "A Robust Analysis of the General Linear Model Based on One Step R-estimates," *Biometrika*, 65, 3, 571-579.

McKean, J. W., Naranjo, J. D., and Sheather, S. J. (1996). "Diagnostics to Detect Differences in Robust Fits of Linear Models."

McKean, J. W. and Schrader, R. M. (1981), "The Geometry of Robust Procedures in Linear Models," *Journal of the Royal Statistical Society B*, 42, 366-371.

McKean, J. W. and Sievers, G. L. (1989). "Rank Scores Suitable for Analyses of Linear Models Under Asymmetric Error Distributions," *Technometrics*, 31, 2, 207-218.

Rousseeuw, P. J. and Leroy, A. M. (1987). *Robust Regression and Outlier Detection*, John Wiley & Sons.

Sheather, S. J. and McKean, J. W. (1987). "A Comparison of Testing and Confidence Interval Methods for the Median", *Statistics & Probability Letters*, 6, 31-36.

Sievers, G. L. (1978). ."Weighted Rank Statistics for Simple Linear Regression", *Journal of the American Statistical Association*, Vol. 73, No. 363, 628-631.

Sievers, G. L. (1979). "A Weighted Dispersion Function For Estimation in Linear Models", *Technical Report No. 62.*

Staudte, R. G. and Sheather, S. J. (1990). *Robust Estimation and Testing*, John Wiley & Sons.

Wiser, R. F.(1990). "Loss Reserving", *Foundation of Casualty Actuary Science*, Casualty Actuarial Society, Chapter 4.

Witt, L. D., Naranjo, J. D., and McKean, J. W. (1995). "Influence Functions for Rank-Based Procedures in the Linear Model."