HEDGING THE RETURN ON EQUITY AND FIRM PROFIT: EVIDENCE FROM CANADIAN OIL AND GAS COMPANIES

by

Jiachi Zhu

Submitted in partial fulfillment of the requirements
for the degree of Master of Science

at

Dalhousie University
Halifax, Nova Scotia
August 2012

DALHOUSIE UNIVERSITY

DEPARTMENT OF MATHEMATICS AND STATISTICS

The undersigned hereby certify that they have read and recommend to the Faculty of Graduate Studies for acceptance a thesis entitled "HEDGING THE RETURN ON EQUITY AND FIRM PROFIT: EVIDENCE FROM CANADIAN OIL AND GAS COMPANIES" by Jiachi Zhu in partial fulfillment of the requirements for the degree of Master of Science.

Dated: August 22, 2012

Supervisor: ______________________________

Readers: ______________________________

____

DALHOUSIE UNIVERSITY

DATE: August 22, 2012

AUTHOR: Jiachi Zhu

TITLE: HEDGING THE RETURN ON EQUITY AND FIRM PROFIT: EVIDENCE FROM CANADIAN OIL AND GAS COMPANIES

DEPARTMENT OR SCHOOL: Department of Mathematics and Statistics

DEGREE: M.Sc. CONVOCATION: October YEAR: 2012

Signature of Author

TABLE OF CONTENTS

List of Tables

LIST OF FIGURES

ABSTRACT

In this thesis, we analyse the relationship between the hedging activities and return on equity, and the relationship between profit on hedging and other factors. We gather the annual reports of Canadian oil and gas companies and get hand-collected data which includes information of seventy five companies. Fully conditional specification is used to impute the missing values. Instrumental variable estimation and finite mixture of regression models are then used to predict the return on equity and hedging gain. We find the instrumental variable estimation is better than the OLS estimation to deal with the hedging data since it eliminates the endogeneity. By finite mixture of regression models, we show that different firms have different hedging strategies, which cause different profits in hedging. We also find the companies with large total assets prefer to hedge.

Key words: hedging, oil and gas, Canadian, fully conditional specification, instrumental variable model, finite mixture of regression models.

CHAPTER 1

INTRODUCTION

With the rapid development of financial markets, the risk management has been incredibly important. Many corporations use hedging activities to reduce the risk in their business operations.

The aim of this thesis is to examine role of hedging activities in Canadian oil and gas companies, to find out if there is a relationship between hedging activities and the return on equity and the firm profits for oil and gas companies. In this thesis, we are going to analyse an hand-collected data which gathered real companies' information and use different statistical models to verify the relationship between the hedging activities and profits.

1.1 Definitions of Hedging

In finance, a hedging means a specific investment activity used to reduce the risk of another investment. It includes a class of methods that enable the investors to profit when the return on investment is not as good as they expected. The original procedure trades two different commodities or underlying assets which have *similar market quotations, opposite direction, same quantity* and *break-even* . **Similar market quotation** means that the two products are equal in their market supply and demand relations, which means the prices of both products will increase or decrease simultaneously if the equilibrium of supply and demand curve changes. **Opposite direction** defines an operation that investors purchase one of two products when selling the other one at the same time, therefore they always get profit in one trade while get loss in the other, regardless the changes in price. The aim of this operation is to offset the exposure in price fluctuations so that to minimize the price risks. **Break-even** identifies the point where the revenue is just sufficient to cover the expanses, or we say that the break-even is the point at which the total revenue equals to the total cost. At the beginning, hedging activities were used to offset the risk and get close to the break-even point. However, with the development of financial markets, the investors no longer satisfied to use the hedging to merely reduce the risk, they desired more in hedging. The investors discovered that traditional hedging activities were too conservative in the new market conditions. Thus many financial vehicles of hedging have been created. They include, but not limited to, insurance policies, forward contracts, swaps and options. In particular, the option, one of the financial derivatives, is getting widely used. Within the method of hedging, the investors often use three different kinds of options. They are *call options, put options*, and *double options*. The definitions and examples of these different types of options will be given in the following sections.

In practice, there was a fund called the "hedge fund" that was created in the 1950's, it is used to gather the capitals from private investors and invest in the hedging investments. It's also capable of evading or neutralizing the risks in investments by using futures or options. But the question is why not all investors choose the hedging or hedge fund if they are so good.

In fact, there are many phenomena in the financial market that can influence the price dramatically, especially in energy prices. Because oil and gas can't be stored for a very long time, the reserves of oil and gas in the market is unpredictable, and some other conditions may also change their prices, such as instability of the government, weather, wars, etc.

Here is a real example. In September 2007, a hedge fund company named Amaranth Advisors sold over US \$3.0 billion gas put options to the gas companies. Believing that the price of gas could not increase rapidly in winter, many gas companies bought these put options to protect the price fall. Amaranth Advisors had faith that the price of gas would increase in winter based on the past experience, where the gas price always goes up dramatically. Unfortunately, on September 27^{th}, 2007, the price of gas decreased to \$7.67/BTU (British heat unit), although it was \$11/BTU three weeks ago. Most of gas companies started to exercise their put options so that Amaranth Advisors had to buy all the gas at a fixed price level that was much higher than the market price. Under this condition, Amaranth Advisors company has lost over US \$6.0 billion in the unsuccessful hedging activity.

From the example above, we understand that not all the hedging activities are as good as we expect, some may lead to a huge loss. Therefore we ask if the hedging activities are advantageous to oil and gas companies? Grant Oh, the global market chief deputy of Bank of Deutschland, believes that hedging is capable to lower the price volatility since the options in hedging usually represents the transactions for the commodity trades in future, so it has a long time span. In contrast with this view, Cameron Raether, the manager of Energy Department of Pabco Company, argues that some hedging activities should be responsible for intensifying the rupture of relationship between the price and fundamentals.

1.2 Call Option

A call option, often called a "CALL", is a financial contract between the buyer and the seller of it. The buyer of the call option acquires a right, but not the obligation, to buy or to exercise it for a certain quantity of a particular economic commodity or financial instrument at a certain price (the strike price) from the seller prior to and at a certain time. The European option can be exercised at the predetermined time, while the American option can be exercised at any time prior to or at the predetermined time. The seller has the obligation, not the right, to sell the economic commodity or financial instrument to the buyer if the buyer decides to exercise the options. The buyer pays a fee, called a premium, for obtaining this right, and the seller gets the premium for selling this rights.

There are several different types of call options, such as *stock call option, stock index call option, foreign currency call option,* and *commodity call option* . However, the key features of all these options are basically the same. That is, if the market price of the specific economic commodity or financial instrument has a higher probability to increase in future, then the buyer of the call option expects to profit from it. If the future market price is higher than the sum of the strike price and premium, the buyer can exercise the call option to buy this commodity or financial instrument and then sell it at the future market price to gain the benefit. When the market price is lower than the total value of the strike price plus premium in future, the buyer can abandon this option and lose the paid premium.

1.3 Put Option

The put option, which is often simply labelled as a "PUT", is a financial contract between the buyer and the seller of it. The buyer of the put option has the right, but not the obligation, to sell or exercise it for a certain quantity of a particular commodity or financial instrument to the seller of the option prior to and at a certain time (the expiration date or during the period of contract specified) for a certain price (the strike price). Same as the call option, the European option can be exercised at the predetermined time and American option can be exercised at any time prior to or at the predetermined time. If the buyer exercises the right, the seller has to buy the underlying instrument from the buyer at the strike price, regardless of the current market price. The buyer pays a fee, the premium, for obtaining this right, and the seller gets the premium for selling this right.

Compared with the call option, the purpose of purchasing the put option is to avoid the falling price. The investor would buy the put option for a particular underlying instrument if he/she thinks that the market price has a higher probability to fall in future. Therefore, the buyer has the right to sell the underlying instrument at a higher price (strike price). The buyer can discard this put option if the market price is higher than the strike price.

For instance, assume that an oil company, spends \$100 per barrel to produce oil, it is likely that the oil price may change in future, either to \$150, or to \$80. To avoid the possibility of 20% loss, the company chooses to purchase a put option for oil, which gives the company the right to sell oil at a fixed price, \$90 per barrel with a premium[1]. If the price falls below \$90, the company can exercise the put option and sell oil at \$90 per barrel to the put option seller, regardless the current oil price. But if the oil price is at or higher than \$90, this put option can be left expired.

It may generate the company 50% profit or 20% loss if selling oil in future without hedging the price changes. But, after employing the hedging activity (put option) in the oil price , the pay-off of this oil company is changed.

The potential profit to hedge the oil price by using a put option becomes

$$profit = (\$150 - premium)/\$100 \tag{1.1}$$

[1] The premium here refers to the cost of the put option paid by the options buyer.

the potential loss becomes

$$loss = (\$100 - \$90 + premium)/\$100 \tag{1.2}$$

Both profit and loss are diminished, compared to selling the oil directly.

Here is another example, in August 2001, China United Telecommunications corporation claimed in its financial report that the number of its clients had increased rapidly. This report gave investors a huge confidence. Therefore, many investors started to purchase the shares expecting a substantial gain in future. Among them, an investment company in Hong Kong spent HK \$0.2 billion to buy the shares. Unfortunately, on August 15^{th}, the stock price of China United Telecommunications dropped almost 10% because some multinational investment organizations sold their shares. By August 16^{th}, this Hong Kong investment company had lost HK \$20 million.

To prevent more losses from this stock, the company began to hedge the Hang Seng Index (the Hong Kong stock market). On June 22^{nd}, 2000, China United Telecommunications corporation had already been listed on the Hong Kong stock market, and has accounted for more than 12% in the Hang Seng Index, which means the investors can partially manipulate the entire index value by controlling China United Telecommunications corporation's stock. Hence, this investment company bought 1667 put options of the Hang Seng Index.

As they expected, the share price of China United Telecommunications corporation started to fall from HK \$12.30 to HK \$9.30 till August 20^{th}, which caused the investment company almost HK \$50 million loss. However, this company profited HK \$60 million from its 1667 put options when the Hang Seng Index decreased 720 points. Finally, it earned HK \$10 million.

1.4 Futures Contract

The futures contract, also called "double option", is a standardized contract between the buyer and seller which determines the future delivery of a certain quantity of an underlying asset in a future date at an agreed price. The buyer of the futures contract, has a "long" position, he/she has an obligation only to buy a certain quantity of an underlying asset at an agreed price (strike price). But the underlying asset will be delivered at a specific future time. the seller has a "short" position, he/she has the responsibility to sell a certain quantity of the underlying asset at the agreed price. This trade is carried out at a specified future time.

For example, to hedge the price changes, the oil and gas firm may use the futures contract to protect itself from the unfavourable move of the oil and gas price. If the oil and gas prices indeed fall as expected, the company can short the oil and gas by using the futures. This firm can profit from the future contract by *closing the position*[2]. If the oil and gas prices are expected to rise, the company can short the oil and gas futures contract to profit from the commodity market.

1.5 Literature Review

Since hedging is a very important procedure in the risk management, many empirical studies exhibit the role of hedging in the financial and commodity risk management. According to the Modigliani-Miller theorem, the imperfection in the financial market, asymmetric tax, costs and information, provides an incentive to hedge to increase the values to the firms. There are three main functions of hedging.

Firstly, Smith and Stulz (1985) found that, hedging could be used to increase the firm values to reduce the proportion of bankruptcy. In the same article, they pointed out that this effect is higher when the cost of financial distress is higher, which means hedging can decrease the financial distress. Froot et al. (1993) analysed the hedging activities when firms face the financial constraints and concluded that the underinvestment from outside financing would occur when the inside cash flow [3] is low enough, and hedging can avoid the underinvestment by generating extra cash. He also illustrated that the less the positive relationship between the investment opportunities and the cash flow at risk [4], the higher the hedging's efficiency. In conclusion, hedging can reduce the cost of financial distress and avert the underinvestment of external financing by adding firm value.

Secondly, Smith and Stulz (1985) suggested that for enterprises facing a convex tax liability function, they can reduce expected tax costs by hedging taxable incomes. The tax liability function is used to compute the total tax that the companies have to pay to the

[2] A close position represents an act of taking the opposite position of the current position so that to end a particular security or commodity investment, such as somebody sells the same quantity of a specified commodity when he bought it before, or buys the same quantity of this commodity when he sold it before.

[3] Cash flow is "the movement of cash into or out of a business, project, or financial product." See more details in wiki-pedia.

[4] Cash flow at risk is a risk measure to a firm's cash flow, computed by using the concept of value-at-risk.

authority at the end of the taxable event, and can be calculated by applying the appropriate tax rate to the taxable event's tax base. Graham and Smith (1999) employed the simulation method to investigate the convexity which is induced by tax-code provisions and they found that the tax liability functions are convex in more than 75% cases. Graham and Smith (1999) also showed that carryforwards and carrybacks improve the probability to hedge incomes. They confirmed that when firms face convex tax functions and use the hedging activities to incomes, they can save 5.4% of expected tax payments from every 5% reduction in taxable incomes by using their theory. Nevertheless, Graham and Rogers (2002) did not find any evidence to support that hedging is responding to the tax convexity by measuring the net long and short notional values. However, they indicated that hedging in taxable incomes has the ability to increase the capacity of debt and enhance interest tax deductions. From the literatures above, we can conclude that the hedging activities could be used to reduce the expected tax payments of taxable incomes.

Thirdly, hedging is used to moderate the risk exposure of corporation's management. Smith and Stulz (1985) analysed taxes, contracting costs and the impact of hedging policy on the firm's investment decisions, they discovered that the value-maximizing corporations will hedge for three reasons. One of them is called managerial risk aversion, where the hedging can be used to reduce the risk exposure. Barbara Pirchegger (2006) considered a two period LEN-type agency model with a risk averse agent and a risk neutral principal to indicate that hedge accounting [5] depends on how the overall risk exposure of the corporation is distributed over time. If the risk exposures are quite different over the two periods the principal prefers hedge accounting and when the risk exposures are not so different, no hedging accounting is preferred. Aziz A. Lookman (2011) examined a sample of firms in the oil and gas exploration and production (E&P) business and concluded that hedging a big risk is associated with lower firm value whereas hedging a small risk is associated with higher firm value. Viral Acharya et al. (2012) found that higher default risk tends to decrease the futures risk premium as a supply disruption benefits the long side of the futures contract.

Generally speaking, hedging is employed to maximize the shareholder value frequently, which increases the firm's internal and external cash flow over time. Nowadays, hedging is

[5] Most of the companies use derivative financial instruments to hedge their exposure to different risks, and the accounting for derivative financial instruments under International Accounting Standards is named hedge accounting.

mainly used in two types of risk management, financial and commodity.

For the financial risk management, Gagnon et al. (1998) indicated that the hedging in currency can obviously reduce the risk by applying the constructed currency portfolios. And Allayannis and Weston (2001) examined the effect of currency derivatives on relative market value[6] by using Tobin's Q ratio of the U.S. non-financial firm value and discovered a positive relationship between the currency hedging and firm value measured by Tobin's Q ratio. They indicated that hedging has the positive correlation with the relative firm value. Jin and Jorion (2006), however, argued that the hedging activities in Allayannis and Weston's literature is hard to control because of the risk exposure variation, endogenity of firm value and so on, by their analysis, therefore, the effect of hedging does not seem to affect market values [7]. On the other hand, Bartram et al. (2006) analysed the interest rate hedging in multi-industry companies and illustrated that effect of hedging has a positive relation to the firm value, which is consistent with Allayannis and Weston's (2001) conclusion.

For the commodity risk management, Rajgopal (1999) examined the risk disclosures of thirty eight U.S. oil and gas companies in the Securities and Exchange Commission (SEC) market, and concluded that the reserve of oil and gas will influence positively the relationship between oil and gas price and stock returns. Jin and Jorion (2006) found that, unlike the oil and gas reserve can improve the relationship, hedging has the ability to reduce the relationship between the stock returns and oil and gas prices. Meanwhile, as mentioned before, they also showed that hedging may not affect the firm value from analysis of 119 U.S. oil and gas companies, this is against the result from Allayannis and Weston (2001). On the other hand, Carter et al. (2006) looked the data from the jet fuel hedging behaviour of firms in the US airline industry during 1992-2003, they pointed out that hedging in the jet fuel has the positive impact on firm values of the airline companies. However, different from the oil and gas companies, the airline firms are consumers of oil, so few investors would use their firms' stock to conclude on their oil prices. This may be the reason of the contrast with the conclusion of Jin and Jorion (2006). In addition, Viral Acharya et al. (2012) solved the a general equilibrium model which the managerial costs of default are the motivation for firm hedging and showed that the implications of this

[6]Market value is a particular asset price which can be traded in a competitive market setting.

[7]The most important issue is they studied the hedging activities of 119 U.S. oil and gas producers, which can't be categorized to the financial part.

model are qualitatively the same as in the model with risk.

The following tree Table 1.1 summarizes the main contributions of these literatures, and more details about the literature will be shown in the Appendix.

Figure 1.1: Treeplot of Hedging

Hedging Function

- Reduce distress cost
  - Smith and Stulz (1985) *(state-preference model)*
  - Froot et al. (1993) *(variation of CSV)*
- Reduce the tax payments
  - Smith and Stulz (1985) *(state-preference model)*
  - Graham and Smith (1999) *(simulation methods)*
  - Graham and Rogers (2002) *(tax function convexity)*
- Moderate the risk exposure
  - Smith and Stulz (1985) *(state-preference model)*
  - Barbara Pirchegger (2006) *(two period LEN-type model)*
  - Aziz A. Lookman (2011) *(multiple Tests)*
  - Viral Acharya et al. (2012) *(general equilibrium model)*

1.6 Structure of the Thesis

The rest of the thesis is organized into five chapters. Chapter two illustrates the data collection and some accounting concepts. Chapter three solves the missing values problems by applying the multiple imputations, it also examines the basic statistics and compare the data before and after the multiple imputations. Furthermore, it separates the data into two parts, the hedging-related variables part, which contains all the variables related to the hedging activities, and non-hedging-related part, which includes the variables that exist regardless of whether the companies apply the hedging activities or not. Chapter four analyses the post-imputation data by employing *instrumental variable model*[8]. Chapter five applies *finite mixture of regression models*[9] to test the data and verify our conclusions. At last, chapter six summarizes the main conclusions of this thesis.

[8]The IV model is used to predict the Y variables by applying the data in last year.
[9]We use the finite mixture models to predict the Y variables by the other variables of this firm.

CHAPTER 2

DATA

2.1 Data Description

2.1.1 Oil and Gas Environment in Canada

There are several important issues that must be paid attention to when we collect the hedging data of Canadian oil and gas companies. Firstly, because of the abundant oil and gas resources and stable economic markets in Canada, more than 150 oil and gas companies are engaging exploration, production, refinery and marketing. Many of them are small companies with the total assets under CAD$500 million. According to Haushalter (2000), the large oil and gas companies are more likely to hedge in America. Many small oil and gas companies are oil and gas exploration companies, they may not hedge as much, therefore we focus on oil and gas companies which have CAD$500 million or more. Secondly, some oil and gas companies in Canada are the subsidiaries of international oil and gas corporations. Hence, their hedging activities may not only be in the Canadian market, but also be in the global markets, this may cause some outliers in the hedging activities. Thirdly, after the recent U.S. financial crisis and global economic recession from 2008, the Canadian Generally Accepted Accounting Principles (GAAP) are implemented in the companies' financial reporting, the annual reports in Canada since then are therefore different from their previous reports. To collect data according to the GAAP, the sample is limited to the 2009-2010 period. Fourthly, although the oil and gas sector in Canada is a major sector in Canada, Canadian economy accounts for only a tiny part of the global economy.

2.1.2 Data Collection & Formation

Our analysis is based on the sample of 79 oil and gas companies over the period from 2009 to 2010. We extract a list of oil and gas firms in Canada that have the total assets of CAD $500 million or more in 2010 from " Bloomberg ". This gives us the names of collecting about 100 companies. Then we collect the data from their annual reports in the System for Electronic Document Analysis and Retrieval (SEDAR) website [1]. Only forty-two companies in the list have the annual reports in SEDAR during the period from 2009 to 2010. Therefore we have to hand collect the annual report data for the other fifty-eight firms from their own websites.

We only retain the companies which have met the following criteria: the companies that report the complete information in Risk Management of Management's Discussion and Analysis and Financial Instruments in Notes of Consolidated Financial Statement; the companies that report the informations about the hedging activities (include hedging or not hedging); and the companies that have the annual reports that have the same pattern which is abode by the Canadian Generally Accepted Accounting Principles (GAAP) during the period of 2009-2010. After screening the list of all companies, our final data consists 79 companies from 2009 to 2010, or 158 firm-year observations, in which 76 firms have presented both of two years data [2].

In these 79 companies, we extract the following variables from their annual reports for our analysis: *hedging style; realized hedging gain in gas, realized hedging gain in oil and realized hedging gain in both oil and gas; derivative in assets, derivative in liability and derivative in revenue; total assets, total liability and total revenue, percentage of derivative in assets, percentage of derivative liability and percentage of derivative revenue; current ratio; return on equity; difference between realized price and the benchmark price, and indicator for hedging activities or not*. When comparing the oil and gas realized prices with the benchmark prices in markets, we use the average monthly Alberta gas price (AECO) as the benchmark gas price, and the average West Texas Intermediate (WTI) oil price as the benchmark oil price. Table 2.1 shows the variable names and their abbreviations in our analysis.

[1]The annual reports from SEDAR are available in www.sedar.com. It's an official website which provides most of the public financial documents and reports of the firms located in Canada under the Canadian Securities Administrators (CSA).

[2]One of these 79 companies misses the annual report in 2010, while the other two firms miss both of 2009 and 2010 annual reports.

Table 2.1: Table of Variables and Their Abbreviations

	Abbreviation	Complete Name
Hedging Related Variables	GG	Gas Hedging Gain
	OG	Oil Hedging Gain
	TG	Total Hedging Gain
	DA	Derivative in Assets
	PA	Derivative in Assets in percentage
	DL	Derivative in Liabilities
	PL	Derivative in Liabilities in percentage
	DR	Derivative in Revenues
	PR	Derivative in Revenues in percentage
Non-Hedging Related Variables	GDP	Difference between Firm's Price and Market Price in Gas
	ODP	Difference between Firm's Price and Market Price in Oil
	TA	Total Assets
	TL	Total Liabilities
	TR	Total Revenues
	CR	Current Ratio
	ROE	Return on Equity
	INDI	Indicators for Hedging or Not

Notes: This table shows the complete names of variables and their abbreviations. These abbreviations will be used throughout the thesis.

To analyse the relationships between hedging activities & return on equity, and the impact of hedging on the firm profits, we assume that the variables return on equity (ROE), realized hedging gain in gas (GG), realized hedging gain in oil (OG), and realized hedging gain in gas and oil (TG) be the predicted variables while the others be the predictor variables[3].

The hedging information can be retrieved in the annual reports in two ways,

A. three financial statements.

(1) Consolidated Balance Sheets.

(2) Consolidated Statement of Comprehensive Loss/Income.

(3) Consolidated Statement of Cash Flow.

B. Analysis and Financial Instruments in Notes of Consolidated Financial Statement.

We can discover the numerical information from these three statements, and the Analysis and Financial Instruments in Notes of Consolidated Financial Statement displays the details of the hedging derivatives such as realized or unrealized put, call option and fixed price contract, etc. The most hedging styles selected by the oil and gas companies in Canada is fixed price contract, swaps, put option, call option and futures contract, as noted by Dan et al. (2005)[4].

2.2 Interpretations of Accounting Concepts

As detailed above, we have collected the data of the eighteen variables from the companies' annual reports during 2009 to 2010. Now we review some important definitions of accounting based on the book entitled "*Corporate Finance, Portfolio Management, and Equity Investments*" (Bruce Kuhlman, 2005).

First of all, the annual report is one typical style of the company's financial reports, it shows their financial performance to consumers, investors, banks and any other parties by presenting the financial statements, based on which the investors and managers could make the economic decisions. There are three key financial statements. They are *Balance Sheet, Income Statement* and *Statement of Cash Flow*.

[3] Since there are strong linear relationship between GG, OG and TG, which can be written as $TG = OG + GG$, we may just need to analyse TG instead of OG and GG.

[4] Dan (2005) also notes that the main hedging instruments used by Canadian oil and gas companies are fixed-price contracts, forwards, received-fixed swaps and options.

The *Balance Sheet* reports the firm's financial position at one point of the time. It lists Assets, Liabilities and shareholder's equity. Assets are probable current and future benefits controlled by a particular entity in the past transactions, or we can briefly say, assets are company's resources. Liabilities are similar with the assets, which imply the probable future costs, they occur from the obligations of a particular entity to transfer asset or provide services to other entities in the future as a result of past transactions. Shareholder's equity means the residual values of net assets for a company after deducting its liabilities. Hence, the relationship among them can be shown as

$$Assets = Liabilities + Shareholder's\ Equity \tag{2.1}$$

We use *derivative in assets, derivative in liabilities, total assets* and *total liabilities* from this part.

The *Income Statement* presents the financial performance of the company over a period of time, from which we obtain the variables *derivative in revenue* and *total revenue*. The income statement also lists: Revenues, Expenses and Gains & Losses. Revenues are inflows of major operation activities such as producing commodity, selling goods, rendering services, etc. On the contrary, expenses mean the outflows from their major operation activities. Gains & losses stand for the increases and decreases in the equity or net assets from these transactions.

The *Statement of Cash Flow* tells us the company's cash receipts and payments. The cash flow statement records the operating cash flow, investing cash flow, and financing cash flow. Operating cash flow includes the cash effects of transactions for firm's normal business. And investing cash flows are from the acquisition or sale of properties, plants and equipments; of securities; of a subsidiary; and of investments in the other companies. Just as its name expresses, financing cash flow involves issuance or retirement of the firm's debt and equity securities and includes dividends paid to stockholders.

The other variables are either calculated by ourselves or selected from the Analysis and Financial Instruments in Notes of Consolidated Financial Statement. Since we have extracted derivatives in assets, liabilities and revenues from *Balance Sheet* and *Income Statement*, we can calculate the percentages of derivatives in some form of assets/liabilities/revenues. CR (current ratio) is computed by using current assets divided by current liabilities which can be found in *Balance Sheet*. ROE equals to net income divided by shareholder equity, where the denominator is found in *Balance Sheet* and the numerator is from *Income Statement*. Realized hedging gain in gas, oil and total, difference between realized price and the benchmark price are calculated in the Analysis and Financial Instruments in Notes of Consolidated Financial Statement. Indicators are used to distinguish whether the companies applied the hedging activities during 2009 - 2010 or not. The hedging style variable represents the different methods in hedging activities which the companies applied.

Chapter 3

Multiple Imputation of Missing Data

Our data contain some missing values since the annual reports are not entirely comparable due to various reasons, such as different accounting practices and different regulations that these firms are subject to. There are 43 companies with one or more missing values distributed in variables GDP, ODP, TA, TL, TR, CR, ROE, GG, OG and TG. In this chapter we will use the multiple imputation to deal with this problem.

Multiple imputation (Rubin 1987) is a statistically valid method of dealing with the missing value problems in the complicated incomplete data. There are two common approaches in multiple imputation: (1) joint modeling (Schafer 1997) and (2) fully conditional specification (FCS) (Van Buuren et al. 2007).

The joint modeling approach (Schafer 1997), has been applied in the log-linear, multivariate normal, and general location models since it imputes missing values from their conditional distributions by the MCMC (Markov Chain Monte Carlo) method. It specifies a parametric multivariate density, $P(Y|\theta)$, where θ and Y represent respectively the model parameter and the data. Using this specification, we can generate imputations from the posterior predictive distribution $P(Y^{miss}|Y^{obs})$. This approach works well when the multivariate distribution, $P(Y|\theta)$, is a reasonable description of the data.

The fully conditional specification (FCS) approach (van Buuren et al. 2005) specifies a multivariate imputation model on a variable-by-variable basis by a number of separately specified conditional densities for the variable that contains missing values. FCS implicitly specifies $P(Y|\theta)$ by using a separate conditional density $P(Y_j|Y_{-j}, \theta_j)$ for each Y_j, where θ and Y represent the model parameter and the data, respectively, Y_j is the j^{th} variable in data Y and Y_{-j} means the all the variables in data Y exclude Y_j. FCS iterates over all of the separately specified conditional densities to impute missing data. This approach is attractive when a joint distribution is difficult to specify. In this thesis, because we face difficulties to specify the multivariate distributions, we will employ the fully conditional specification approach.

Comparing to the joint modeling, Van Buuren et al.(2005, 2007) indicated FCS has several important advantages over the joint modeling approach. Firstly, FCS allows to create more flexible models because it deals with k one-dimensional problems instead of a k-dimensional problem. Secondly, FCS can help to preserve investments in specialized imputation methods that are difficult to formulate as a part of a multivariate density $P(Y|\theta)$. It is easy to incorporate imputation methods that preserve unique features in the data, such as bounds, interactions, bracketed responses, and so on. Thirdly, it is easier to generalize to models under non-ignorable missing data mechanisms. In addition, FCS is also easier to communicate to users by specifying a separate imputation model for each variable.

3.1 Missing Values in the Dataset

The hedging data of oil and gas companies, retrieved from the annual reports of selected companies for the period from 2009 to 2010 has 158 observations, which should contain 2686 values but only with 2054 available. There are 632 missing values distributed in thirteen out of seventeen variables. The variables which contain the most missing values are realized gas hedging gain (103 missing values), realized oil hedging gain (97 missing values), difference between firm gas price and market gas price (89 missing values) and difference between firm oil price and market oil price (82 missing values). The others only have a few missing values, which appear in the realized gain of both oil and gas hedging (73 missing values), derivative in assets (57 missing values), derivative in liabilities (55 missing values), derivative in revenues (55 missing values), and total assets (4 missing values), total liabilities (4 missing values), total revenues (5 missing values), current ratios (4 missing values) and return on equity (4 missing values).

The missing values are resulted from the different regulations, of different companies. It is reasonable to treat these missing values as missing completely at random (MCA). Therefore, the fully conditional specification is employed in our analysis.

3.2 Fully Conditional Specification Approach

In FCS, let Y_j $(j = 1, 2, 3, ..., k)$ be one of the k variables that have missing values, Y_j^{obs} represents the observed values in Y_j and Y_j^{miss} stands for the missing values in Y_j. Therefore $Y^{obs} = (Y_1^{obs}, Y_2^{obs}, Y_3^{obs}, ..., Y_k^{obs})$ and $Y^{miss} = (Y_1^{miss}, Y_2^{miss}, Y_3^{miss}, ..., Y_k^{miss})$ are used to denote, respectively, the observed and missing values in the data Y. We use Y_{-j} to denote a set of the variables in Y excluding Y_j, $Y_{-j} = (Y_1, Y_2, ..., Y_{j-1}, Y_{j+1},$
$Y_{j+2}, ..., Y_k)$.

FCS assumes that Y is a hypothetically complete data with k-variate with the joint distribution $P(Y|\theta)$. The challenge is how to get the distribution of θ, regardless it is explicit or implicit. To get the posterior distribution of θ, FCS uses the conditional distributions given below:

$$\begin{aligned} &P(Y_1|Y_{-1}, \theta_1),\\ &P(Y_2|Y_{-2}, \theta_2),\\ &...\\ &P(Y_k|Y_{-k}, \theta_k). \end{aligned} \tag{3.1}$$

Starting from a simple draw from the observed marginal distributions, the tth iteration of FCS is a Gibbs sampler consists of the following successive draws

$$\begin{aligned} &\theta_1^{*(t)} \sim P(\theta_1|Y_1^{obs}, Y_2^{(t-1)}, ..., Y_k^{(t-1)}),\\ &Y_1^{*(t)} \sim P(Y_1^{miss}|Y_1^{obs}, Y_2^{(t-1)}, ..., Y_k^{(t-1)}, \theta_1^{*(t)}),\\ &...\\ &\theta_k^{*(t)} \sim P(\theta_k|Y_k^{obs}, Y_1^{(t)}, ..., Y_{k-1}^{(t)}),\\ &Y_k^{*(t)} \sim P(Y_k^{miss}|Y_k^{obs}, Y_1^{(t)}, ..., Y_{k-1}^{(t)}, \theta_k^{*(t)}). \end{aligned} \tag{3.2}$$

where $Y_j^{(t)} = (Y_j^{obs}, Y_j^{*(t)})$ represents the j^{th} imputed and actual variable at iteration t. Unlike the MCMC techniques in joint modeling, convergence can be quite fast since the previous imputations $Y_j^{*(t-1)}$ only enter $Y_j^{*(t)}$ through its relation with the other variables. And there is no information about Y_j^{miss} is used to draw $\theta_j^{*(t)}$, which is also different from joint modeling of the MCMC approach. Van Buuren et al. (2005, 2007) also mentioned that the number of iterations is fixed to a small number, about 5 or 10 times.

3.3 Multiple Imputation based on FCS

In this section, we employ the fully conditional specification to impute the missing data. According to Van Buuren et al. (2005), we assume the distribution of $Y = (Y_1, Y_2, ...Y_{17})$ can be specified by a parameter called θ. After that, we calculate the posterior distribution of θ based on the observed data, $P(\theta|Y^{obs})$, and predict the missing values of Y for each variable iteratively by applying the following formulas:

$$\begin{aligned}
&\theta_1^{*(1)} \sim P(\theta_1|Y_1^{obs}, Y_2^{(0)}, ..., Y_{17}^{(0)}), \\
&Y_1^{*(1)} \sim P(Y_1^{miss}|Y_1^{obs}, Y_2^{(0)}, ..., Y_{17}^{(0)}, \theta_1^{*(1)}), \\
&\theta_2^{*(1)} \sim P(\theta_2|Y_2^{obs}, Y_1^{(1)}, Y_3^{(0)}, ..., Y_{17}^{(0)}), \\
&Y_2^{*(1)} \sim P(Y_2^{miss}|Y_2^{obs}, Y_1^{(1)}, Y_3^{(0)}, ..., Y_{17}^{(0)}, \theta_2^{*(1)}), \\
&... \\
&\theta_{17}^{*(1)} \sim P(\theta_{17}|Y_{17}^{obs}, Y_1^{(1)}, ..., Y_{16}^{(1)}), \\
&Y_{17}^{*(1)} \sim P(Y_{17}^{miss}|Y_{17}^{obs}, Y_1^{(1)}, ..., Y_{16}^{(1)}, \theta_{17}^{*(1)}).
\end{aligned} \tag{3.3}$$

where imputation models of $Y^* = (Y_1^*, Y_2^*, ..., Y_{17}^*)$ can be shown as $Y_j^*|Y_{-j} \sim N(\mu_{-j}^* + \theta_j^* Y_{-j}, \sigma_{-j}^{2*})$, where μ_{-j}^*, θ_j^* and σ_{-j}^{2*} were draws from the appropriate posterior distributions. And we also estimate the θ by $\Theta_i = ((Y_{-i}^{obs})^t Y_{-i}^{obs})^{-1} (Y_{-i}^{obs})^t Y_i^{obs}$. Following Van Buuren(2005, 2007), we iterate the above loop for five times, to get that all the missing values are converged on five times of loops.

After multiple imputing all the missing values in the data, we need to compare the post-imputation data with the pre-imputation data and find out the differences between them. Table 3.1 shows the basic statistics for pre- and post-imputation data.

Table 3.1: Basic Statistics for Pre- and Post-imputation Data of 158 Observations

	Pre-Imputation			Post-Imputation		Difference	
Variable	Missing N	Mean	SE Mean	Mean	SE Mean	Difference in Mean	Difference in SE
GG	103	113.50	54.00	64.40	18.30	49.10	35.70
OG	97	36.30	25.90	18.88	9.44	17.42	16.46
TG	73	99.30	35.90	82.00	19.20	17.30	16.70
DA	57	59.20	18.30	59.90	12.50	-0.70	5.80
PA	0	0.29	0.05	0.29	0.05	0.00	0.01
DL	55	73.90	22.90	53.30	14.70	20.60	8.20
PL	0	1.36	0.26	1.36	0.26	0.00	0.00
DR	55	96.00	35.30	95.60	24.40	0.40	10.90
PR	0	1.77	0.47	1.77	0.47	0.00	0.00
GDP	89	0.28	0.09	0.20	0.05	0.08	0.04
ODP	82	-9.41	0.72	-10.68	0.51	1.27	0.21
TA	4	5425.00	829.00	5320.00	808.00	105.00	21.00
TL	4	2693.00	462.00	2641.00	450.00	52.00	12.00
TR	5	1930.00	340.00	1894.00	329.00	36.00	11.00
CR	4	1.94	0.21	1.96	0.21	-0.02	0.00
ROE	4	0.05	0.01	0.05	0.01	0.00	0.00
INDI	0	0.59	0.04	0.59	0.04	0.00	0.00

Notes: The meaning of variables' abbreviations can be found in Table 2.1. The data contains 158 observations, which includes 4 outliers and 4 miss all the informations.

Comparing the pre- and post-imputation data, we find that except the variables DA and CR, the means of all other variables becomes smaller after multiple imputation. And the standard error (SE) of all the variables are smaller or equivalent after we impute the data. This shows the imputation smooths the entire data and shrinks the diversities among oil and gas companies in Canada when many of them may choose the similar strategies during 2009 to 2010.

3.4 Basic Data Exploration

The dataset contains 158 observations in total (79 companies) after imputing the missing values in each variable. We will delete four outliers (Athabasca and Provident in 2009 and 2010) and four observations that contain large number of missing values[1] (Suncor and Niko in 2009 and 2010). Therefore, there are 150 observations (75 companies). And seventeen variables which include twelve predictor variables, four predicted variables and one indicator for hedging. We classify all variables except the indicator variable into two groups, one is called hedging-related variables, which include GG, OG, TG, DA, PA, DL, PL, DR and PR; the other group includes non-hedging-related variables, namely GDP, ODP, TA, TL, TR, CR and ROE. We log-transform the variables TA, TL, TR and CR since they contain huge differences in values between large companies and small companies, we use the log-transformation to decrease this difference to make them more normally distributed.

[1]These two companies contain missing values in TG, DA, PA, DL, PL, DR, PR, GDP and ODP, the imputation will not be reliable.

Table 3.2: Basic Statistics for Seventy-five Companies

	Eighty-eight Hedging Observations				Sixty-two Non-hedging Observations			
Variable	Mean	SE Mean	Minimum	Maximum	Mean	SE Mean	Minimum	Maximum
GG	68.70	29.30	-35.00	2250.00	58.40	15.40	-35.00	808.00
OG	25.50	15.80	-84.00	1330.00	9.44	4.39	-84.00	81.00
TG	94.00	31.50	-77.00	2250.00	64.90	12.10	-77.00	344.00
DA	62.50	19.30	-151.00	1272.20	56.30	12.90	-126.00	361.00
PA	0.49	0.09	-1.09	3.63	0.00	0.00	0.00	0.14
DL	79.80	24.60	0.00	1991.00	15.74	3.18	0.00	106.00
PL	2.26	0.41	0.00	23.52	0.08	0.07	0.00	4.53
DR	103.80	38.00	-210.00	2214.40	83.90	24.20	-165.50	632.00
PR	3.15	0.77	-16.61	25.41	-0.19	0.15	-6.57	3.15
GDP	0.39	0.07	-0.50	2.89	-0.08	0.08	-1.89	1.58
ODP	-11.03	0.68	-23.19	9.65	-10.19	0.79	-22.39	9.65
TA	7801.00	1280.00	108.00	46589.00	1799.00	441.00	131	20580.00
TL	3929.00	725.00	13.00	28705.00	812.00	205.00	6.00	9403.00
TR	2371.00	418.00	9.00	15127.00	1216.00	524	0.00	25092.00
CR	1.09	0.10	0.01	4.51	3.19	0.43	0.41	18.57
ROE	0.04	0.01	-0.45	0.25	0.05	0.02	-0.30	0.58
INDI	1.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00

There are 44 companies, which have employed the hedging activities during 2009 and 2010, the other 31 companies (62 observations) didn't employ hedging activity. Theoretically, the values of hedging-related group variables should be 0 for the non-hedging companies because they did not employ hedging activities. However, the means of hedging-related variables in non-hedging companies are not zero. This is because some companies still could get profits or needed to pay their liabilities due to the hedging activities in previous years. The managements of their companies must report this in the annual reports even if they do not employ any hedging activities in 2009 and 2010.

Table 3.2 compares the hedging companies [2] with non-hedging companies, it shows that the hedging activities make differences in some variables (such as GDP, ODP, TA, TL, TR and CR) but not in others (such as ROE).

By using the independent two sample t-test of GDP, we find that the sample mean of GDP in hedging companies is significantly greater than that of GDP in non-hedging companies. However, the mean of ODP in hedging companies is smaller than that of non-hedging companies. It makes sense that proper hedging activities reduce the market risks and make the profits, while, improper hedging strategies cut down the profits, and cause losses. During the period from 2009 to 2010, most oil and gas companies in Canada made the right decisions in hedging in gas. But oil spill in the Gulf of Mexico and economic recession in Europe in 2010, seriously influenced the oil price and the hedging activities were not good, reflected in the ODP variable. As Table 3.2 shows, relative to the non-hedging companies, the standard error of GDP and ODP in hedging companies is smaller than those of non-hedging companies' standard error. That means the hedging activities also affected variances. Because of the properties of hedging, it must reduce the volatility in selling prices when hedging reduces the risks.

By using the independent two sample t-tests, we find the means of TA, TL and TR in hedging companies are significantly larger than those in non-hedging companies. Values of TA, TL and TR typically indicate the sizes of the company. Most large companies hedge risks against the market price volatility. Most of transnational enterprises produce dozens or, even hundreds of times more than small oil and gas companies. Thus, they are intolerant to unexpected price fluctuations. That's why these transnational companies would rather use hedging activities to avoid the loss, even the hedging reduces the profits. On the

[2] The companies which apply the hedging activities in 2009 and 2010

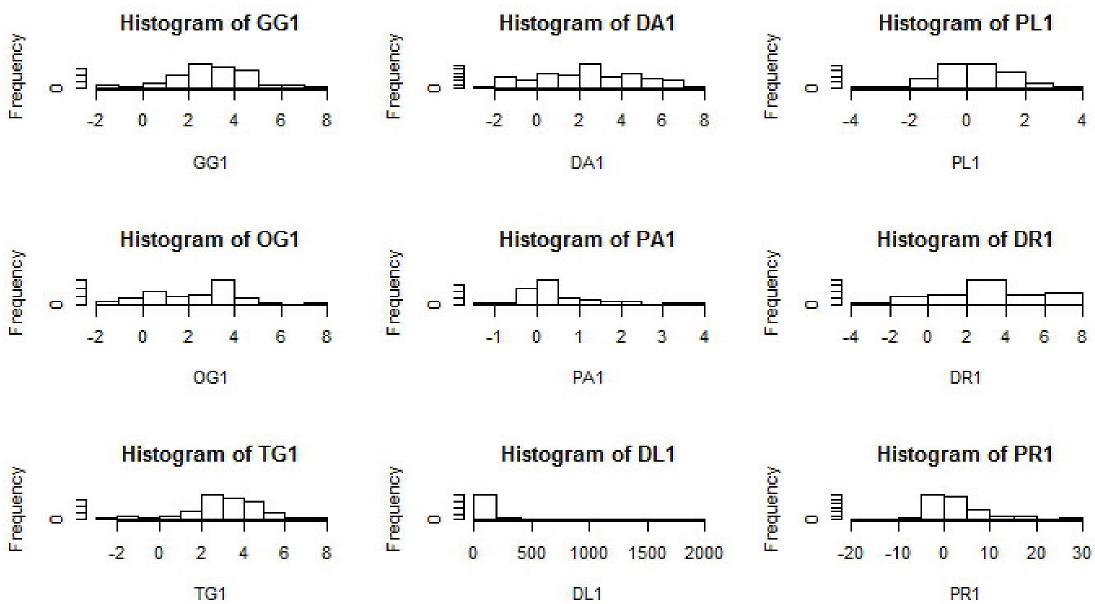
other hand, because of the insufficient capitals, small companies do not have enough fund to sustain their expenses in hedging, they also can't afford for high costs in production, transition and storage. Hence, they have to cut any unnecessary cost. This is the reason that many small Canadian oil and gas companies do not hedge.

By applying the two-sample t-test, we find that the sample mean of CR in hedging companies is significantly smaller than that of non-hedging companies. The variable CR represents the ratio of the current assets divided by the current liabilities. Generally, the current ratio (CR) is in the range of 0.5-2.0, a higher ratio may imply that the companies have more idle current assets or have small current liabilities. A lower ratio may tell us that the frequency of cash flow in the companies is slow. The hedging derivatives do not need a large cash flow, and thus it decreases the current ratio (CR).

Finally, there is no significant difference in ROE between hedging companies and non-hedging firms. This may imply that the hedging activities do not affect the variable ROE, which means hedging activities may not be able to help investors to earn a higher return on equity. Nevertheless, the hedging activities indeed can enhance the enterprises' operating capacity significantly by, for example, to increase the companies' funds utilization efficiency or improve cash flow. Hence, there must exist some relationship between ROE and hedging activities.

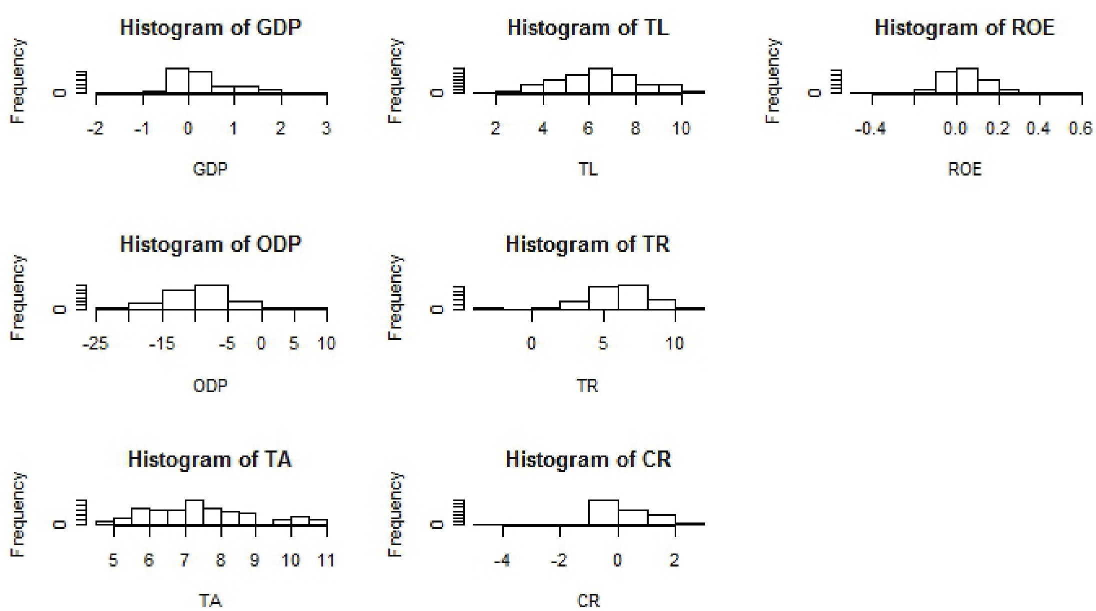
Figure 3.1 and Figure 3.2 show the histograms of all hedging-related variables of hedging companies and non-hedging-related variables of all the companies, respectively.

Figure 3.1: Histograms of Hedging-related Variables



Notes: The meanings of variables' abbreviations are shown in Table 2.1. In Figure 3.1, we only retain the companies which employ the hedging activities in 2009 and 2010 since there exist many 0's corresponding to the non-hedging companies.

Figure 3.2: Histograms of Non-hedging-related Variables



Notes: The meanings of variables' abbreviations are shown in Table 2.1. The distributions of non-hedging-related variables are also quite normal.

From these two Figures, we find that the distributions in hedging-related variables are quite normal, except the variable DL. This is because many of Canadian oil and gas companies are small companies. Their liabilities are distributed in many aspects, such as accounts payable and future income tax liability, in which hedging activities just occupy a small percentage, when the other large companies spend much more in hedging. Same as the hedging-related variables, the non-hedging-related variables are also quite normally distributed.

CHAPTER 4

INSTRUMENTAL VARIABLE ESTIMATION

In this Chapter, we focus on the instrumental variable (IV) estimation. We apply the IV estimation to predict the Y variables by using other variables in last year. We will model the relationship between ROE in 2010 and other variables in 2009 and between TG in 2010 and the other predictors in 2009.

IV estimation is used to deal with the causal relationships when controlled experiments are not feasible, it can circumvent the endogeneity problems in economic data. It is efficient when the predictor variables are correlated with the error terms of a regression relationship. In our data, there exist situations that one or more X variables are endogenous, such as DA and TA, DL and TL. They either influence or be influenced by each other.

Section 4.1 describes the IV estimation method, Section 4.2 applies IV estimation for our data. Section 4.3 applies the Hausman test to compare the IV model with the original model, and concludes the IV model fits data better. We then employs the IV estimation to predict TG in Section 4.4.

4.1 Review of IV Estimation

Instrumental variable estimation has been widely used in the literature. Philip G. Wright (1928) was the first person to derive the theory of IV estimation. Nelson and Startz (1990) discussed the exact small sample distribution of the IV estimator and compare it with the asymptotic distribution. Maddala and Jeong (1992) pointed out that the bimodality is a consequence of the special case considered by Nelson and Startz's report. Imbens and Angrist (1994) demonstrated that the linear IV estimation can be interpreted under weak conditions as a weighted average of local treatment effects, when the weights depend on the elasticity of the endogenous X variables to changes in the instrumental variables. Pearl (2000) used counterfactuals and graphical criteria to formally define the instrumental variables. Angrist and Krueger (2001) applied the techniques of IV to a survey of the history. Stock, Wright, and Yogo (2002) presented that because of the observability of endogenous covariates and the instruments, the strength of the instruments can be directly assessed. Miguel, Satyanath, and Sergenti (2004) also employed the IV in their literature, they checked the weather shocks to identify the changes in economic growth on civil conflict. Heckman (2008) illustrated the relationship between IV and causality in econometrics.

A simple review of IV estimation is given below. In the simple linear regression,

$$Y = X * \beta + \epsilon \tag{4.1}$$

Where Y represents the dependent variable, X is the independent variable, β is a parameter, and ϵ is an unobserved error term representing all causes of Y other than X. Generally, X and ϵ are assumed uncorrelated. Thus from random sample of k observations, the ordinary least squares estimation can be written as

$$\hat{\beta}_{OLS} = \frac{x'y}{x'x} = \frac{x'(x\beta + \epsilon)}{x'x} = \beta + \frac{x'\epsilon}{x'x} \tag{4.2}$$

here x, y and ϵ stand for the column vectors. Under certain regularity conditions the second term has an expected value conditional on x of zero and converges to zero as the number of observations increase, so the estimator has mean $E(\hat{\beta}_{OLS}) = \beta$, the estimation is unbiased and consistent. However, this assumption can be violated in some situations, in which $E(x'\epsilon) \neq 0$, therefore the OLS estimation for regression coefficients are biased and inconsistent.

Under the above situation, we need other methods to handle the correlation between X and ϵ, and IV estimation is one of them. There are 3 circumstances which may lead us to use the IV method.

(1) There are measurement errors in X variables.

(2) The predicted variable has different lags from that of predictor variables which

can cause X and ϵ correlated.

(3) In many economics conditions, one or more X variables are endogenous, in which case X and Y may be influenced together by some other factors.

Our data belongs to the third circumstance. Both X and Y variables are potentially influenced by some other unobservable factors.

In the IV estimation, we need to select one or more variables $Z = (Z_1, Z_2, ..., Z_k)$, that are correlated with X but not with ϵ. The model used in the first step can be written as

$$X = \alpha_0 + Z_1 * \alpha_1 + Z_2 * \alpha_2 + Z_3 * \alpha_3 + ... + Z_k * \alpha_k + \upsilon \ , \tag{4.3}$$

or

$$X = Z * \alpha + \upsilon \ . \tag{4.4}$$

Here α is the coefficient of Z, and υ stands for the residuals of the model. Whether the instrumental variables Z are weak or not can be tested by a F-test. They are judged as weak if the F-test statistics is less than 10. If the instrumental variable are weak, the estimation for α may be biased and the sampling distributions of the estimators may be not normal, thus the inference of the model may not be reliable.

If the instrumental variables are strong, we can estimate (Y) using the model

$$Y = \hat{X} * \beta + \epsilon \ , \tag{4.5}$$

where

$$\hat{X} = Z * \hat{\alpha} \tag{4.6}$$

The IV estimate of β is

$$\hat{\beta_{IV}} = (Z'X)^{-1}Z'Y = (Z'X)^{-1}Z'(X\beta + \epsilon) = \beta + (Z'X)^{-1}Z'\epsilon \ . \tag{4.7}$$

Z and ϵ are uncorrelated here, therefore, in the second term of equation (4.1.7), $E(Z'\epsilon) = 0$, so that the estimation becomes unbiased and consistent.

4.2 Application of IV Estimation for ROE

We first build a linear model based on 2010 observations only. The model is

$$\begin{aligned} ROE^{2010} =& \beta_0 + \beta_1 GDP^{2010} + \beta_2 ODP^{2010} + \beta_3 TA^{2010} + \beta_4 TL^{2010} + \\ & \beta_5 TR^{2010} + \beta_6 CR^{2010} + \beta_7 DA^{2010} + \beta_8 PA^{2010} + \beta_9 DL^{2010} + \\ & \beta_{10} PL^{2010} + \beta_{11} DR^{2010} + \beta_{12} PR^{2010} + \epsilon_i \end{aligned} \tag{4.8}$$

This model is not a very good fit for the data, since the R^2 is 0.324, and adjusted R^2 is 0.192. The backwards deletion according to AIC resulted a reduced model:

$$ROE^{2010} = \beta_0 + \beta_1 ODP^{2010} + \beta_2 TA^{2010} + \beta_3 TL^{2010} + \beta_4 TR^{2010} + \beta_5 DL^{2010} + \epsilon_i \tag{4.9}$$

the superscript 2010 stands for the observations in 2010. The reduced model still has the R^2 and adjusted R^2 less than 0.4.

We will use the residual plots to check the randomness of residuals.

Figure 4.1: Residuals v.s Non-Hedging-Related Variables

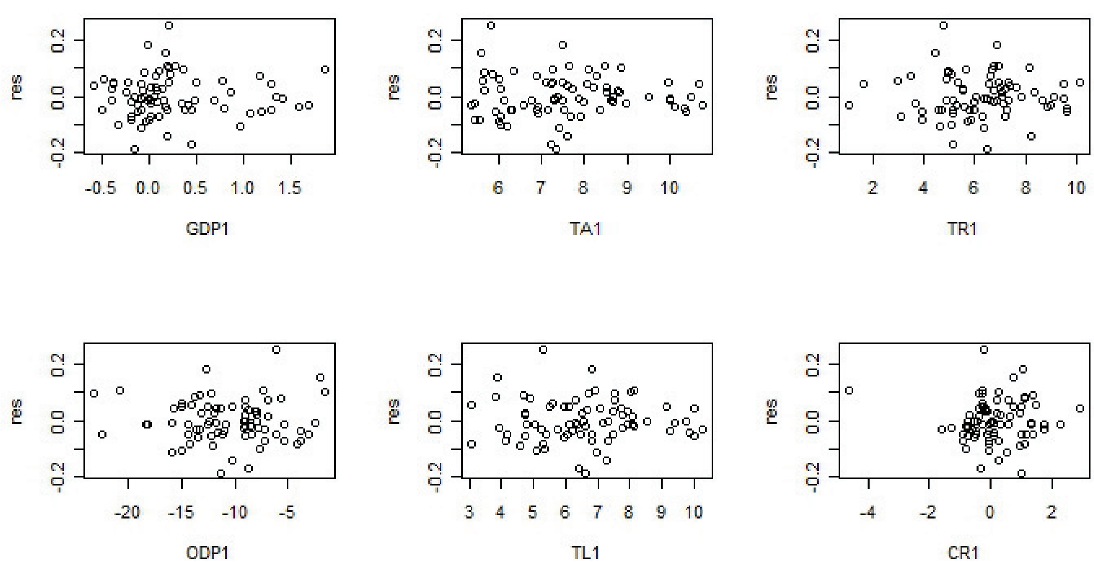
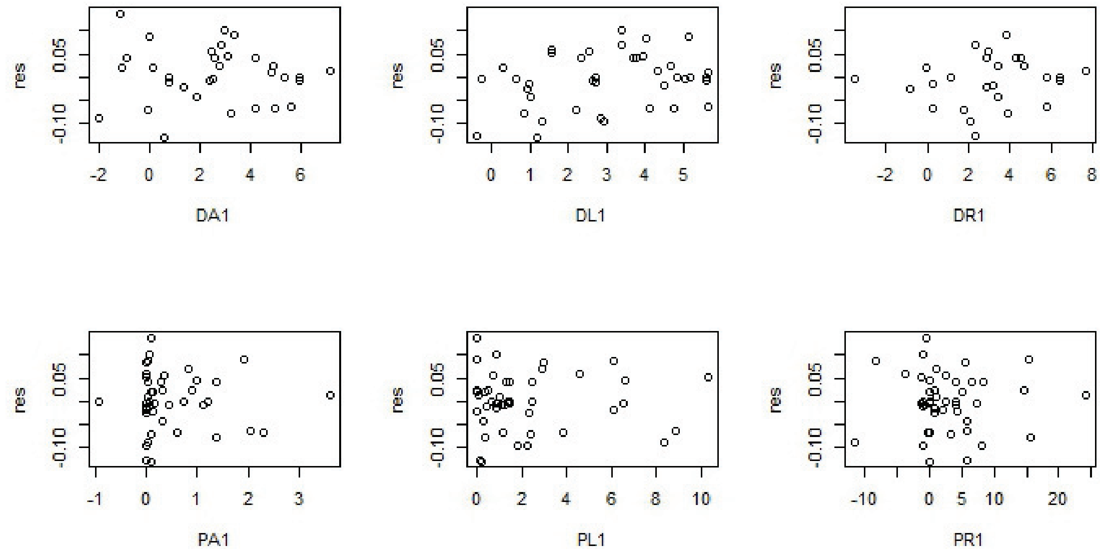


Figure 4.2: Residuals v.s Hedging-Related Variables



From the residual plots presented in Figures 4.1-4.2, we can see the residuals with first six variables $(GDP^{2010}, ODP^{2010}, TA^{2010}, TL^{2010}, TR^{2010}, CR^{2010})$ are quite random. For six hedging related variables $(DA^{2010}$, PA^{2010}, DL^{2010}, PL^{2010}, DR^{2010}, $PR^{2010})$, there exist a lot of 0's due to the absence of hedging activities for non-hedging companies. In order to show more clearly only the hedging variables, we take out those 0's here. We can see the slight non-randomness in PA, PL and PR plots.

In this particular dataset, we will use the data from year 2009 as the instrumental variables and predict ROE and TG for year 2010. The aim is to predict the current ROE or TG by the data from previous year.

After testing all the X variables by using the IV estimation and checking the validity, we select several variables in 2009 to be the instrumental variables according to the F-test[1],

[1] The principle is mentioned above, which can be simply expressed as rejecting the variable as instrumental variable when the F-test of model $\hat{X} = \hat{\alpha_0} + \hat{\alpha_1} * Z_{1i} + \hat{\alpha_1} * Z_{2i} + \hat{\alpha_1} * Z_{3i} + ...$ is less than 10; and choosing not to reject if the F-test is larger or equal to 10.

they are TA, TL, TR, DA, PA, DL and DR. We then build a new model,

$$\begin{aligned} ROE^{2010} = {} & \beta_0 + \beta_1 * GDP^{2010} + \beta_2 * ODP^{2010} + \beta_3 * \hat{TA}^{2010} + \beta_4 * \hat{TL}^{2010} \\ & + \beta_5 * \hat{TR}^{2010} + \beta_6 * CR^{2010} + \beta_7 * \hat{DA}^{2010} + \beta_8 * \hat{PA}^{2010} + \beta_9 * \hat{DL}^{2010} \\ & + \beta_{10} * PL^{2010} + \beta_{11} * \hat{DR}^{2010} + \beta_{12} * PR^{2010} + \epsilon_i \end{aligned} \tag{4.10}$$

where $\hat{X}_i = (\hat{TA}^{2010}, \hat{TL}^{2010}, \hat{TR}^{2010}, \hat{DA}^{2010}, \hat{PA}^{2010}, \hat{DL}^{2010}, \hat{DR}^{2010})$ represents the variables which are predicted by the instrumental variables in 2009. The best reduced model according to AIC is,

$$ROE^{2010} = \beta_0 + \beta_1 * \hat{TA}^{2010} + \beta_2 * \hat{TL}^{2010} + \beta_3 * \hat{TR}^{2010} + \epsilon_i \tag{4.11}$$

Table 4.1 below shows the model fitting information for the full model and the selected model,

Table 4.1: Results of Full and Reduced IV Models

Models	R^2	adjusted R^2	F-test	P-value
Full	0.35	0.22	2.73	0.005
Reduced	0.29	0.26	10.65	0.000

We note that the R^2 and adjusted R^2 are small in both models, which means these two models are not good fit for the data. The reason may be that the relationship between ROE and other variables is not linear.

4.3 Testing for the Need for Instrumental Variable

Comparing the model in IV estimation with the original model in OLS estimation, it is not clear which model is better to predict the ROE in 2010. Now, we will apply the Hausman test to check if the IV estimation is essentially better than the OLS estimation.

Hausman's specification test (Hausman, 1978) is based on the idea that the difference between two consistent estimators are close to 0. It is often used to test two estimations under the null hypothesis that two estimators are equivalent. One of the estimators is consistent under both the null and the alternative hypothesis, but the other estimator is consistent under null hypothesis, inconsistent under the alternative.

The idea that estimates by any two consistent methods should be close to each other is proposed by Durbin (1954). Wu (1973, 1974) employed a similar method to compare four different tests (two of them included IV) for testing the assumption that regressors in the linear regression model are statistically independent of the disturbance term. He proved that the test statistics is the basis for a new estimator that is the same as the OLS estimator if the null hypothesis of regressors exogeneity is accepted, or is equal to the IV estimator if the null hypothesis is rejected. He also provided the version of the Hausman test statistic which is formed by using the IV estimate of the error variance. Nakamura and Nakamura (1981) proved that the Hausman test is equal to the statistics proposed by Durbin (1954). Hausman and Taylor (1981) illustrated that all different versions of test statistic have a chi-square distribution with the degrees of freedom which is at most the number of potentially endogenous regressors. Gaston and Trefler (1994) applied the Hausman test to check the effects of international trade policy on wages of U.S manufacturing industries in 1983. Chmelarova (2007) reviewed the previous literatures and employed the Hausman test with heteroskedastic data.

We now employ the Hausman test to test the need for IV estimation compared to the OLS estimation. According to the null hypothesis, both the OLS and IV estimators are consistent and converge to each other, but the IV estimator is less efficient if the predictors and error term are uncorrelated. If the predictors and error term are correlated, the OLS estimator is inconsistent but the IV estimator is consistent, so the difference the two estimators does not converge to 0 as the sample size increases.

In the linear model,

$$Y = X * \beta + \epsilon \tag{4.12}$$

Assume we have two estimators, β_{OLS} and β_{IV}.The Hausman test statistics is

$$H = (\hat{\beta}_{IV} - \hat{\beta}_{OLS})' \, (Var(\hat{\beta}_{IV}) - Var(\hat{\beta}_{OLS}))^{-1} \, (\hat{\beta}_{IV} - \hat{\beta}_{OLS}) \tag{4.13}$$

or we can write it as

$$\begin{aligned} H &= \hat{\sigma}^2 \, (\hat{\beta}_{IV} - \hat{\beta})' \, [(\hat{X}'\hat{X})^{-1} - (X'X)^{-1}]^{-1} \, (\hat{\beta}_{IV} - \hat{\beta}) \\ where \;\; Var(\hat{\beta}_{IV}) - Var(\hat{\beta}) &= \hat{\sigma}^2 \, [(\hat{X}'\hat{X})^{-1} - (X'X)^{-1}] \end{aligned} \tag{4.14}$$

This test statistic has an asymptomatic chi-squared distribution with the degrees of freedom equal to the rank of $((\hat{X}'\hat{X})^{-1} - (X'X)^{-1})^{-1}$.

Our model with the OLS estimation is

$$\begin{aligned} ROE_{OLS} = \beta_0 + \beta_1 GDP + \beta_2 ODP + \beta_3 TA + \beta_4 TL + \beta_5 TR + \beta_6 CR \\ + \beta_7 DA + \beta_8 PA + \beta_9 DL + \beta_{10} PL + \beta_{11} DR + \beta_{12} PR + \epsilon_i \end{aligned}, \qquad (4.15)$$

and the model with IV estimation is,

$$\begin{aligned} ROE_{IV} = \beta_0 + \beta_1 * GDP + \beta_2 * ODP + \beta_3 * \hat{TA} + \beta_4 * \hat{TL} + \beta_5 * \hat{TR} + \beta_6 * CR \\ + \beta_7 * \hat{DA} + \beta_8 * \hat{PA} + \beta_9 * \hat{DL} + \beta_{10} * PL + \beta_{11} * \hat{DR} + \beta_{12} * PR + \epsilon_i \ . \end{aligned} \qquad (4.16)$$

Note that the OLS estimators are consistent and more efficient under the null hypothesis, but they are inconsistent if the null hypothesis is false. That means there is no need to use IV estimation if the null hypothesis is true since it will be less efficient than OLS estimation.

The test statistic of Hausman test for these two models, is 25.784, with the degree of freedom equal to 13, which gives us the p-value 0.0181 in chi-square test. In this case, we need to reject the null hypothesis. Thus, the IV estimation is better than the OLS estimation in our data.

4.4 Application of Instrumental Variable Estimation to Predict the Total Hedging Gains (TG)

As the ROE model above, we build a linear model for TG as follows,

$$\begin{aligned} TG^{2010} = \beta_0 + \beta_1 GDP^{2010} + \beta_2 ODP^{2010} + \beta_3 TA^{2010} + \beta_4 TL^{2010} + \beta_5 TR^{2010} \\ + \beta_6 CR^{2010} + \beta_7 DA^{2010} + \beta_8 PA^{2010} + \beta_9 DL^{2010} + \beta_{10} PL^{2010} \\ + \beta_{11} DR^{2010} + \beta_{12} PR^{2010} + \beta_{13} ROE^{2010} + \epsilon_i \end{aligned} \qquad (4.17)$$

This equation is similar to the linear model for ROE, except we add one more predictor, ROE. The OLS has a larger R^2 (0.88) and adjusted R^2 (0.86), which proves this model is a good fit for the data. Also, the F-test is about 34.7 and p-value is $2.2 * e^{-16}$, which shows the linear relationship between TG and the other variables is much stronger. Furthermore, stepwise model selection is employed to select the model with smallest AIC, which results in the model:

$$\begin{aligned} TG^{2010} &= \beta_0 + \beta_1 GDP^{2010} + \beta_2 TA^{2010} + \beta_3 TL^{2010} + \beta_4 CR^{2010} + \beta_5 DL^{2010} \\ &\quad + \beta_6 DR^{2010} + \beta_7 PR^{2010} + \beta_8 ROE^{2010} + \epsilon_i \ . \end{aligned} \tag{4.18}$$

For IV estimation, we choose seven predictors which meet the criteria that the F-test statistics has to be larger then 10. Hence the IV model is

$$\begin{aligned} TG^{2010} &= \beta_0 + \beta_1 * GDP^{2010} + \beta_2 * ODP^{2010} + \beta_3 * \hat{TA}^{2010} + \beta_4 * \hat{TL}^{2010} \\ &\quad + \beta_5 * \hat{TR}^{2010} + \beta_6 * CR^{2010} + \beta_7 * \hat{DA}^{2010} + \beta_8 * \hat{PA}^{2010} + \beta_9 * \hat{DL}^{2010} \\ &\quad + \beta_{10} * PL^{2010} + \beta_{11} * \hat{DR}^{2010} + \beta_{12} * PR^{2010} + \beta_{13} * ROE^{2010} + \epsilon_i \end{aligned} \tag{4.19}$$

same as before, $\hat{X}^{2010}$ represents the instrumental variable which are predicted by the data in 2009. And the reduced IV model is

$$\begin{aligned} TG^{2010} &= \beta_0 + \beta_1 GDP^{2010} + \beta_2 ODP^{2010} + \beta_3 * \hat{TL}^{2010} + \beta_4 CR^{2010} + \beta_5 * \hat{DA}^{2010} \\ &\quad + \beta_6 * \hat{DL}^{2010} + \beta_7 PR^{2010} + \beta_8 ROE^{2010} + \epsilon_i \end{aligned} \tag{4.20}$$

The model also fits very well with R^2 and adjusted R^2 equal to 0.87 and 0.85, respectively, and F-test is 53.56, which is much larger than 10, it shows the IV estimation can be used in TG.

We still apply the Hausman tests to check if there exists any advantage of IV estimation compared with the OLS estimation.

Two different models which contain OLS estimation and IV estimation are,

$$\begin{aligned} TG_{OLS} &= \beta_0 + \beta_1 GDP^{2010} + \beta_2 ODP^{2010} + \beta_3 TA^{2010} + \beta_4 TL^{2010} + \beta_5 TR^{2010} \\ &\quad + \beta_6 CR^{2010} + \beta_7 DA^{2010} + \beta_8 PA^{2010} + \beta_9 DL^{2010} + \beta_{10} PL^{2010} \\ &\quad + \beta_{11} DR^{2010} + \beta_{12} PR^{2010} + \beta_{13} ROE^{2010} + \epsilon_i \ , \end{aligned} \tag{4.21}$$

$$\begin{aligned} TG_{IV} = \beta_0 + \beta_1 * GDP^{2010} + \beta_2 * ODP^{2010} + \beta_3 * \hat{TA}^{2010} + \beta_4 * \hat{TL}^{2010} \\ + \beta_5 * \hat{TR}^{2010} + \beta_6 * CR^{2010} + \beta_7 * \hat{DA}^{2010} + \beta_8 * \hat{PA}^{2010} + \beta_9 * \hat{DL}^{2010} \\ + \beta_{10} * PL^{2010} + \beta_{11} * \hat{DR}^{2010} + \beta_{12} * PR^{2010} + \beta_{13} * ROE^{2010} + \epsilon_i \ . \end{aligned} \tag{4.22}$$

Same as Section 4.3, we test the null hypothesis that the difference between coefficients of OLS estimation and IV estimation are not systematic. The alternative hypothesis is IV estimator is consistent, where OLS is inconsistent. We get the test statistic as 23.686, at the degree of freedom of 14, which resulted approximately a p-value 0.05. Hence, the IV estimation is also necessary in TG. The IV estimation is theoretically capable to predict TG by applying $X = (GDP, ODP, \hat{TA}, \hat{TL}, \hat{TR}, CR, \hat{DA}, \hat{PA}, \hat{DL}, PL, \hat{DR}, PR, ROE)$.

CHAPTER 5

FINITE MIXTURE OF REGRESSION MODELS

In Chapter 4, we have applied the IV estimation to our data, and showed that using IV estimation method is better than an OLS fitting. However the model still does not fit well, illustrated by the small R^2 values. Intuitively different firms have different hedging regulation and strategies, depending on their sizes and types etc.. Among those firms which apply similar hedging strategies, the relationship may be better modelled by a linear regression model. Thus we attempt to use the finite mixture of regression models in this Chapter. The predicted variables will still be ROE and TG.

Finite mixture of regression (FMR) models are powerful models to fit data when data can be better described by several sub-populations. The memberships of data to the sub-populations need not to be pre-specified, instead a ”soft” membership of each observation to sub-populations can be derived out from the model. Finite mixture of regression models are model based clustering based on regression model fitting to each sub-population.

We will first review the FMR models in Section 5.1, we then employ FMR models to predict ROE using non-hedging related variables and all variables respectively in Sections 5.2 and 5.3. Finally we apply FMR models to predict TG in Section 5.4.

5.1 A review of finite mixture of regression models

Finite mixture of regression (FMR) models are used to model the data where observations come from several different groups and the group affiliations are unknown. It provides a natural representation of multi-modal distributions by a finite number of latent classes.

Although it's been used over a hundred years, it only becomes popular in last twenty years with the development of the available computing power. FMR models have been applied in many fields, include biology, physics, economics, and marketing.

The expectation-maximization (EM) algorithm (Dempster et al. ,1977) is commonly used to estimate the parameters within a maximum likelihood framework for mixture models. Wedel and DeSarbo (1991) extended the finite mixture models by mixing standard linear regression models as well as generalized linear models. Jacobs, Jordan, Nowlan, and Hinton (1991) developed the finite mixture regression (FMR) models in machine learning under the mixture of experts models. Diebolt and Robert (1994) estimated the finite mixture models with MCMC sampling within the Bayesian framework with a fixed number of components. The important development of finite mixture models have been summarized by Wedel & Kamukura (2000) and Skrondal & Rabe Hesketh (2004), they also applied and improved the use of finite mixture of regression models in market segmentation and social science to substitute the cluster analysis and cluster-wise regression techniques. Leisch (2003) created a R package "FlexMix" for estimation of FMR models, which we will use to analyse our data.

We adopt the notations from Leisch (2003) to write the density function for a k-component FMR model as

$$\begin{aligned} f(y|x;\theta;\pi) &= \sum_{i=1}^{k} \pi_i f_i(y|x;\theta_i) \\ \pi_i \geqslant 0, &\quad \sum_{i=1}^{k} \pi_i = 1 \end{aligned} \tag{5.1}$$

where y is the predicted variable and x represents the predictors. The parameters $\pi_i = (\pi_1, \pi_2, ..., \pi_k)$ is the prior probability of k components, when θ_i stands for the component specific parameters for density function f_i. We can collect all parameters in a vector

$$\psi = (\pi_1, \pi_2, ..., \pi_k; \theta_1^T, \theta_2^T, ..., \theta_k^T)^T \tag{5.2}$$

Model 5.1.1 is also termed as latent class regression if f_i is a univariate normal density with mean $\beta_i' x$ and standard deviation σ_i, where $\theta_i = (\beta_i', \sigma_i)'$.

Generally, maximizing the log-likelihood function directly is computationally demanding, since the log-likelihood function

$$log\ L = \sum_{n=1}^{N}(log(\sum_{i=1}^{k}\pi_i f_i(y_n|x_n,\theta_i))) \tag{5.3}$$

has summation behind the logarithm.

EM algorithm maximizes instead an expected complete data log-likelihood by augmenting data with their latent class variables. More specifically:

E-step: the posterior probabilities of n^{th} observation belongs to the i^{th} component given the current parameter estimate $\hat{\psi}$ is calculated as

$$\hat{P}_{ni} = \Pr(i|x_n, y_n, \hat{\psi}) \tag{5.4}$$

so the prior probabilities are updated:

$$\hat{\pi}_i = \frac{1}{N}\sum_{n=1}^{N}\hat{P}_{ni}\ . \tag{5.5}$$

M-step: maximizing the expected log-likelihood function given the prior probabilities:

$$\max_{\theta_i, i=1,\cdots,k}\ \sum_{n=1}^{N}\sum_{i=1}^{k}\hat{P}_{ni}\ log(\hat{\pi}_i f_i(y_n|x_n,\theta_i))\ . \tag{5.6}$$

The E-step and M-step are iterated until the likelihood improvement goes under the specified criterion or the program reaches the maximum number of iterations.

The posterior probability that observation (x,y) is from the component c is given by

$$\Pr(c|x,y,\psi) = \frac{\pi_c f(y|x,\theta_c)}{\sum_i \pi_i f(y|x,\theta_i)} \tag{5.7}$$

The posterior probabilities are often used to assign the observations into different components by choosing the maximum posterior probability.

The R package “flexmix” developed by Leisch (2003) aims to provide the convenience for the rapid development of new mixture models and can in principle fit most of the finite mixture models. As the EM algorithm can only converge to a local maximum of the likelihood, thus multiple times of fitting with different starting values are needed. The function “stepflexmix” facilitates such multiple fittings.

Standard methodology for model-based clustering is to use the EM algorithm to estimate the finite mixture models corresponding to each number of clusters, and then employ BIC to select the number of mixture components, taken to be equal to the number of clusters (Fraley and Raftery, 1998). As an approximation to the integrated likelihood, Bayesian Information Criterion (BIC) has been commonly used for model selection. However, BIC has some drawbacks, such as if the correct model is not in the family of considered models, BIC criterion will overestimate the correct size, regardless the cluster separation. Biernacki, Celeux, and Govaert (2000) proposed the integrated classification likelihood (ICL) criterion, which is to assess the number of mixture components through integrated likelihood of complete data. The function “ICL” in the R package “flexmix” computes the ICL criterion to select the number of mixture components.

5.2 Predicting ROE using FMR models with Non-hedging-related Variables

For each company, there are two observations for year 2009 and 2010 respectively. Thus ideally we should use finite mixture of linear mixed regression models. We could not find a package to fit such models and it is beyond this thesis to develop these models. Thus in the rest of this chapter, all the models are fit by ignoring the correlation between observations.

We have separated the variables into two parts. *Hedging-related variables* include variables that describe hedging activities, and *non-hedging-related variables* are the variables not related to the hedging activities. As we mentioned above, ROE is equal to the net income divided by shareholder equity, it can also be calculated as $ROE = Net\ Income/(TA - TL)$, where Net income is a part of TR. We find that all the elements in this equation are from non-hedging-related variables. Hence, we will first check the relationship between ROE and non-hedging-related variables by using FMR models.

Using “stepflexmix” function in “flexmix” (Leisch, 2003), we can easily fit FMR models for a set of increasing number of components. For each choice of number of components, “stepflexmix” repeatedly fit models using different starting values for K times with K specified by user, and returns a list of the best solutions found. We then compare models with different number of components by applying ICL criterion. As a reference, we have also calculated the AIC and BIC values for each model.

Table 5.1: Results of Selecting Number of Components with Non-hedging-related Variables

	iterations	converged	k	k0	log-likelihood	AIC	BIC	ICL
1	2	TRUE	1	1	130.628	-245.256	-221.1709	-221.1709
2	102	TRUE	2	2	158.5879	-283.1759	-231.9951	-223.5226
3	119	TRUE	3	3	223.5759	-395.1519	-316.8753	-296.5006
4	200	FALSE	4	4	240.613	-411.2259	-305.8537	-244.7269
5	85	TRUE	5	5	250.0242	-412.0483	-279.5804	-193.6537
6	163	TRUE	6	6	284.1216	-462.2432	-302.6795	-256.4638
7	158	TRUE	7	7	306.3999	-488.7999	-302.1405	-255.0656
8	111	TRUE	8	8	324.5143	-507.0286	-293.2735	-250.5963
9	142	TRUE	9	9	335.6881	-511.3762	-270.5254	-244.756
10	80	TRUE	9	10	356.3639	-552.7277	-311.8769	-277.2798

Notes: k_0 represents the number of components that we suggest, and k means the actual number of components that software clustered.

For each number of components, $k = (1, 2, ..., 10)$, We estimate the models 100 times. Table 5.1 shows the best log-likelihoods resulted from 100 times running for each k value. The smallest ICL happens at $k = 5$. Therefore we choose five components here.

We next verify how separate the five components are. We employ the command "ratio" in "flexmix" (Leisch ,2003) to calculate how much overlaps exist between components. For the i^{th} component, the ratio is defined as the number of observations assigned to it divided by the number of observations whose posterior probabilities $p_{ni} > 0$. For a well separated component, a large proportion of observations with non-zero posteriors probabilities should be assigned to the component, giving a ratio close to 1. Table 5.2 shows the resulting ratios for our data.

Table 5.2: Ratio values for five components for the Non-hedging-related Variables model

	prior	size	post>0	ratio
Comp.1	0.0884	11	58	0.190
Comp.2	0.5179	71	135	0.526
Comp.3	0.1268	23	29	0.793
Comp.4	0.0660	10	10	1.000
Comp.5	0.2009	35	53	0.660

Table 5.2 shows four out of the five ratios are greater than 0.5. That means the overlap of the components is small. However there is one component with small ratio.

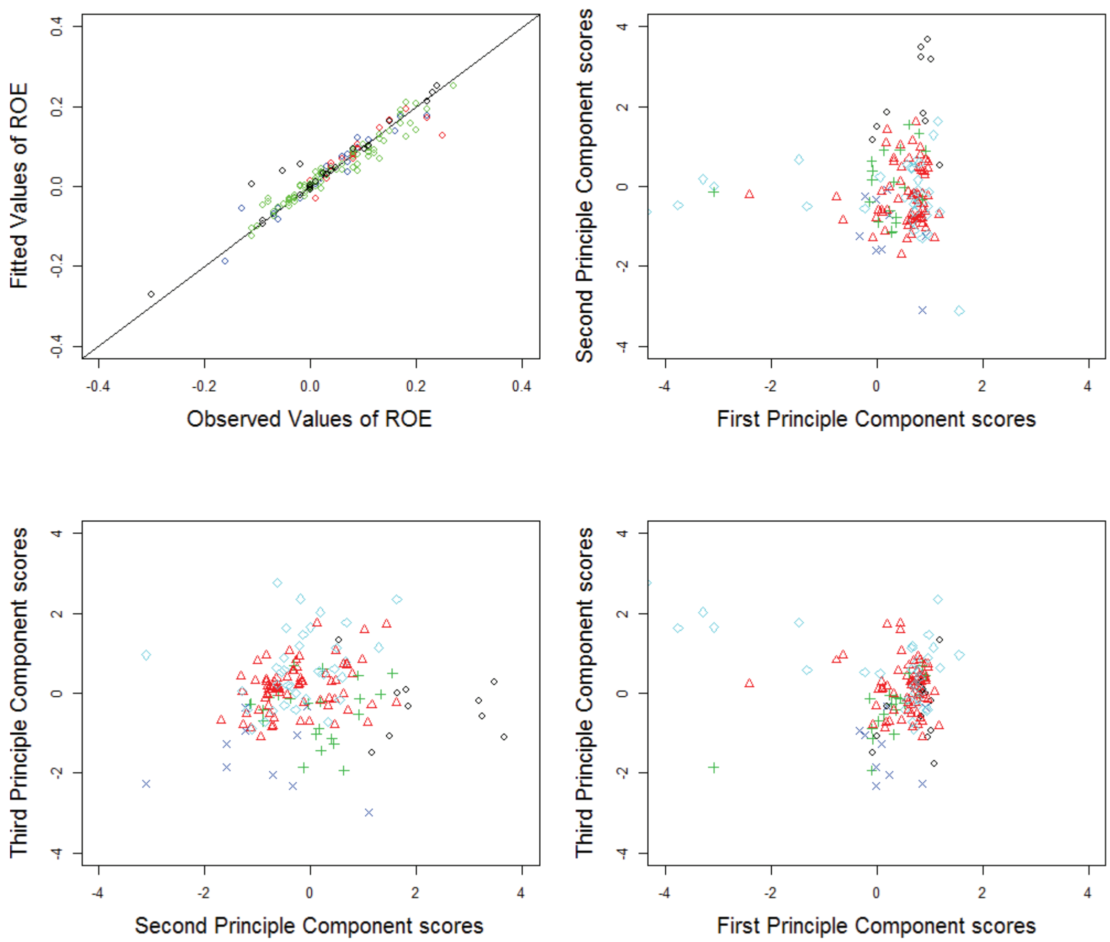
Using "summary" function in package "flexmix" (Leisch, 2003), we can extract all the parameter estimates and their standard errors, t-statistics and p-values, based on which we can perform some basic variable selection within each component. We remove the insignificant predictors once a time within each component and refit the models until all the variables are significant. The final results are shown in Table 5.3:

Table 5.3: FMR Models with Non-hedging-related Variables as predictors

	Comp 1	Comp 2	Comp 3	Comp 4	Comp 5
Intercept	2.34	0.26		0.23	-0.13
GDP	-0.20	0.05	-0.03	0.02	0.02
ODP	0.04			0.02	
TA	-0.27	-0.18	-0.11	0.01	-0.02
TL	-0.07	0.07	0.05	-0.05	0.02
TR	0.07	0.10	0.09	0.03	
CR	-0.02	-0.03	0.08	0.04	

By checking the observations assigned to each component, we find the variable TA is an important issue for ROE. Component four contains most of companies with TA values larger than five thousands million Canadian dollars. Component one consists the smallest TA values.

Figure 5.1: Fitted Values v.s Observed Values and FRM components presented on Principle Components of predictor variables



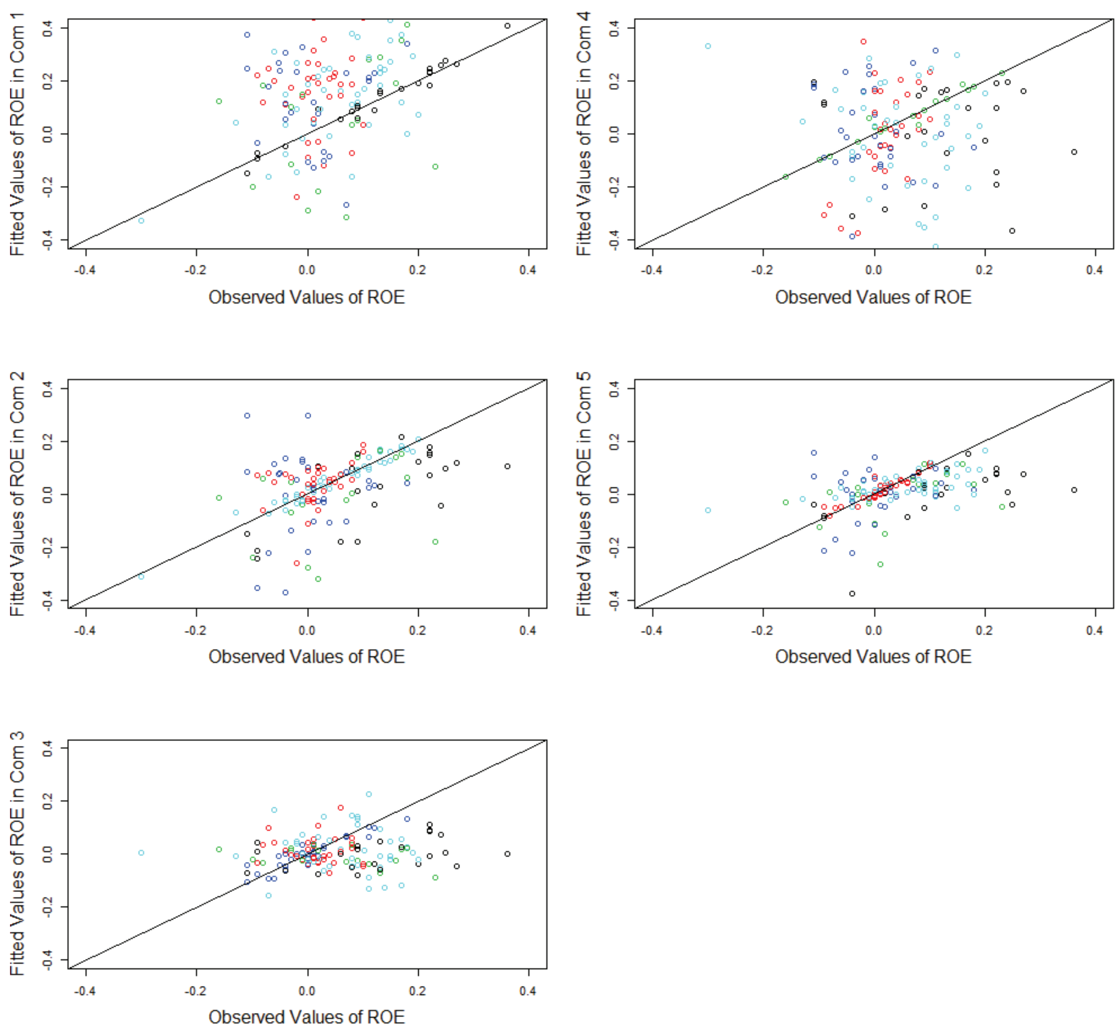
Notes: Component 1, 2, 3, 4 and 5 are black, red, green, dark blue and light blue, respectively.

The top left panel of Figure 5.1 shows that the fitted values of ROE are close to the observed ROE values, which means the FMR models with non-hedging-related variables predict the ROE well. To find which companies are clustered into the same components, we present the five components with five different colours in the scores plots using the first three principal components of the predictor variables in the other three panels in Figure 5.1. The best separation is found in the second and third PC scores plot. The loadings for the principal components [1] can be used to find the features of the components that are more clearly separated in the plot. It seems the companies in component 1 has larger CR and small GDP; the component 4 contains companies with small CR and GDP; most of the companies in component 3 have small GDP and the companies in component 5 contain large value of GDP and TR.

We also plotted fitted values from different components versus observed values of ROE. They are shown in Figure 5.2.

[1]The first principal component is a linear combination of the variables: $-0.245*GDP+0.011*ODP+0.574*TA+0.559*TL+0.515*TR+0.128*CR$.
The second principal component is a linear combination of the variables: $0.441*GDP-0.595*ODP-0.013*TA-0.004*TL-0.036*TR+0.671*CR$.
The third principal component is a linear combination of variables: $0.688*GDP-0.659*ODP+0.026*TA-0.021*TL+0.277*TR-0.118*CR$.

Figure 5.2: Fitted Values of ROE from Each Component v.s Observed Values of ROE



Notes: Component 1, 2, 3, 4 and 5 are black, red, green, dark blue and light blue, respectively.

From Figure 5.2, we can see that the points assigned to each component are accurately predicted by the regression model from the corresponding component. The points assigned to other components are spread around the reference line. Thus, FMR models well separate the components, since the prediction of ROE from each component is accurate, and the models of different components are totally different.

5.3 Predicting ROE using FMR models with All Variables

To observe if Hedging-related variables will improve the prediction accuracy and how hedging activity can change the ROE variable, we now re-do the same exercise as in Section 5.2 by including all the Hedging-related variables.

We still use "stepflexmix" function (Leisch (2003)) to fit FMR models for a set of increasing number of components, k = (1,2,..., 10). Then we choose the number of components by the smallest ICL.

Table 5.4: Results of Selecting Number of Components with Hedging-related Variables

	iterations	converged	k	k0	log-likelihood	AIC	BIC	ICL
1	2	TRUE	1	1	133.8701	-237.7402	-192.5806	-292.5806
2	54	TRUE	2	2	181.8397	-301.6794	-208.3497	-296.2016
3	99	TRUE	3	3	279.4081	-464.8162	-323.3163	-313.8661
4	158	TRUE	4	4	308.1305	-490.261	-300.591	-286.9139
5	200	FALSE	5	5	388.4381	-618.8761	-381.0359	-352.5508
6	103	TRUE	6	6	429.735	-669.47	-383.4597	-354.1821
7	143	TRUE	7	7	487.7699	-753.5397	-419.3592	-409.7379
8	67	TRUE	8	8	543.1177	-832.2353	-449.8847	-433.5711
9	118	TRUE	8	9	554.4463	-854.8926	-472.5419	-458.9622
10	53	TRUE	9	10	603.8797	-921.7594	-491.2385	-478.1179

Notes: k_0 represents the number of components that we suggest, and k is the actual number of components that the software clustered.

From Table 5.4, the smallest ICL occurs at $k = 4$. Thus there are four components in our FMR models. Table 5.5 presents the ratios of overlap for these 4 components.

Table 5.5: Probability of Ratio Test in Four Components with All Variables

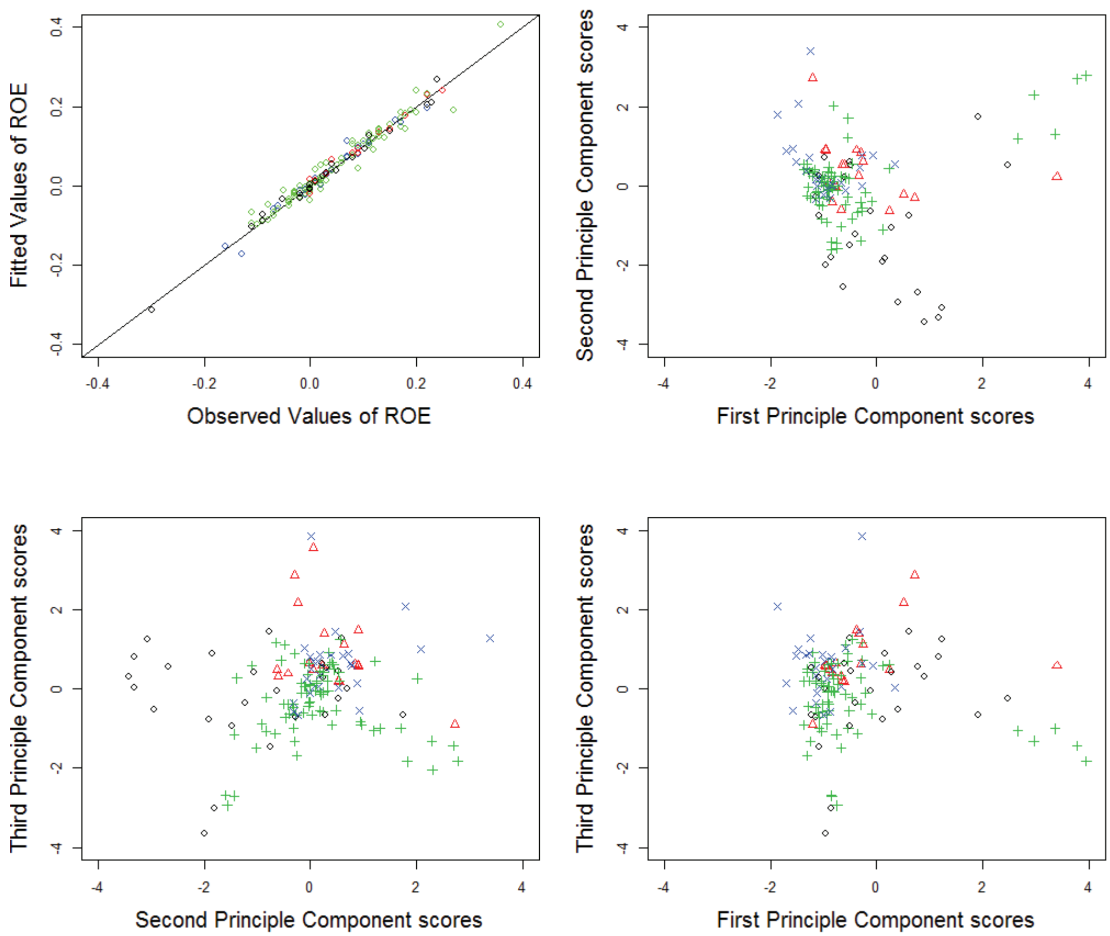
	prior	size	post>0	ratio
Comp.1	0.173	28	34	0.824
Comp.2	0.110	18	21	0.857
Comp.3	0.534	73	140	0.521
Comp.4	0.183	31	38	0.816

After performing the model selection procedures within each component, the final models are presented in Table 5.6.

Table 5.6: FMR Models with All Variables

	Comp 1	Comp 2	Comp 3	Comp 4
Intercept	0.994	0.181		-0.132
GDP	0.383	0.018		
ODP				0.005
TA	-0.316	-0.112		0.108
TL	0.034	0.034		-0.032
TR	0.205	0.085	0.056	
CR	-0.149		0.027	
DA				
PA	-0.332	0.077		0.142
DL				
PL	0.012	-0.008	-0.002	0.013
DR				
PR	0.046	0.011	0.005	0.005

Figure 5.3: Fitted Values v.s Observed Values and FMR components presented on Principle Components of predictors



Notes: Component 1, 2, 3 and 4 are black, red, green and blue, respectively.

The first panel in Figure 5.3 shows that the FMR models can predict ROE accurately. The other three panels shows the FMR components information over the first three PC's of the predictor variables. Same as in Section 5.1, we find that the plot of second and third principle component gives better separation effect[2]. We find component 1 contains most of the companies with small PR, and we also find that ROE in this component is not high either; the companies in component 2 has large CR; component 3 has large number of companies, and most of them have small CR; and the companies in component 4 also have large CR, and their PR is not low. From the above results, we find that except the non-hedging variables, ROE is also related to PR, which is one of the hedging-related variables. That proves not only the non-hedging variables, we also need to use hedging-related variables to predict ROE.

We then check the fitted values from different components versus the observed values of ROE. The results are similar with Figure 5.2, which means the separation of components are reasonable, when the prediction of ROE from different components are quite different.

5.4 Predicting TG using FMR models with All Variables

In this Section we apply FMR models to predict the realized total hedging gains (TG). Same as before, we use "stepflexmix" function (Leisch (2003)) to fit the models using different starting values with component number k range from 1 to 10 and find the component number with the smallest ICL criterion. We also calculated the AIC and BIC values for each model as a reference:

[2]The first principal component is a linear combination of the variables: $0.384 * DA + 0.279 * PA + 0.256 * DL + 0.099 * PL + 0.380 * DR + 0.247 * PR - 0.211 * GDP - 0.012 * ODP + 0.411 * TA + 0.389 * TL + 0.350 * TR - 0.074 * CR$
The second principal component is a linear combination of the variables: $-0.169 * DA - 0.267 * PA + 0.346 * DL + 0.254 * PL - 0.170 * DR + 0.479 * PR - 0.433 * GDP + 0.150 * ODP + 0.236 * TA + 0.235 * TL + 0.269 * TR + 0.251 * CR$.
The third principal component is a linear combination of the variables: $0.112 * DA + 0.393 * PA - 0.149 * DL + 0.221 * PL + 0.087 * DR + 0.133 * PR - 0.228 * GDP + 0.522 * ODP - 0.143 * TA - 0.187 * TL - 0.169 * TR + 0.573 * CR$.

Table 5.7: Results of Selecting Number of Components for FMR models to predict TG

	iteration	converged	k	k0	log-likelihood	AIC	BIC	ICL
1	2	TRUE	1	1	-855.7027	1741.4054	1786.565	1786.565
2	25	TRUE	2	2	-725.799	1513.5981	1606.928	1633.113
3	97	TRUE	3	3	-599.8038	1293.6075	1435.107	1436.247
4	134	TRUE	4	4	-522.0484	1170.0968	1359.767	1367.537
5	72	TRUE	5	5	-470.669	1099.338	1337.178	1357.178
6	62	TRUE	6	6	-443.5009	1077.0018	1363.012	1383.889
7	125	TRUE	7	7	-386.0729	994.1458	1328.326	1354.001
8	84	TRUE	7	8	-382.3582	986.7165	1320.897	1293.901
9	57	TRUE	8	9	-322.1724	898.3448	1280.695	1335.472
10	80	TRUE	8	10	-343.5727	941.1453	1323.496	1336.863

From Table 5.7, the smallest ICL can be found at $k = 7$. Therefore we separate the data into 7 components. After that, we verify how separate the seven components are. Table 5.8 gives the results of ratios.

Table 5.8: Probability of Ratio Test in Seven Components

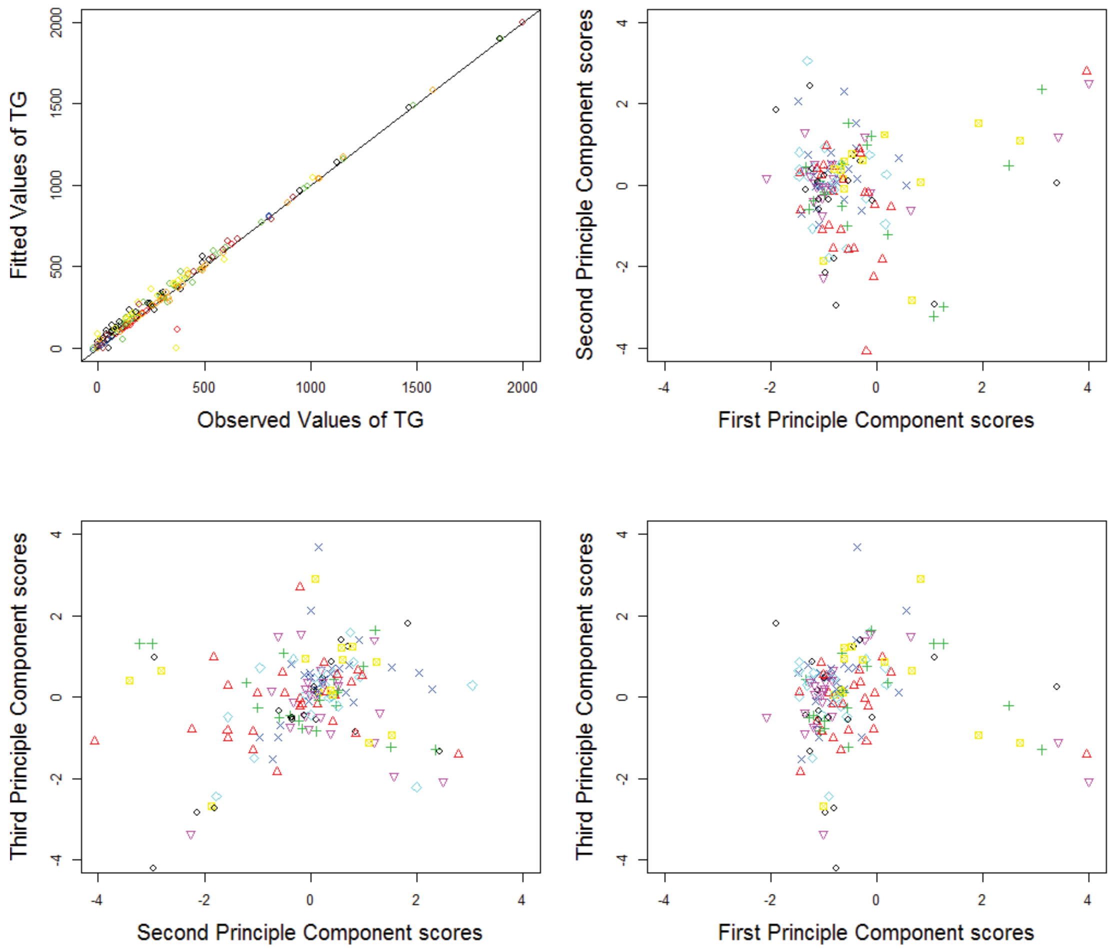
	prior	size	post>0	ratio
Comp.1	0.1214	19	21	0.905
Comp.2	0.1596	27	30	0.900
Comp.3	0.1314	17	34	0.500
Comp.4	0.1798	27	53	0.509
Comp.5	0.1399	22	26	0.846
Comp.6	0.1764	24	45	0.533
Comp.7	0.0916	14	17	0.824

Table 5.8 shows a good separation result, four ratios are larger than 0.8. Using backwards deletion within each component, we get the final model shown in Table 5.9.

Table 5.9: FMR models to predict TG

	Comp 1	Comp 2	Comp 3	Comp 4	Comp 5	Comp 6	Comp 7
Intercept	-142.62	33.33	49.97	-24.84	-1.95	-23.52	393.79
GDP	37.51	-15.77	-37.23	6.39	-18.48	-4.02	-64.98
ODP	7.05	0.21	1.05	0.56	0.85		-6.02
TA	43.96	-14.16		5.08	-1.02	13.25	67.15
TL	-1.85	18.61	-8.51		13.35	-12.32	7.03
TR	-8.12	-2.72	6.51	-1.72	-7.66	2.02	-23.64
CR	3.83	7.13	8.04	-3.12	6.90	-1.56	-7.82
ROE	-192.26		119.21		-24.09	47.51	458.52
DA	-2.04	-1.24	0.40	-0.63	-0.98		-2.75
PA	-38.81	8.83	-21.50	9.52	37.60	-8.97	30.59
DL	0.56	0.52	-0.18	0.47		0.46	
PL	6.53	-5.43	6.66	0.75	1.09	-3.02	3.26
DR	1.03	0.66		0.48	0.83		1.72
PR	1.11	0.20	9.09	0.47	1.16	2.36	-1.69

Figure 5.4: Fitted Values v.s Observed Values and FMR components on Principle Components of predictors



The first panel in Figure 5.4 shows the fitted values of TG are close to the observed TG values, thus FMR models can fit the data well. The other three panels in Figure 5.4 represent the seven components with seven different colours using the score plots of first three principal components of the predictor variables. However, the separation of different components can't be shown on these plots.

The prediction of ROE by different components are quite different after checking the plots of fitted values from different components versus the observed values of TG. Thus, the separation of components is good.

By carefully checking the observations assigned to different components, we find component four contains small PR and CR, and companies in component six have small CR but large PR. Component seven contains the companies with the biggest TA, and most of companies with large value of TA are in component seven (mean of TA is 7785.95). The other large companies are assigned to component five (mean of TA is 6908.77). Compared to the mean of ROE (0.13) in component seven, the companies in component five have smaller ROE mean (0.01). On the contrary, component two and four contain most of the companies with small TA (means of them are 4068.52 and 3767.68, respectively). Between these two components, Component two has larger CR than that of component four.

For the hedging-related variables, component six includes the companies with larger PA (0.97), component three contains the companies with larger PL (2.67). This situation shows the companies without large TA and applied hedging activities are assigned to these two components. The hedging companies are also found in component five and seven since they have the companies with the large values of DA (125.47 and 161.73) and DR (0.32 and 0.38). This proves again that the companies with large TA are willing to hedge.

Table 5.10 shows the features of each component.

Table 5.10: Components Characteristics

	Characteristics
Com 1	Contains the medium-sized companies which do not employ hedging activities.
Com 2	Contains the companies with the small TA, TL and TR, and employ the hedging activities.
Com 3	Contains the medium-sized companies with large PL, and employ the hedging activities.
Com 4	Contains the companies with the small TA, TL and TR, and do not employ the hedging activity. Meanwhile, PR and CR in it are also small.
Com 5	Contains the companies with large TA, TL and TR, and with small ROE. Most of them employ the hedging activities.
Com 6	Contains the medium-sized companies with large PA, and employ the hedging activities. It has large PR, and small CR.
Com 7	Contains the companies with large TA, TL and TR, and with large ROE. Most of them employ the hedging activities.

Chapter 6

Concluding Remarks

The main aim of this thesis is to find out the role of the hedging activities on return on equity and realized hedging gains. We gather the data from the annual reports of 75 oil and gas companies in Canada during 2009-2010. In order to deal with missing value problem, we employed multiple imputation method, more specifically fully conditional specification, to impute the missing values in the data. We then applied IV estimation and finite mixture of regression models, to analyse the data to find whether there exists relationship between hedging activities and ROE or TG.

Using IV estimation to check the relationship of the variables between 2010 and 2009, to predict ROE and TG using the data from previous year, we concluded that IV estimation is better than the OLS estimation. In IV estimation of predicting ROE, we discovered that ROE is strongly connected with TA. ROE is not only related with non-hedging variables, but also influenced by hedging variables.

We then employed the finite mixture of regression models which can separate the data into several components. We find that there exists relationship between ROE and hedging-related variables, especially PR. High value of PR may increase ROE, which proves that hedging activities can influence the return on equity. We also verify again that big companies love to employ hedging activities. CR is also an important variable since this variable is closely related to how data were separated into different components according to FMR models. FMR models are also used to predict the relationship between TG and the other variables and we find that different firms have different hedging strategies, which are associated with different patterns of profits in hedging.

6.1 Further Research

There are many issues left for the further research. Firstly, we do not have enough resources to test the relationship between companies' hedging activities and their stock returns, they may be strongly related. Secondly, as we compare hedging companies with non-hedging companies, we also find that CR presents the diversity that is caused by hedging activities. Further analysis using the obtained data on relationship of hedging and CR should be interesting. Thirdly, due to the underestimation or mis-report of hedging results in companies' reports, some biases caused may influence the analysis results. Finally, we have ignored the correlation between data in 2010 and 2009, This may influence the result.

Appendix A

A.1 Aspects of Hedging

Figure A.1: Aspects of Hedging

- Hedging
  - Reduce financial distress cost and avert the underinvestment
    - Smith and Stulz (1985)
    - **[state-preference model]**
    - Froot et al. (1993)
    - **[variation of CSV]**
  - Reduce the tax payments
    - Smith and Stulz (1985)
    - **[state-preference model]**
    - Graham and Smith (1999)
    - **[simulation methods]**
    - Graham and Rogers (2002)
    - **[tax function convexity]**
  - Moderate the risk exposure
    - Smith and Stulz (1985)
    - **[state-preference model]**
    - Barbara Pirchegger(2006)
    - **[two period LEN-type agency model]**
    - Aziz A. Lookman (2011)
    - **[multiple tests]**
    - Viral Acharya et al. (2012)
    - **[general equilibrium model]**

A.2 Table of Literatures

NAME	Smith, Clifford W. and Rene M. Stulz (1985)
TITLE	The determinants of firm's hedging policies
DATA	No data
METHOD	State-preference model
CONCLUSION	Use model to present three reasons for hedging: tax, cost of financial distress and managerial risk aversion
NAME	Froot, Kenneth A., David S. Scharfstein and Jeremy C. Stein (1993)
TITLE	Risk management: coordinating corporate investment and financing policies
DATA	No data
METHOD	Simple model of the benefits of hedging and the variation of costly-state verification(CSV)
CONCLUSION	When external finance cost more than internally generated sources of funds, the firms will hedge. And there are 7 implications
NAME	Gagnon, Louis, Gregory J. Lypny, and Thomas H. McCurdy (1998)
TITLE	Hedging foreign currency portfolio
DATA	No data
METHOD	Model the joint evolution of daily spot portfolio returns and log-differences of corresponding futures prices in a GARCH system
CONCLUSION	Accounting for portfolio effects in constructing a multi-currency hedge leads to efficiency and utility gains

NAME	Rajgopal, Shivaram (1999)
TITLE	Early evidence on the informativeness of the SEC's market risk disclosures: The case of commodity price risk exposure of oil and gas producers
DATA	Risk disclosures of thirty eight U.S. oil and gas companies in the Securities and Exchange Commission (SEC) market
CONCLUSION	The reserve of oil and gas will influence positively the relationship between oil and gas price and stock returns
NAME	Graham, John R. and C.W. Smit (1999)
TITLE	Tax incentives to hedge
DATA	Historical data of 84,200 firm-year observations during 1980-1994
METHOD	Simulation methods: coporated convex tax functions and estimations of the potential tax savings from hedging.
CONCLUSION	Hedging can significantly reduce the expected tax cost
NAME	Allayannis, George and James P. Weston (2001)
TITLE	The use of foreign currency derivatives and firm market value
DATA	720 U.S. non-financial firms between 1990 and 1995
METHOD	Tobin's Q ratio
CONCLUSION	Hedging has the positive relationship with the firm value
NAME	Graham, John R. and Daniel A. Rogers (2002)
TITLE	Do firms hedge in response to tax incentives

DATA	International Swaps and Derivatives Association (ISDA) reports and Sample firms sample firms at the end of fiscal 1994 or 1995
METHOD	Using an explicit measure of tax function convexity
CONCLUSION	No evidence that firms hedge in response to tax convexity and also find that hedging can increase debt capacity
NAME	Jin, Y., and P. Jorion (2006)
TITLE	Firm Value and Hedging: Evidence from US Oil and Gas Producers
DATA	Hedging activities of 119 U.S. oil and gas producers from 1998 to 2001
METHOD	Measurement of the Q ratio similar to Tobin's Q
CONCLUSION	Hedging has no value effect for a sample of oil and gas firms
NAME	Change Dan, Hong Gu and Kuan Xu (2005)
TITLE	The Impact of Hedging on Stock Return and Firm Value: New Evidence from Canadian Oil and Gas Companies
DATA	Oil and gas companies with market value more than Cdn $500 million in 2004, which 33 companies have been selected with the System for Electronic Document Analysis and Retrieval (SEDAR) during the period of 2000-2002
METHOD	Generalized additive model
CONCLUSION	Factors that affect stock returns and firm value are nonlinear, oil and gas companies will hedge when the unfavorable price occurs. And hedging has significant impact in firm value

NAME	Barbara Pirchegger (2005)
TITLE	Hedge Accounting Incentives for Cash Flow Hedges of Forecasted Transactions
DATA	No data
METHOD	Two period LEN-type agency model with a risk averse agent and a risk neutral principal
CONCLUSION	Whether hedge or not is depending on how the firm's overall risk exposure is allocated over periods. If risk exposures different over periods the firm prefers hedge accounting

NAME	Carter, David A., Daniel A. Rogers, and Betty J. Simkins (2006)
TITLE	Does fuel hedging affect firm value? Evidence from the U.S. airline industry
DATA	Jet fuel hedging behavior of firms in the US airline industry during 1992-2003
METHOD	First, estimate a monthly market model using an equally-weighted airline industry return. Second, using a three standard deviation price change to illustrate the effects of an extreme move in underlying asset prices.
CONCLUSION	Jet fuel hedging is positively related to airline firm value and the principal benefit of jet fuel hedging by airlines comes from reduction of underinvestment costs.

NAME	Bartram, Sohnke, Gregory Brown and Frank Fehle (2006)
TITLE	International evidence on financial derivatives usage
DATA	Sample of 7,309 non-financial firms from 48 countries
METHOD	For the simultaneous effects of the different factors on the likelihood of derivatives use, they use the LOGIT model. To test the relation between derivatives use and country-specific factors, they use by-country regression and LOGIT model.
CONCLUSION	Interest rate hedging, has a positive impact on firm value
NAME	Aziz A. Lookman (2011)
TITLE	Does Hedging Increase Firm Value? Comparing Premia for Hedging 'Big' versus 'Small' Risks
DATA	Unique hand-collected data set of firms in the oil and gas E&P sector
METHOD	Use different types of tests to analyse the data.
CONCLUSION	Hedging a 'big' risk is associated with lower firm value whereas hedging a 'small' risk is associated with higher firm value.
NAME	Viral Acharya et al. (2012)
TITLE	Online Appendix to Limits to Arbitrage and Hedging: Evidence from Commodity Markets
DATA	The EDGAR database which contains quarterly or annual statements for 94 firms
METHOD	General equilibrium model where the managerial costs of default are the motivation for firm hedging.
CONCLUSION	Higher default risk decrease the futures risk premium when supply disruption benefit the long side of the futures contract.

Table A.1: Table of Literatures

Bibliography

[1] Acharya, Viral et al. (2012), "Online Appendix to Limits to Arbitrage and Hedging: Evidence from Commodity Markets", May 18, 2012.

[2] Allayannis, George and James P. Weston (2001), "The Use of Foreign Currency Derivatives and Firm Market Value." *Review of Financial Studies 14*, 243-276.

[3] Bartram, Sohnke, Gregory Brown and Frank Fehle (2006), "International Evidence on Financial Derivatives Usage." *forthcoming in Financial Management*.

[4] Biernacki, C., G. Celeux, and G. Govaert (2000), "Assessing a Mixture Model for Clustering with the Integrated Completed Likelihood", *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22, 719-725.

[5] Carter, David A., Daniel A. Rogers, and Betty J. Simkins (2006), "Does Fuel Hedging Affect Firm Value? Evidence from the U.S. Airline Industry." *Financial Management 35*, 53-86.

[6] Chmelarova Viera (2007), "The Hausman Test, and Some Alternatives, with Heteroskedastic Data", Dissertation.

[7] Dan, Gu and Xu (2005), "The Impact of Hedging on Stock Return and Firm Value: New Evidence from Canadian Oil and Gas Companies." *Dalhousie, Department of Economics in its series Department of Economics at Dalhousie University working papers archive with number hedging.*

[8] Durbin, J. (1954), "Errors in Variables", *Review of the International Statistical Institute*, 22, 23-32.

[9] Fraley, C. and A. E. Raftery (1998), "How Many Clusters? Answers via Model-based Cluster Analysis", *The Computer Journal 41*, 578-588.

[10] Froot, Kenneth A., David S. Scharfstein, and Jeremy C. Stein (1993), "Risk Management: Coordinating Corporate Investment and Financing Policies." *Journal of Finance 48*, 1629-1658.

[11] Gagnon, Louis, Gregory J. Lypny, and Thomas H. McCurdy (1998), "Hedging Foreign Currency Portfolio," *Journal of Empirical Finance 5*, 197-220.

[12] Gaston, Noel and Daniel Trefler (1994), "Protection, Trade, and Wages: Evidence from U.S. Manufacturing", *Industrial and Labor Relations Review*, 47(4), 574-593.

[13] Graham, John R. and C.W. Smith (1999), "Tax Incentives to Hedge." *Journal of Finance 54*, 2241-2262.

[14] Graham, John R. and Daniel A. Rogers (2002), "Do Firms Hedge in Response to Tax Incentives?" *Journal of Finance 57*, 815-839.

[15] Greene, W. H. (2003), "Econometric Analysis" 5th edition, New Jersey, Prentice Hall.

[16] Halili, Abbas K. and Jiahua Chen (2007), "Variable Selection in Finite Mixture of Regression Models" *Journal of the American Statistical Association*, Volume 102, Issue 479, 2007.

[17] Hausman, J. (1978), "Specification Tests in Econometrics", *Econometrica*, 46(6), 1251-1271.

[18] Hausman, J. and W. Taylor (1981), "A Generalized Specification Test", *Economics Letters*, 8, 239-45.

[19] Heckman, J. (2008), "Econometric Causality" *National Bureau of Economic Research working paper*, #13934.

[20] Imbens, G. and J. Angrist (1994), "Identification and Estimation of Local Average Treatment Effects" Econometrica 62, 467476.

[21] Jin, Y., and P. Jorion (2006), "Firm Value and Hedging: Evidence from US Oil and Gas Producers.", *Journal of Finance (Forthcoming)*.

[22] Leisch, Friedrich (2003), "FlexMix: A General Framework for Finite Mixture Models and Latent Class Regression in R" *Report Series*, Report No. 86, December 2003.

[23] Lookman, Aziz A.(2011), "Does Hedging Increase Firm Value? Comparing Premia for Hedging 'Big' versus 'Small' Risks", *Social Science Research Network*, No. 5174.

[24] Maddala and Jeong (1992), "On the Exact Small Sample Distribution of the Instrumental Variable Estimator", *The Econometric Society*, Vol. 60, No. 1 (Jan., 1992), pp. 181-183.

[25] Miguel, E., Satyanath, S. and Sergenti, E. (2004), "Economic Shocks and Civil Conflict: An Instrumental Variable Approach" *Journal of Political Economy*,Volume 112, 725753.

[26] Nakamura, Alice and Nakamura, Masao (1981), "On the Relationships Among Several Specification Error Tests Presented by Durbin, Wu, and Hausman", *Econometrica*, 49(6), 1583-1588.

[27] Nelson and Startz (1990), "Some Further Results on the Small Sample Properties of the Instrumental Variable Estimator", *Econometrica, 58*, 967-976.

[28] Pearl, J. (2000), "Causality: Models, Reasoning, and Inference" Cambridge University Press.

[29] Pirchegger, Barbara (2006), "Hedge Accounting Incentives for Cash Flow Hedges of Forecasted Transactions." *journal European Accounting Review*, Vol.15 (2006), 115-135.

[30] Rajgopal, Shivaram (1999), “Early Evidence on the Informativeness of the SEC’s Market Risk disclosures: The Case of Commodity Price Risk Exposure of Oil and Gas Producers.” *Accounting Review 74*, 251-280.

[31] Rubin, D.B. (1987, 2004), *Multiple Imputation for Non-response in Surveys*, New York: John Wiley. Reprinted as a Wiley Classic, 2004.

[32] S. van Buuren, J.P.L. Brand, C.G.M. Groothuis-Oudshoorn and D.B. Rubin (2005), “Fully Conditional Specification in Multivariate Imputation” *Journal of Statistical Computation and Simulation*, Volume 76, Number 12, 1049-1064(16).

[33] S. van Buuren and Groothuis-Oudshoorn (2010), “MICE: Multivariate Imputation by Chained Equations in R” *Journal of statistical software, Volume: in press*, 1-68.

[34] Schafer J.L. (1997), “Analysis of Incomplete Multivariate Data” Chapman & Hall, London.

[35] Smith, Clifford W. and Rene M. Stulz (1985), “Determinants of Firm’s Hedging Policies.” *Journal of Financial and Quantitative Analysis 20*, 391-405.

[36] Stock, J., J. Wright and M. Yogo (2002), “A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments” *Journal of the American Statistical Association*, 20(4), 51829.

[37] Torsten , Achim Zeileis, Giovanni Millo, and David Mitchell (2012), “Testing Linear Regression Models” *CRAN*, 2012-05-31.

[38] Wu, DeMin (1973), “Alternative Tests of Independence between Stochastic Regressors and Distrurbances”, *Econometrica*, 41 (4), 733750.

[39] Wu, DeMin (1974), “Alternative Tests of Independence between Stochastic Regressors and Disturbances: Finite Sample Results”, *Econometrica*, 42(3), 529-546.