

GARCH MODELS FOR FORECASTING VOLATILITIES OF THREE MAJOR  
STOCK INDEXES: USING BOTH FREQUENTIST AND BAYESIAN APPROACH

A THESIS

SUBMITTED TO GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE

MASTERS OF SCIENCE

BY

LI, YIHAN

UNDER THE SUPERVISION OF

DR. MUNNI BEGUM

BALL STATE UNIVERSITY

MUNCIE, INDIANA

MAY 2013

## **Acknowledgement**

First of all, my greatest gratitude goes to my supervisor, Dr. Munni Begum, for her excellent guidance, and an excellent atmosphere for doing this thesis paper. I also want to thank her for her patience for correcting my writing, and SAS code writing. I would like to thank Dr. Rahmatullah Imon for supporting me, and providing precious advice for this thesis. I would also like to thank Dr. Rathin Rathinasamy and Dr. Tung Liu for guiding my research for last several months, and helping me to extend my background in Finance and Economics. I also sincerely thank Dr. Dale Umbach for helping and supporting through all this time as a statistics graduate student.

I would like to thank Sudipta Saha, who as a good friend were always willing to help and give his best suggestions. Many thanks to Jenni Felts, Dallis Pike, Xinyue Yang, Kejun Xu ,and Weilian Dong, and others for supporting me through my masters' degree and thesis. My thesis would not have been possible without their support. I would like thank to my father Dr.Gangyan Li and my mother Mrs. Xiaoming Han from bottom of my heart. They were always supporting me and encouraging me with their best wishes.

# Table of Contents

Acknowledgement	i
Table of Contents	ii
List of Figures	iii
List of Tables	iv
<b>1 Introduction</b>	<b>1</b>
Introduction	1
1.1 Purpose of Study	2
1.2 Significance of Study	3
1.3 Definition of terms	4
1.4 Literature Review	5
<b>2 Data Analysis</b>	<b>9</b>
2.1 The Nikkei 225 Index	12
2.2 The Standard&Poor 500 Index	15
2.3 The DAX Index	17
2.4 Conclusion	22
<b>3 Methodology</b>	<b>23</b>
3.1 Introduction	23
3.2 The ARCH Model	24
3.3 The Traditional GARCH Model	26
3.4 Exponential GARCH Model	27
3.5 Integrated GARCH Model	29
3.6 Power GARCH Model	30
3.7 Threshold GARCH model	30
3.8 The Quadratic GARCH model	31
3.9 Normality Assumption for Conditional Time Series Data	32
3.10 Estimation Methods	33
3.10.1 Maximum Likelihood Estimation for the Normal Distribution	33
3.10.2 Maximum Likelihood Estimation for the Student's t Distribution	34
3.10.3 The Bayesian Estimation of GARCH model	35
3.11 Summary	36
<b>4 Empirical Results</b>	<b>37</b>
4.1 The Nikkei 225 Index	39
4.2 The Standard&Poor 500 Index	44
4.3 The DAX Index	48
4.4 The Bayesian Estimation	52
<b>5 Conclusion</b>	<b>54</b>
<b>References</b>	<b>56</b>

## **List of Figures**

Figure 1: Plot of daily closing prices for Nikkei 225 stock index (1992 to 2012)	12
Figure2: Plot of daily returns for Nikkei 225 stock index (1992 to 2012)	14
Figure 3: Plot of the daily closing price for S&P 500 Index (1992 to 2012)	15
Figure 4: Plot of the daily returns for S&P 500 Index (1992 to 2012)	17
Figure 5: Plot of daily prices for DAX index stock index (1992 to 2012)	18
Figure 6: Plot of daily returns for DAX index (1992 to 2012)	19

## List of Tables

### Chapter 2

Table 1: Descriptive Statistics & ADF Tests	21
---	----

### Chapter 4

#### 4.1 The Nikkei 225 Index

Table 1: Goodness-of-fit Test	41
-------------------------------	----

Table 2: Forms of GARCH models for Volatility	43
---	----

#### 4.2 The Standard & Poor Index

Table 3: Goodness-of-fit Test	45
-------------------------------	----

Table 4: Forms of GARCH models for Volatility	47
---	----

#### 4.3 The DAX Index

Table 5: Goodness-of-fit Test	49
-------------------------------	----

Table 6: Forms of GARCH models for Volatility	51
---	----

#### 4.4 The Bayesian Estimation

Table 7 Parameter estimation based on the Bayesian Approach, compared with MLE	53
--	----

## **Chapter 1: Introduction**

In the last two decades, several countries have experienced substantial global economic crises. In Japan, there was the Economic Bubble that unfolded during the years of 1986 to 1991. Noguchi (1994) used the terms “extraordinary period” to describe this special time in Japanese economic history, because before December of 1986, the Japanese economy had a spectacular rise. After December of 1986, there was a great depreciation of the Japanese currency (Yen) that brought on the recession (Noguchi, 1994). In the USA, the National Bureau of Economic Research (NBER) officially announced that the country had been in an economic recession since December of 2007, which resulted from the decline in GDP growth (Borbely 2009). The economy of Germany practically stagnated in the beginning of 2000s. These three countries are developed nations and are the leading countries in their respective regions. Japan has the third largest economy in the world, and it is one of the leading countries in Asia. The United States has not only the largest economy in the North America, but also has the largest economy in the whole world. Germany leads the European economy.

The economic crisis seriously affected the global financial stock markets and made it difficult for investors to make investment decisions. Finding a way to measure the risk of the stock markets return came into sharp focus among the investors worldwide.

Volatility or standard deviation, also described as the total risk, is one of the widely used measures of risk used in financial time series modeling. Nevertheless, forecasting the volatility of the financial stock markets is needed minimizing the risk of making ineffective investment decisions.

### 1.1 Purpose of Study

The time series models are used to address the behavior of volatility. While the traditional time series models operate under the assumption that the variance does not change over time, the conditional variance does not depend upon the previous information. However, under in real world scenario, the variance does change over time, and very much depends on the past information. In order to handle this situation, Engle (1982) first introduced the Autoregressive Conditional Heteroscedastic (ARCH) model, which asserts that the conditional variance does change over time. Bollerslve (1986) expanded the ARCH model to Generalized ARCH (GARCH) model, which allowed more flexible lag structure.

McDonald (2006) describes the GARCH model as an important and widely used tool to forecast volatility model that attempts to statistically capture the ebb and flow of volatility. A few scholars have developed variant forms of the GARCH model. For example, Nelson (1991) proposed the exponential GARCH (EGARCH) models, and it overcomes some of the limitation of the traditional GARCH model. Engle and Bollerslve (1986) first introduced the Integrated GARCH (IGARCH) model that allows the current information to have significant impact on forecasting volatility. Ding, Granger, and Engle (1993) first mentioned the Power GARCH (PGARCH) model that it provides an alternative way to model the volatility with the long memory property. Glosten,

Jaganathan and Runkle (1993) and Zakoian (1994) developed the Threshold GARCH (TGARCH) model. Engle and Ng (1993) first found the Quadratic non-parametric GARCH (QGARCH) model.

The primary purpose of this paper is to use the above forms of GARCH models to address and forecast volatility of the stock market returns of Japan, the United States, and Germany. These three countries are the leading countries in their regions, which are Asia, North America and Europe respectively. The data set for Japan is based on the daily closing stock price from the Nikkei 225 index, the United States data set is from the Standard & Poor 500 index, and the data set for Germany is drawn from the DAX index. In order to further analyze these three data sets, both the GARCH models with normal distribution and the GARCH models with the student's t distribution will be considered in this paper. Moreover, The Maximum Likelihood Estimation will be used for all the forms of the GARCH models, and the Bayesian Estimation will be used to fit the standard GARCH model with student's t distribution. The other purpose of this paper is to compare the results and choose the precise method to forecast volatility for these three stock market return series.

In this study, firstly we need to figure out the precise form of the GARCH model to forecast volatility for these three stock indexes. We will also be using parameter estimation result to interpret properties for these three stock indexes.



## 1.2 Significance of Study

Results of this study will provide a better understanding of addressing and forecasting volatility used by the family of GARCH models. Even though there is a great amount of research has been done using the GARCH model to study the behavior of volatility, the focus of most of the researches has been usually on one country, method or one estimation method each of the time. It is necessary to study more alternative models, and more international evidence on the volatility forecasting.

This paper investigates the three stock markets from three different countries, and six forms of the GARCH models with alternative distributions used for forecasting volatility. Specifically, we carry out the analysis with alternative distributions, which are the normal distribution and the student's t distribution. Moreover, we are not only just considering the Maximum Likelihood Estimation to fit the three data sets into the forms of GARCH models, but also the Bayesian Estimation will be considered in this paper as well.

## 1.3 Definition of terms

**Volatility:** is a measurement for price variation of financial market over the time.

**Heteroscedastic:** the variance of the data will change within the time, and depends on the past information.

**Homoscedastic:** the variance of the data will remain the same, and does not depend on the past information.

**The Nikkei 225 index:** is a stock market index for the Tokyo Stock Exchange (TSE). It contains the 225 top-rated, blue-chip Japanese companies listed in the first section of the TSE, and it is a price weighted index, currency unit is Japanese Yen (CME Group, 2007).

**The S&P 500 index:** the full name is the Standard & Poor's 500, is one of the American stock market indexes based on the common stock prices of 500 top publicly traded American companies, ranked by market value. It represents 83% of the market capitalization of all regularly traded stocks in the United States (Siegel, Schwartz.2006).

**The DAX index:** the full name of the DAX is the Deutscher Aktien index. It is a blue-chip stock market, which is consisting of the 30 major German companies trading on the Frankfurt Stock Exchange. It represents nearly 80% of the market capital authorized in Germany (Clark, Mansfield, and Tickell, 2001).

#### 1.4 Literature Review

The primary purpose of this study is to use both the family of the GARCH models with the normal distribution and the student's t distribution to forecast volatility in financial data from three different regions of world and compare the results to observe the differences.

This review emphasizes the volatility, the significance of forecasting volatility. This chapter provides the general background information about volatility, and how to use GARCH models to forecast the volatility.

Volatility in the financial market received massive attentions during the last two decades and it is the measurement for the variation of stock price of the financial market all over the time and it can be treated as the measurement of risk. A great volatility could help stock, bond and foreign exchange markets to make important public policy issue. French, Schwert and Stambauch(1987) examines the realitonship between stock returns and volatility, and successfully find the evidence of expected risk premium is positively related with volatility. Kim, Morley, and Nelson (2004) found the positive relationship

between stock market volatility and the equity premium. However, volatility is hard to directly get. Busse (1999) showed that volatility timing is an important factor in the returns of mutual funds and has led to higher risk-adjusted returns. Brandt and Jones (2006) mentioned that the volatility of financial stock market is time-varying and predictable, but forecasting the future level of volatility is complicated, because correct estimators of the parameters of a volatility model can be difficult to find.

Forecasting volatility is one of the most important tasks in financial markets, since volatility is the measurement of the variation among the prices. The precise forecasting of volatility will become the key input in order to help investors to make correct investment decisions, or to create effective financial portfolios. Moreover a precise volatility forecasting results can help investors to have a clear mind about the certain level of risk, which involves in their investment decision.

Some scholars used Copula approach to address and forecast volatility .Ning, Xu, and Wirjanto (2010) noted that the advantage of copula method is in estimating volatility with an empirical distribution function which is distribution-free. Wilson and Ghahramani (2010) developed a stochastic volatility model, namely the Gaussian Copula Process Volatility (GCPV), and used it to predict volatility. They observed that GCPV can produce more robust results than GRACH models. Sokolinskiy and Dijk (2011) developed a novel approach to modeling and forecasting realized volatility measure based on copula functions. They found that Gumbel Copula-based RV model achieved the best forecasting volatility performance for S&P 500.

The popular sources of forecasting volatility are the time series models. Some conventional time series model are based on the assumption of the variance of error terms

of expected values are not depends on the past values. However, under the real circumstances, the variance of error terms actually depends on the past information. In order to capture the more accurate results, Engle (1982) proposed the ARCH models which states the current variance of error terms depends on the previous values, and he successfully used the ARCH to estimate the variance of the United Kingdom inflation. Bollerslev(1986) extended the ARCH model as the Generalized ARCH(GARCH) model, and bears much resemblance to the extension of the standard time series autoregressive (AR) process to the general autoregressive moving average (ARMA) process.

Since the invention of the GARCH model, it has been the most popular way for researchers, analysts, and investors to forecast volatility. Eagle (2001) pointed out GARCH (1, 1) model is the simplest and most robust of the family of volatility models. Floros (2008) found strong evidence that the volatility of daily returns can be characterized by the GARCH models from Egypt and Israel stock markets. Akgiray (1989), Pagan and Schwert (1989) and Brooks (1996) showed that the GARCH models are fitted to the US stock data satisfactorily.

There is several extension forms of the GARCH model have been proposed as well. Nelson (1991) proposed the exponential GARCH (EGARCH) model. Brandt and Jones (2006) used the EGARCH models to forecast the volatility of S&P 500. Engle and Bollerslev (1986) proposed the Integrated GARCH (IGARCH) model. Ding, Ganger, and Engle (1993) first mentioned the Power GARCH (PGARCH) model. Lucey and Tully (2006) used PGARCH models to determine the volatility of global gold prices. Glosten, Jaganathan and Runkle (1993) and Zakoian (1994) and developed the Threshold GARHC (TGARCH) model. Chiang (2001) successfully uses the TGARCH model to test the

relationship between stock returns and volatility for seven Asian stock markets. Lastly, Engle and Ng (1993) first proposed the Quadratic GARCH (QGARCH) model.

Details of this paper are organized as follows. Chapter 2 briefly introduces the primary statistics analysis for the three data sets. Chapter 3 presents the methodology of the forms of the GARCH models we considered in this paper. Chapter 4 is the empirical results. Finally, Chapter 5 is the conclusion of this thesis paper.

## Chapter 2: Data Analysis

We consider three data sets in this research. The first data set is based on the daily closing stock price from the Nikkei 225. The second data set is built from daily closing stock prices from the S&P 500 and the third data set is based on the daily closing stock prices from the DAX stock index. The daily closing stock prices were obtained from the YAHOO! FINANCE historical price database. The following logarithmic price change formula is used to calculate the daily returns:

$$(1) \quad R_t = \ln\left(\frac{p_t}{p_{t-1}}\right)$$

where  $R_t$  represents the daily return,  $P_t$  is the daily closing stock price at time  $t$ , and  $P_{t-1}$  means the daily closing stock price at time  $t-1$ .

In order to explore these three datasets in details, we calculate a number of important statistics, such as Skewness, Kurtosis, Jarque-Bera and Augmented Dickey Fuller (ADF).

Skewness is a measurement of the lack of symmetry. Positive skewness of a variable under consideration implies that its distribution has a long right tail. On the other hand, negative skewness shows that the distribution of a variable under consideration has a long left tail.

The Kurtosis is a measurement of whether the data is peaked or not relative to a normal distribution, and it is a descriptor of the shape of a probability distribution. A higher positive kurtosis shows the distribution has higher, acute peaks around the mean and has fatter tails. On the other hand, a distribution with negative kurtosis shows it has lower, wider peaks around the mean and has thinner tails.

Given the large number of observations of these three data sets, the Jarque-Bera test will be an alternative way to test normality. It is a type of Lagrange multiplier test that was developed to test normality, heteroscedasticity, and serial correlation or autocorrelation of regression residuals (Jarque and Bera 1980). The Jarque-Bera statistics is calculated from skewness and kurtosis, and they follow chi-squared distribution with two degrees of freedom as below:

$$(2) \quad \text{The Jarque – Bera Statistic} = N * \left[ \frac{\text{skewness}^2}{6} + \frac{(\text{kurtosis}-3)^2}{24} \right] \sim \chi^2(2)$$

where N is the number of observations, and if the value for Jarque-Bera statistic is greater than the critical value, then the null hypothesis of the normality is rejected.

Moreover, the Augmented Dickey-Fuller (ADF) test is utilized to test stationarity for time series datasets. The hypothesis for a time series is:

$H_0$ : The series is non-stationary

$H_A$ : The series is stationary

The ADF statistic is compared with the critical values of t-distribution to draw conclusions about stationarity (Dickey and Fuller, 1981).

For the time series model, the series is expected to be stationary. If the original series is non-stationary, there are several ways to transform the non-stationary series into a stationary one. The most common method is called differencing, which consists of

taking the difference between consecutive observations. The differencing formula is represented as follows:

$$X_t = Y_{t+1} - Y_t.$$

$X_t$  is the new price after taking the difference procedure at time  $t$ ,  $Y_{t+1}$  is the original price at the time  $t+1$ , and  $Y_t$  is the original price at the time  $t$ . After using the differencing method, the new data set will have one less observation than the original data. If the new data set is still non-stationary, repeat the differencing procedure again until stationary is obtained. Another common approach is to consider natural log transformation on the price ratio which is presented as follows:

$$x_i = \ln\left(\frac{y_{i+1}}{y_i}\right)$$

where  $x_i$  is the return after using the log-transformation,  $y_{i+1}$  is the original daily price at time  $t+1$ , and  $y_i$  is the original daily price at time  $t$ . The new data set will also have one observation less than the original data set. Notice that the log-transformation method produces the same equation on log scale.

In this research, all data will operate under the financial time series model. Floros (2008) mentioned that fat tails and volatility clustering are the two important characteristics of the financial time series. Positive kurtosis results prove that the distribution of the data has the fat tails. Kirchler and Huber (2007) mentioned that volatility clustering manifests itself as periods of tranquility interrupted by periods of turbulence. So the visual inspection of the plots of daily returns will test the data to determine whether or not it has volatility clustering or not.



The chapter is organized as follows: Sections 2.1, 2.2, and 2.3, all of which provide the data information for the Nikkei 225, the S & P 500, and the DAX index, respectively. Section 2.4 presents the conclusion of this chapter.

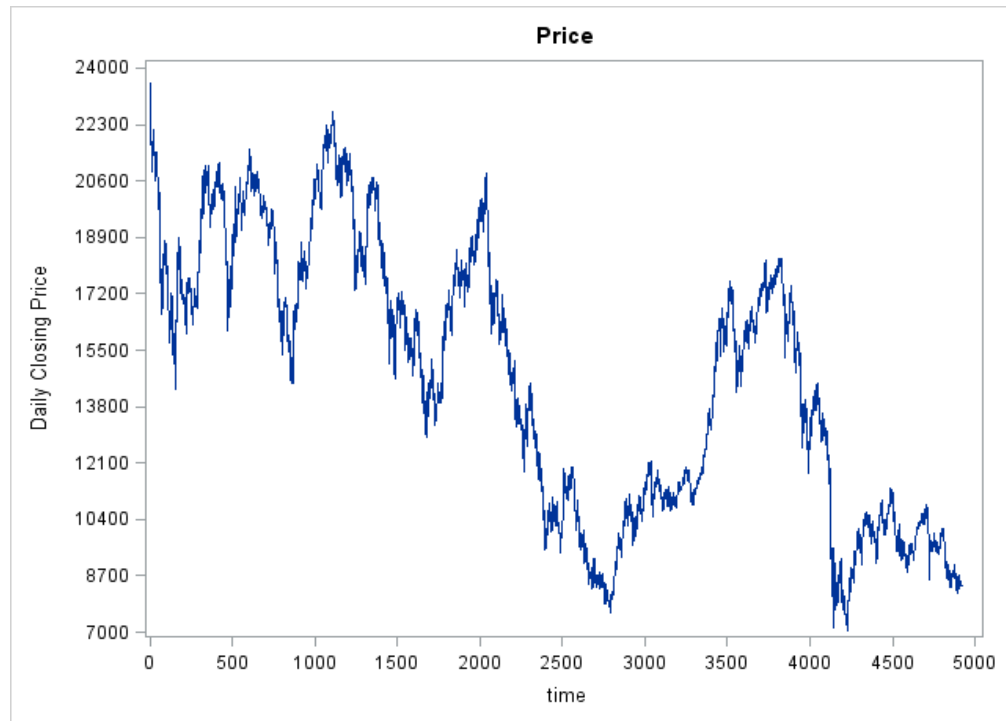
## 2.1 The Nikkei 225 Index

The data employed in this section is comprised of 4919 observations which were obtained from the Nikkei 225 index (Japanese Stock Market) covering the period 7<sup>th</sup> January 1992 to 9<sup>th</sup> January 2012. The ordinary data is the set of daily closing stock prices and the currency is in Japanese Yen.

The Nikkei 225 index is one of the oldest barometers of the Japanese Stock market. It uses the 225 top-rated, blue-chip Japanese companies listed in the First Section of the Tokyo Stock Exchange (CME Group, 2007).

**Figure1: Plot of daily closing prices for Nikkei 225 stock index (1992 to 2012)**

Currency in Japanese Yen



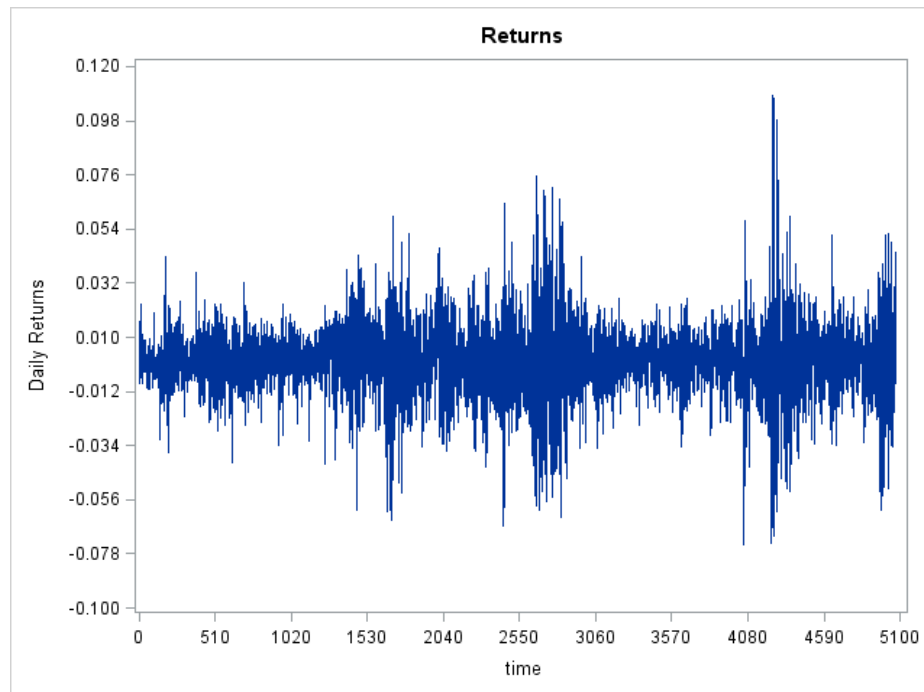
Some basic information is shown on the above graph. There are several peaks that are shown during the period. The data do not vary around the common place or mean, so the non-stationary property exists.

Moreover, Table 1 shows the basic summary statistics of the daily closing price for Nikkei 225 and the results from the ADF test. The mean is 14,631.05, and value for skewness is 0.00416746. The positive skewness shows that the distribution of the data set has long right tails. The Jarque-Bera statistics is 1166.770511, which is greater than the critical value. It rejects the null hypothesis of normality. The value of ADF statistics is -0.94, and proves that the time series is non-stationary, since it is smaller than the critical value at 1%, 5% and 10 % t-distribution at 2 degrees of freedom.

Since the data for the daily closing price is non-stationary, the log-transformation method captures the new data of daily returns. The plot for the daily returns is shown on Figure 2.

**Figure2: Plot of daily returns for Nikkei 225 stock index (1992 to 2012)**

Currency in Japanese Yen



The plot of the daily returns shows that data has the stationary pattern, since all the data is moving around a common location or mean. The sign for volatility clustering happened in the above figure as well.

The summary statistics for the Nikkei225 daily returns is shown on Table 1. The value of the mean is  $-0.00021$ . The skewness is  $-0.1211103$ , which shows the distribution of the data has long left tails. The kurtosis is  $5.3584191$ , which is a highly positive value and proves the distribution has fatter tails. The Jarque-Bera statistics is  $1166.770511$ , which is much greater than the critical value of 1%, 5%, and 10% of the Chi-square distribution with two degrees of freedom. The null hypothesis of normality is rejected.

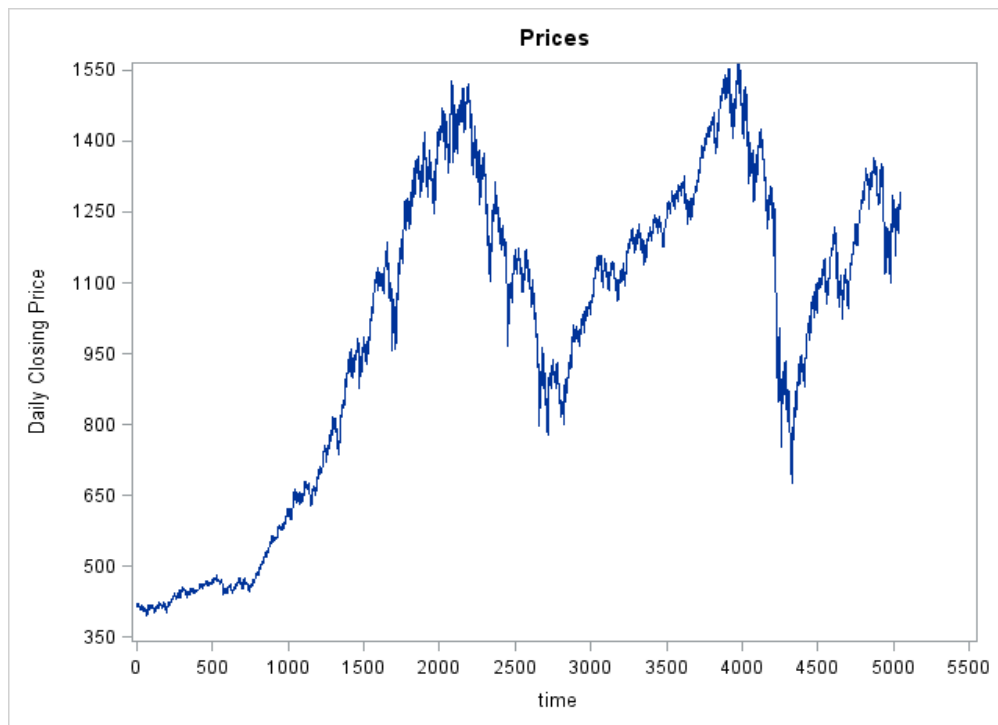
## 2.2 The S& P 500 Index

The data set is built from daily closing prices from the S&P 500 index, and it has 5046 observations from 3<sup>rd</sup> January 1992 to 10<sup>th</sup> January 2012.

The S&P Index is the most widely used benchmark for measuring the performance of the large-capitalization U.S. basic stocks. It covers the top 500 largest companies in the United States, ranked by market value. It represents 83% of the market capitalization of all regularly traded stocks in America (Siegel, Schwartz.2006).

**Figure 3: Plot of the daily closing price for S&P 500 Index (1992 to 2012)**

Currency in U.S Dollar



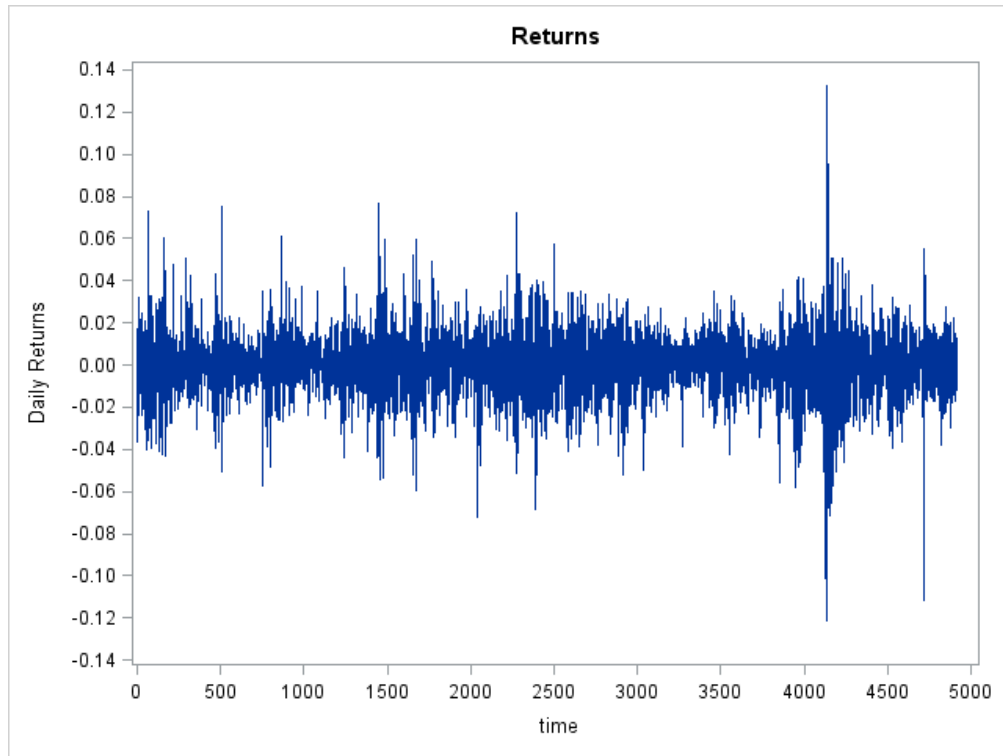
The plot of daily closing prices shown on Figure 3 above visually represents that the S&P 500 daily closing prices are non-stationary. This is evident because the prices do not move around the same location, and there are three peaks during the time period from January 1992 to January 2012.

Table 1 shows the summary statistics for the S&P 500 daily closing price. The value of the mean is 1005.08249. The distribution of this data set has long left tails, since the value for skewness is -0.4770196. The kurtosis is -1.00768823 which implies that the distribution of this data set has wider peaks around the mean. The Jarque-Bera Statistic is 3,568.301652, and it is greater than the critical value of 1%, 5% and 10% for the Chi-Square distribution with two degrees of freedom. The result of the Jarque-Bera test shows that it accepts the null hypothesis of normality. The plot of daily closing prices shows that it is non-stationary. The ADF test result from Table 1 is 1.96, which is smaller than the critical values of 1%, 5% and 10 % of the T-distribution, which means that failed to reject the null hypothesis.

After using the log-return method to capture the daily returns for the S & P 500, the following Figure 4 shows the plot of the daily returns. It obviously showed that the daily returns have a stationary pattern, since the data is fluctuating around a common location. Also, the volatility clustering phenomenon happened based on the plot of the daily returns, since some of daily returns are high for extended periods and then low for extended periods.

**Figure 4: Plot of the daily returns for S&P 500 Index (1992 to 2012)**

Currency in U.S Dollar



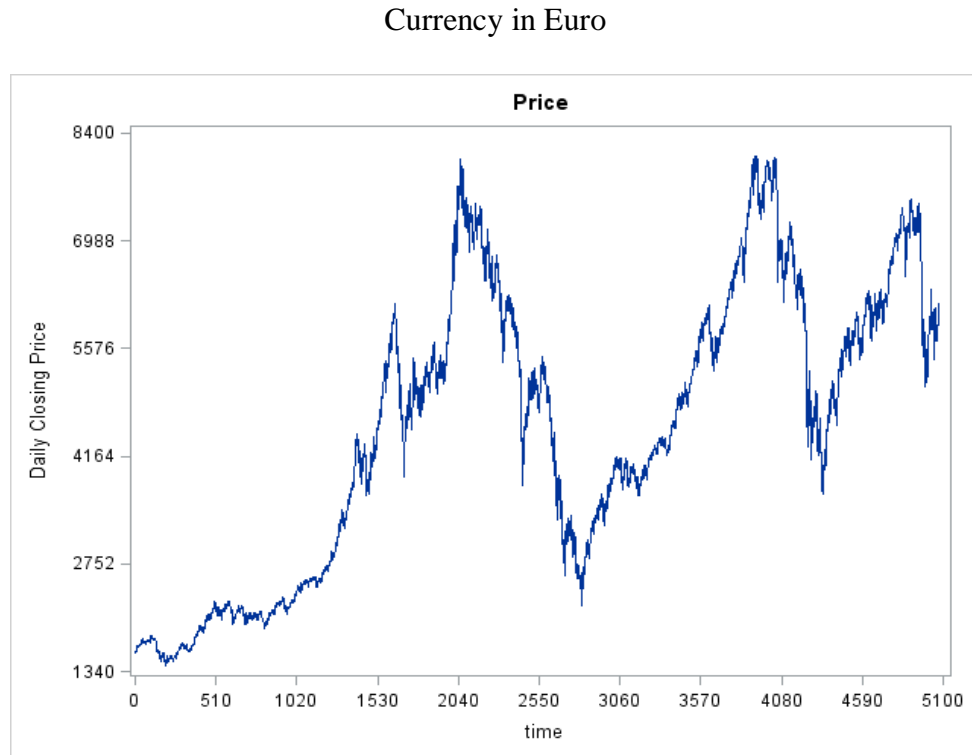
The summary statistics for the daily returns are shown on Table 1; the value of mean is 0.000224, and skewness is  $-0.02369762$ . The skewness exists and the distribution of daily returns has a long left tail. The kurtosis is 6485.383663, and it proved that the distribution of the data has fatter tails. The Jarque-Bera statistics is 6485.383663, which is larger enough to reject the null hypothesis of normality.

### 2.3 DAX index.

The data set was based on the DAX index (Germany Stock Market); it included 5,072 observations from 3<sup>rd</sup> January 1992 to 9<sup>th</sup> January 2012. The data is the daily stock closing price of the DAX index.

The DAX index is a Germany blue chip stock market index which is the measurement for the development of the 30 largest companies on the German equities market and it represents nearly 80% of the market capital authorized in Germany (Clark, Mansfield, and Tickell, 2001).

**Figure 5: Plot of daily prices for DAX index stock index (1992 to 2012)**



The daily closing price plot (Figure 5) visually shows some basic information about this DAX data set. There are three peaks during this period. The inspection of the plot graph showed the closing stock prices are non-stationary, since all the prices do not vary with a common mean or a location.

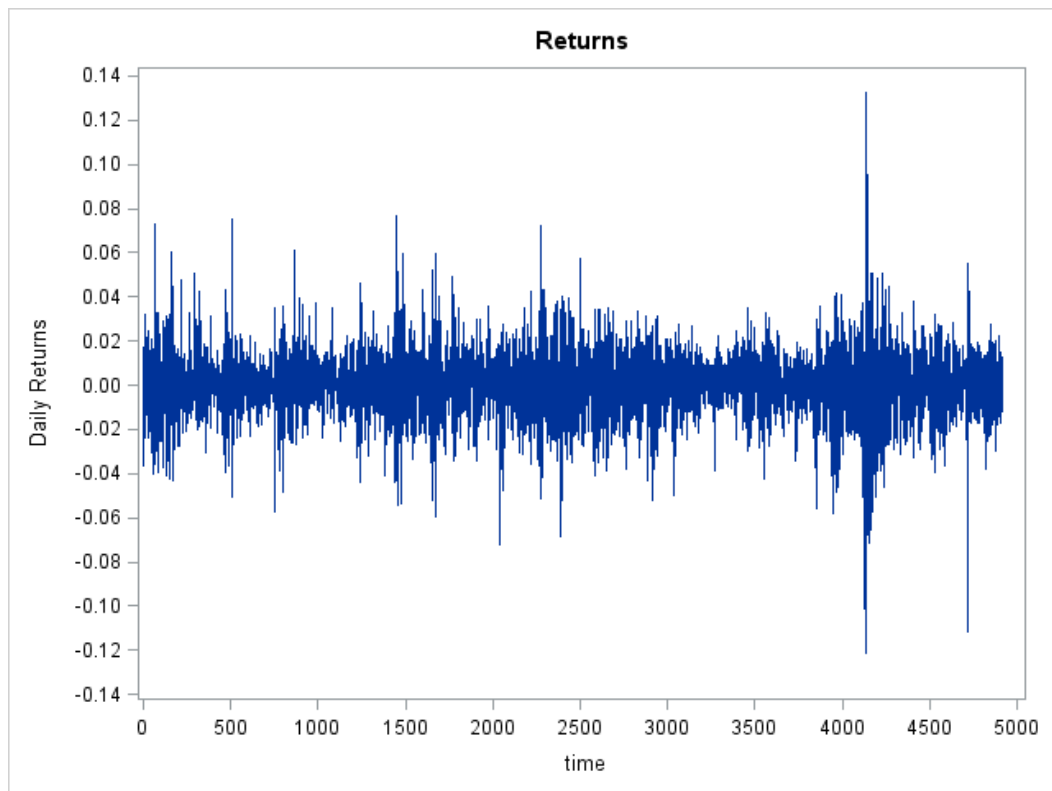
From the Table 1, the value of mean is 4,534.027, and skewness is -0.048021; the skewed exists. The ADF statistics are smaller than the critical values for 1%, 5% and 10% t-distribution, thus we accept the null hypothesis of stationarity. However the

Jarque-Bera statistics is 2099.454, which implies the rejection of the null hypothesis of normality.

After using log-transformation method to get the data for daily returns, the plot of daily returns for the DAX data set is presented on the Figure 6. It is clearly seen that the data for daily returns are stationary because the log-transformed returns from the DAX data set fluctuate around a common mean or location. Moreover, it clearly shows the volatility clustering phenomenon as the daily returns are not independent from one to other.

**Figure 6: Plot of daily returns for DAX index (1992 to 2012)**

Currency in Euro





The table 1 show, in regards to the daily returns, the value of the mean is 0.00261; skewness is -0.00.0607152. It means the distribution for daily returns has a long right tail. The value for kurtosis is 4.29124248, which is positive and greater than 3. It proved the distribution for returns has the fatter tails. The Jarque-Bera statistic is 355.4737583, which is greater than the critical value for 1%, 5% and 10% of chi-square distribution with two degrees of freedom. Thus it rejects the null hypothesis of normal distribution.

#### 2.4 Conclusion

For these three datasets, plots for daily closing price shown in figures 1, 3 and 5 do not show stationarity but all the plots of the daily returns (Figure 2, 4 and 6) appear to be stationary since the log transformed series from all these three sets fluctuate around a common mean or location.

We present the summary statistics for these three datasets along with Skewness, Kurtosis, Jarque-Bera and Augmented Dickey Fuller (ADF) in Table 1.

**Table1: Descriptive Statistics & ADF Tests**

<b>A. Return</b>	<b>Nikkei 225</b>	<b>S&amp;P 500</b>	<b>DAX</b>
Mean	-0.00021	0.000224	0.00261
Median	0	0.00058	0.000775
Std. Dv	0.01548	0.01207	0.01487
Maximum	0.1323458	0.109572	0.107975
Minimum	-0.1211103	-0.0946951	-0.07433
Skewness	-0.1804741	-0.2369762	-0.06072
Kurtosis	5.35848191	8.53366307	4.291242
Jarque-Bera	1166.770511	6485.383663	355.4738
Observations	4919	5046	5072
<b>B. Price</b>	<b>Nikkei 225</b>	<b>S&amp;P500</b>	<b>DAX</b>
Mean	14,631.05	1005.08249	4534.027
Median	15207.86	1099.805	4704.97
Std. Dv	4059	334.77716	1842.771
Maximum	23566	1565.15	8105.49
Minimum	7054.98	394.5	1420.3
Skewness	0.00416746	-0.4770196	-0.04801
Kurtosis	-1.2813125	-1.0076823	-0.15041
Jarque-Bera	3756.826032	3568.301652	2099.454
ADF (Level)	-0.94	1.95	0.88
ADF (1st diff.)	-2.76	-1.68	0.46
Observations	4919	5046	5072

From the above Table, the ADF statistics for the daily closing price for three data sets were smaller than the critical values for 1%, 5% and 10% T-Distribution. This implies that the time series from all three stock exchanges fail to satisfy stationarity.

Since all the daily closing prices from three stock exchanges are non-stationary, we use log-transformation to calculate the daily returns. These daily returns from these three data sets have negative skewness which implies that all the daily return distributions have long left tails. The values of high positive kurtosis imply that the distribution for these three data sets have higher peaks around mean compared with the

normal distribution. The Jarque-Bera statistics showed the Jarque-Bera test rejects the null hypothesis of the normality at the 5% level for Chi-Square Distribution with two degrees of freedom.

Furthermore, all the plots of the daily returns for these three data sets showed the volatility clustering occurs (showed in Figure 2, 4 and 6). In order address such volatility clustering we consider GARCH models to analyze these daily returns from three stock exchange market. A variety of GARCH models are presented in next chapter.

## Chapter 3: Methodology

### 3.1 Introduction

The volatility is the measurement of variation among the prices of financial time series data. Accurate forecasting volatility is an important tool for investors to make the right investment decisions and also helps researchers to better understand the change in the financial market. Time series models are the main methodologies for forecasting volatility in financial data. Some conventional time series models are based on the assumption of homoscedasticity, which means the variance of error terms of expected values remain the same at any given time. However, under the real circumstances, the variance of error terms actually varies all the time, which implies that heteroscedasticity, exists in the data. In order to capture more accurate forecasting results, Robert F Engle (1982) proposed the Autoregressive Conditional Heteroscedasticity (ARCH) model which states that the variance in the data at time  $t$  depends on the previous time  $t-1$ . Tim Bollerslev (1986) generalized the ARCH model and named the model as the Generalized ARCH (GARCH) model which allows for a more flexible lag structure and it bears much resemblance to the extension of the standard times series autoregressive (AR) process to the general autoregressive moving average (ARMA) process.

During the last two decades, the ARCH and GARCH models have been the most popular methods for the researchers, analysts, and investors to forecast volatility. Moreover, some scholars have also developed variant forms of the GARCH model. For example, Nelson (1991) proposed the exponential GARCH (EGARCH) model, and Engle and Bollerslev (1986) proposed the Integrated GARCH (IGARCH) model. Ding, Ganger, and Engle (1993) first mentioned the Power GARCH (PGARCH) model, while Glosten, Jaganathan and Runkle (1993) and Zakoian (1994) and developed the Threshold GARCH (TGARCH) model. Lastly, Engle and Ng (1993) first proposed the Quadratic GARCH (QGARCH) model.

Estimation of the parameters in the above models is mostly based on the likelihood approach. Ardia (2007) used the Bayesian approach to address parameter estimation in the GARCH model.

In this paper, we consider the traditional GARCH model, as well as other extensions of the GARCH model. Both the likelihood and the Bayesian approaches for model fitting are considered. In what follows, we briefly present these models and discuss the estimation methods using both likelihood and Bayesian approaches.

### 3.2 The ARCH model.

The ARCH (q) model was proposed by Robert F Engle (1982). It addresses the heteroskedasticity in the data over time and models this explicitly in terms of defining a model for conditional variance. The conditional density function  $f(y_t | y_{t-1})$  represents the prediction of current random variable  $y_t$  which is based upon the past information. The conditional variance denoted by  $V(y_t | y_{t-1})$  depends on the past information and

changes within the time. Engle's (1982) ARCH model allows explicitly address the conditional variance, which depends on the past information, and can be written as

$$(1) \quad y_t = \varepsilon_t = \varepsilon_t \sqrt{h_t}$$

$$h_t = \omega + \alpha_1 \varepsilon_{t-1}^2$$

Where  $\varepsilon_t$ s are independently and identically distributed as standard normal  $N(0,1)$  and are not dependent on the  $y_{t-1}, y_{t-2} \dots y_{t-i}$ .  $h_t$  is the conditional variance which depends on the past time  $t-1$ . If we add the assumption of normality with zero mean, the ARCH model can be expressed as the following:

$$(2) \quad y_t \mid \psi_{t-1} \sim N(0, h_t)$$

$$h_t = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2$$

Where  $y_t$  is the time series data at time  $t$ ,  $q$  is the order of ARCH process and  $\omega$  and  $\alpha_i$  are unknown parameters.

In order to address the effects of the exogenous variables  $x_t$  the ARCH regression model can be written as follows

$$(3) \quad y_t = x_t' b + \varepsilon_t$$

where  $x_t$  is a vector of explanatory variables, and  $b$  is a vector of unknown regression parameters.

Engle (1982) assumes that the mean of conditional density function of  $y_t$  is  $x_t' b$ , a linear combination of lagged endogenous and exogenous variables, which included past information set  $\psi_{t-1}$  with a vector of unknown parameters  $b$ . It can be formally expressed as

$$(4) \quad y_t \mid \psi_{t-1} \sim N( x_t' b, h_t )$$

where  $\psi_t$  is all the past available information of time  $t-1$ ,  $y_t$  is the time series data at time  $t$ , and the conditional variance  $h_t$  can be expressed as the following:

$$(5) \quad h_t = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2$$

$$\varepsilon_t = y_t - x_t' b$$

Where  $q$  is the order of the moving average ARCH terms and  $\omega$  and  $\alpha_i$  are unknown parameters.

### 3.3 Traditional GARCH Model

Bollerslev (1986) introduced a new general class of ARCH models, named generalized autoregressive conditional heteroscedastic (GARCH) models, which allows for both a long memory and a more flexible lag structure. ARCH models concern with the conditional variance which is linearly associated with the past variances only. The GARCH models added the previous conditional variances into the formulation as well. The GARCH model with the assumption of normality with zero mean can be expressed as

$$(6) \quad y_t = \varepsilon_t = \varepsilon_t \sqrt{h_t}$$

$$y_t | \psi_{t-1} \sim N(0, h_t)$$

where  $\psi_{t-1}$  denotes the past information through time  $t-1$ , and the conditional variance  $h_t$  can be written as

$$(7) \quad h_t = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j}$$

where  $p \geq 0; q > 0$ ;

$$\omega > 0, \alpha_i > 0, i = 1, 2, 3 \dots q$$

$$\beta_j \geq 0, j = 1, 2, 3 \dots p$$

where  $q$  is the order of autoregressive ARCH terms,  $p$  is the order of moving average terms of GARCH model, and  $\alpha$  and  $\beta$  are the vectors of unknown parameters.

If we add exogenous variables  $x_t$  and assume normality with non-zero mean for the variable  $y_t$ ,  $y_t$  can be expressed as,

$$(8) \quad y_t = x_t' b + \varepsilon_t$$

where  $x_t$  is a vector of explanatory variables, and  $b$  is a vector of unknown parameters.

So the GARCH regression model can be given by

$$(9) \quad y_t | \psi_{t-1} \sim N(x_t' b, h_t)$$

$$\varepsilon_t = y_t - x_t' b$$

$$h_t = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j}$$

where  $p \geq 0; q > 0$  ;

$$\omega > 0, \alpha_i > 0, i = 1, 2, 3 \dots q$$

$$\beta_j \geq 0, j = 1, 2, 3 \dots p$$

where  $\psi_t$  represents all the past available information at time  $t-1$ ,  $q$  is the order of moving average ARCH terms, and  $p$  is the order of autoregressive GARCH terms. Notice that, when  $p=0$ , the GARCH ( $p, q$ ) model will be reduced to the ARCH ( $q$ ) model.

### 3.4 Exponential GARCH Model

Nelson (1991) pointed out some limitations of the GARCH models. First, the GARCH model, by assumption, cannot handle the negative correlation between future values and current values. Second, the GARCH model may over restrict the dynamics of conditional variance by parameter restrictions. Third, the GARCH model makes it hard to interpret whether shocks of conditional variance continue or not.



If time series data  $y_t$  can be represented as

$$(10) \quad y_t = u + \varepsilon_t$$

where  $u$  is the expected value,  $\varepsilon_t$  is white noise with zero mean and it can be written as

$$(11) \quad \varepsilon_t = z_t \sqrt{h_t}$$

where  $z_t$ s are independently and identically distributed as standard normal  $N(0,1)$ .

Thus,

$$(12) \quad z_t = \frac{\varepsilon_t}{\sqrt{h_t}}$$

In order to address the above limitations, Nelson (1991) proposed an alternative form of the GARCH model named the Exponential Generalized Autoregressive Heteroscedastic (EGARCH) model. He considered the natural log of conditional variance as linear in some functions of time and past function  $Z_t$  to ensure the conditional variance of time series data  $y_t$  remains positive. So the natural log of conditional variance of EGARCH is given by

$$(13) \quad \ln(h_t) = \omega + \sum_{i=1}^q g(z_{t-i}) + \sum_{j=1}^p \beta_j \ln(h_{t-j})$$

where

$$(14) \quad g_i(z_t) = \alpha_i z_t + \gamma [ |z_t| - E|z_t| ]$$

and the term  $[ |z_t| - E|z_t| ]$  represents a magnitude effect of the GARCH model. If we combine equation (13) with equation (14), and consider  $E|z_t|$  as 0, the EGARCH model can be simplified as

$$(15) \quad \ln(h_t) = \omega + \sum_{i=1}^q \alpha_i \left| \frac{\varepsilon_{t-i}}{\sqrt{h_{t-i}}} \right| + \sum_{j=1}^p \beta_j \ln(h_{t-j}) + \gamma \left| \frac{\varepsilon_{t-i}}{\sqrt{h_{t-i}}} \right|$$

where

$$p \geq 0; q > 0;$$

$$\omega > 0, \alpha_i > 0, i = 1, 2, 3 \dots q$$

$$\beta_j \geq 0, j = 1, 2, 3 \dots p$$

where  $\varepsilon_t$  is white noise with zero mean,  $h_t$  is the conditional variance,  $q$  is the order of moving average ARCH terms, and  $p$  is the order of autoregressive GARCH terms. The  $\alpha$  and  $\beta$  are the unknown vector of parameters. The  $\gamma$  parameter is the measurement of the asymmetry effect. If  $\gamma = 0$ , then the impact to conditional variance is symmetry.

### 3.5 Integrated GARCH Model

Engle and Bollerslev (1986) proposed the Integrated GARCH (IGARCH) model. It adds an assumption, that the sum of the parameters  $\alpha_i$  and parameters  $\beta_j$  equals one. It has an important property of persistent variance which allows the current information to have significant impact of forecasting conditional variance for all horizons. Thus under the IGARCH model

$$(16) \quad \sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j = 1$$

and

$$(17) \quad \sum_{j=1}^p \beta_j = 1 - \sum_{i=1}^q \alpha_i.$$

The conditional variance  $h_t$  from the traditional GARCH model is considered as,

$$(9) \quad h_t = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j}$$

Then combining equations (16) and (9), the conditional variance for IGARCH (1, 1) model can be written as

$$(18) \quad h_t = \omega + \alpha_1 \varepsilon_{t-1}^2 + (1 - \alpha_1) h_{t-1}$$

where  $y_t = \varepsilon_t = \varepsilon_t \sqrt{h_t}$ ,  $q$  is the order of autoregressive ARCH term.

where  $0 < \alpha_1 < 1$ .

### 3.6 Power GARCH model

The Power GARCH (PGARCH) model was first introduced by Ding, Z, Ganger, C.W.J and Engle in 1993. PGARCH provides an alternative way to manipulate the volatility with the long memory property, and imposes a Box-Cox power transformation to the conditional standard deviation and to the asymmetric absolute residuals.

The PGARCH ( p, q ) can be expressed as

$$(19) \quad \varepsilon_t = \varepsilon_t \sqrt{h_t} \text{ , where } \varepsilon_t \text{ s are independently and identically distributed as}$$

standard normal  $N(0,1)$

$$h_t^{\frac{1}{2}\lambda} = \omega + \sum_{i=1}^q \alpha_i (|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i})^\lambda + \sum_{j=1}^p \beta_j h_{t-j}^{\frac{1}{2}\lambda}$$

where  $\omega > 0$ ,  $\lambda \geq 0$ ,

$$\alpha_i \geq 0 \text{ for } i= 1, 2, 3, \dots, q \text{ , } -1 < \gamma_i < 0 \text{ for } i= 1, 2, 3, \dots, q.$$

$$\beta_j \geq 0 \text{ for } j = 1, 2, 3, \dots, p.$$

and  $h_t$  is the conditional variance, q is the order of moving average ARCH terms, and p is the order of autoregressive GARCH terms. The  $\alpha$  ,  $\beta$  , and  $\gamma$  are the vectors of unknown parameters.

### 3.7 Threshold GARCH model

The Threshold GARCH (TGARCH) model was proposed by Glosten, Jaganathan, and Runkle(1993) and Zakoian(1994). It can release some restrictions of the linear function of conditional variance. Under the TGARCH model, an additional coefficient parameter  $\gamma_i d_{t-i}$  is added to each  $\varepsilon_{t-i}^2$ . So the TGARCH model can be expressed as follows

$$(19) \quad y_t = u + \varepsilon_t$$

$$h_t = \omega + \sum_{i=1}^q (\alpha_i + \gamma_i d_{t-i}) \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j}$$

where

$$d_{t-i} = \begin{cases} 1, & \text{if } \varepsilon_{t-i} < 0 \\ 0, & \text{if } \varepsilon_{t-i} > 0 \end{cases}$$

If  $\varepsilon_{t-i} > 0$ , it represents good news in the financial market, and the conditional variance with good news can be written as

$$(20) \quad h_t = \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j}$$

The impact to conditional variance is  $\alpha_i \varepsilon_{t-i}^2$  when good news happens.

If  $\varepsilon_{t-i} < 0$ , it shows that bad news happens in the financial market, and the conditional variance with bad news can be written as

$$(21) \quad h_t = \sum_{i=1}^q (\alpha_i + \gamma_i) \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j}$$

The effect of conditional variance is  $(\alpha_i + \gamma_i) \varepsilon_{t-i}^2$  when bad news happens. The parameter  $\gamma_i$  is the measurement of the leverage effect. If  $\gamma_i > 0$ ,  $\alpha_i + \gamma_i > 0$ , it means bad news increases the conditional variance.

Moreover,  $\alpha_i \geq 0$  for  $i = 1, 2, 3, \dots, q$ , and  $\beta_j \geq 0$  for  $j = 1, 2, 3, \dots, p$ . The  $q$  is the order of moving average ARCH terms, and  $p$  is the order of autoregressive GARCH terms. The  $\alpha$ ,  $\beta$  are the vectors of unknown parameters.

### 3.8 The Quadratic GARCH model

Engle and Ng (1993) first introduced the Quadratic GARCH (QGARCH) model. QGARCH model is a partially non-parametric model and allows the consistent estimation

of the new conditional variance impact under a certain range. So the QGARCH model can be given by

$$(22) \quad y_t = \varepsilon_t = \varepsilon_t \sqrt{h_t}$$

where  $\varepsilon_t$  s are independently and identically distributed as standard normal  $N(0,1)$

$$h_t = \omega + \sum_{i=1}^q \alpha_i (\varepsilon_{t-i} - \psi_i)^2 + \sum_{j=1}^p \beta_j h_{t-j}$$

Notice that if  $\psi_i > 0$ ; bad news ( $\varepsilon_{t-i} < 0$ ) has a larger effect on the conditional variance, and good news ( $\varepsilon_{t-i} > 0$ ) has less impact on the conditional variance. Moreover,  $\alpha_i \geq 0$  for  $i = 1, 2, 3, \dots, q$ , and  $\beta_j \geq 0$  for  $j = 1, 2, 3, \dots, p$ . The  $q$  is the order of moving average ARCH terms, and  $p$  is the order of autoregressive GARCH terms. The  $\alpha$ ,  $\beta$  are the vectors of unknown parameters needed to be estimated.

### 3.9 Normality Assumption for Conditional Time Series Data

As in most classical time series models the traditional GARCH models also assumes that the conditional distribution of the time series data  $y_t$  is normal. Thus the conditional distribution of  $y_t$  given  $y_{t-1}$  can be written as,

$$(23) \quad f(y_t | y_{t-1}) = \frac{1}{\sqrt{2\pi h_t}} \exp\left(-\frac{(y_t - x_t' \beta)^2}{2h_t}\right),$$

where  $h_t$  is the conditional variance of all forms of GARCH Models, presented above.

For example from equation (9) we have the following form of  $h_t$

$$h_t = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j} \quad (p \geq 0, q > 0)$$

Notice that the above conditional variance function of the traditional GARCH model can be replaced by any conditional variance function  $h_t$  from other forms of GARCH models in section 3.4-3.7.

However, because of positive kurtosis and Jarque-Bera tests from the Chapter 2, there is substantial evidence to point out that the normality assumption may not hold for the data considered for this thesis. In such cases a more robust distribution against the violation of normality assumption, the student's  $t$  distribution can be considered. In order to analyze the time series data with excess kurtosis, Bollerslev (1987) proposed the student's  $t$  GARCH models, which added an additional parameter, degrees of freedom, for the student's  $t$  distribution, and it can be written as

$$(24) \quad f(y_t | y_{t-1}) = \frac{\Gamma[\frac{1}{2}(V+1)]}{\pi^{\frac{1}{2}}\{\Gamma[\frac{1}{2}V]\}} [(V-2)h_t]^{-\frac{1}{2}} \left\{1 + \frac{y_t^2}{(V-2)h_t}\right\}^{-\frac{1}{2}(V+1)} \quad V > 2$$

where  $V$  is the degrees freedom of the students'  $t$  distribution.  $h_t$  is the conditional variance drawn from the GARCH model and can be written as follows:

$$h_t = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j} \quad (p \geq 0, q > 0)$$

Noticed again, the above conventional GARCH conditional variance formulation can be replaced by any conditional variance function  $h_t$  from other forms of the GARCH models.

### 3.10 Estimation Methods

In order to use the time series models to forecast volatility, the first step is to fit the data into the time series model, which can be done by estimating the unknown parameters of the model.

In this paper, we use both the maximum likelihood approach to estimate unknown parameters for all the forms of the GARCH models, and the Bayesian approach for the GARCH model with the students'  $t$  innovation. These methods are presented briefly as follows.

### 3.10.1 Maximum Likelihood Estimation for the Normal Distribution

Recall that  $y_t = \varepsilon_t = \varepsilon_t \sqrt{h_t}$ , and  $\varepsilon_t$  is independent and identically distributed with standard normal. So  $\varepsilon_t | \psi_{t-1} \sim N(0, h_t)$ , and the conditional function of  $\varepsilon_t$  is given by

$$(24) \quad f(\varepsilon_t | \psi_{t-1}) = \frac{1}{\sqrt{2\pi h_t}} \exp\left(-\frac{(\varepsilon_t)^2}{2h_t}\right)$$

and multiplication of probability density function can be written as

$$(25) \quad \begin{aligned} f(\varepsilon = \varepsilon_1, \varepsilon_2, \dots, \varepsilon_n | \psi_{t-1}) &= f(\varepsilon_1 | \psi_{t-1}) * f(\varepsilon_2 | \psi_{t-1}) * \dots * f(\varepsilon_n | \psi_{t-1}) \\ &= \frac{1}{(\sqrt{2\pi})^n} \prod_{i=1}^n \exp\left(-\frac{(\varepsilon_i)^2}{2h_t}\right) * \frac{1}{\sqrt{h_t}} \end{aligned}$$

where  $h_t$  is the conditional variance, and is the function of the unknown parameters. So the natural logarithm of the likelihood function can be written as

$$(26) \quad l = -\frac{1}{2} \sum_{i=0}^n [\log 2\pi + \log h_t + \frac{\varepsilon_t^2}{h_t}]$$

The log likelihood is then maximized to estimate the unknown parameters involved.

### 3.10.2 Maximum Likelihood Estimation for the Student's t Distribution

The alternative way to analyze the time series data set with the excess kurtosis is students'  $t$  GARCH model. So the conditional probability density function of  $\varepsilon_t$  is given by

$$(27) \quad f(\varepsilon_t | \psi_{t-1}) = \frac{\Gamma\left[\frac{1}{2}(V+1)\right]}{\pi^{\frac{1}{2}} \Gamma\left[\frac{1}{2}V\right]} [(V-2)h_t]^{-\frac{1}{2}} \left\{1 + \frac{\varepsilon_t^2}{(V-2)h_t}\right\}^{-\frac{1}{2}(V+1)} \quad V > 2$$

and the multiplication of probability density function can be written as

$$(28) \quad \begin{aligned} f(\varepsilon = \varepsilon_1, \varepsilon_2, \dots, \varepsilon_n | \psi_{t-1}) &= f(\varepsilon_1 | \psi_{t-1}) * f(\varepsilon_2 | \psi_{t-1}) * \dots * f(\varepsilon_n | \psi_{t-1}) \\ &= \prod_{i=1}^n \left\{ \frac{\Gamma\left[\frac{1}{2}(V+1)\right]}{\pi^{\frac{1}{2}} \Gamma\left[\frac{1}{2}V\right]} [(V-2)h_t]^{-\frac{1}{2}} \left\{1 + \frac{\varepsilon_t^2}{(V-2)h_t}\right\}^{-\frac{1}{2}(V+1)} \right\} \end{aligned}$$

where  $h_t$  is the conditional variance function with unknown parameters. So the likelihood function can be given by

(29)

$$l = \log[\Gamma(\frac{\nu+1}{2})] - \log[\Gamma(\frac{\nu}{2})] - \frac{1}{2} \sum_{i=1}^n \{ \pi h_t \log[(\nu-2)] + (\nu+1) \log(1 + \frac{\varepsilon_t^2}{h_t(\nu-2)}) \}$$

The log likelihood is then maximized to estimate the unknown parameters involved. Due to the complex form of the log likelihood, either numerical optimization or Bayesian approaches are warranted.

### 3.10.3 The Bayesian Estimation of GARCH model

Ardia(2010) presented the GARCH (1, 1) model with students'  $t$  – innovations can be written as

$$(30) \quad y_t = \varepsilon_t \sqrt{\frac{\nu-2}{2} \theta_t h_t} \quad t= 1, 2, 3, \dots, T$$

where  $\varepsilon_t$  s are independently and identically distributed as standard normal  $N(0, 1)$  and  $\theta_t$ s are independently and identically distributed as Inverted Gamma( $\frac{\nu}{2}, \frac{\nu}{2}$ )

$$h_t = \omega + \alpha_1 y_{t-1}^2 + \beta_1 h_{t-1}$$

Where  $\omega > 0, \alpha_1 \geq 0, \text{ and } \beta_1 \geq 0, \text{ and } \nu > 2$ , and N represents the normal distribution.

The parameter  $\alpha, \beta, \nu, \text{ and } \theta$  are the unknown parameters.

The Maximum Likelihood Estimation only considers maximizing the likelihood function. However, the idea of the Bayesian estimation is to consider all the possible value of unknown parameters. Bayesian estimation is based on the simple Baye's Rule to build the posterior distribution from the likelihood function and the joint prior distribution of the unknown parameters. The posterior distribution can be written as follows,



$$\begin{aligned}
(31) \quad p(\alpha, \beta, v, \text{ and } \theta | y_t) &= \frac{p(y_t | \alpha, \beta, v, \theta) p(\alpha, \beta, v, \theta)}{p(y_t)} \\
&= \frac{l(\alpha, \beta, v, \theta | y_t) p(\alpha, \beta, v, \theta)}{p(y_t)}
\end{aligned}$$

In order to compute the  $p(y_t)$ , we will integrate it as follows

$$(32) \quad p(y_t) = \int p(y_t | \alpha, \beta, v, \theta) p(\alpha, \beta, v, \theta) d\alpha d\beta dv d\theta$$

Due to the complex form of the integrand, it is hard to find a closed form solution for the equation (32). Sampling based numerical methods such as Markov Chain Monte Carlo simulation is considered to carry out parameter estimation.

### 3.11 Summary

The main purpose of this chapter is to introduce the different forms of the GARCH models and the parameter estimation methods using both the Likelihood and Bayesian approaches. In chapter 4 we present the results from the analysis of the three data sets using these forms of GARCH models under both the normal distribution and the student's t distributional assumption. Bayesian estimation approach for the GARCH model with student's t innovations is also considered.

## Chapter 4: Empirical Results

In this chapter, we present results from fitting GARCH, the EGARCH, the IGARCH, the PGARCH, the QGARCH, and the TGARCH models for  $p = 1$  and  $q = 1$  values. We consider three different stock exchanges data from three different countries representing three influential economics in the world. These are the NIKKEI 225 index from Japan, the S&P 500 index from the United States, and the DAX index from Germany. Comparative results in the form of prediction of forecasting errors from these GARCH models for these three datasets are presented in this chapter. The Maximum Likelihood Estimation is used to estimate the parameters of forms of the GARCH models, while the Bayesian approach is used to estimate the traditional GARCH model.

The primary assumption about the forms of GARCH models for conditional time series data is that it follows the normal distribution. All three datasets show excess kurtosis and positive skewness. Thus we consider different forms of the GARCH models with student's  $t$  distribution. We fit twelve models: the normal GARCH, the student's  $t$  GARCH, the normal EGARCH, the student's  $t$  EGARCH, the normal IGARCH, the student's  $t$  IGARCH, the normal TGARCH, the student's  $t$  TGARCH, the normal PGARCH, the student's  $t$  PGARCH, the normal QGARCH, and the student's  $t$  QGARCH.

The Akaike's Information Criteria (AIC) and Log-Likelihood value are used to evaluate the performance of the forms of the GARCH models. They are defined as follows:

$$(1) \quad AIC=2k - 2\ln(L)$$

where L is the maximized value of the likelihood function for the model, and k is the number of the parameters in the model.

Moreover, the sum of squared of the forecasting error is used to determine which provides the most accurate forecasting results. The sum of squared of forecasting errors (SSFE) can be written as

$$(2) \quad SSFE= \sum_{i=1}^T e^2_t$$

where T is the number of forecasting observations, and  $e_t$  is forecasting error obtained by subtracting the forecasted value from the observed value.

In order to calculate the SSFE, we drew the daily returns for next 20 observations from the real situation, and compare values with the corresponding 20 observations predicted from the model. The SSFE tells us how reliable our models are with respect to forecasting.

Furthermore, the parameter estimation of the twelve forms of the GARCH model is presented in this chapter as well. The parameter  $\alpha_1$  and  $\beta_1$  are represented as the ARCH and the GARCH coefficients. Please note that the key assumption of the IGARCH model is the sum of  $\alpha_1$  and  $\beta_1$  equals one. The parameter  $\gamma$  of the PGARCH and the EGARCH models shows whether the data sets have the leverage effect or not. For the TGARCH model, the parameter  $\gamma$  indicates whether bad news increases volatility or not.

The parameter  $\nu$  is the inverse value of the degree of freedom for all the forms of the GARCH model with the student's  $t$  distribution. .

This chapter is organized as follows: Sections 4.1, 4.2, and 4.3, all of which provide the empirical analysis results for the Nikkei 225 index, the S & P 500 index, and the DAX index based on the Maximum Likelihood Estimation, respectively. Section 4.4 presents the empirical results for these three datasets using the traditional GARCH models with student's  $t$  distribution based on the Bayesian approach. Section 4.5 presents the summary for this chapter.

#### 4.1 The Nikkei 225

The main results of this section are presented in the following Tables 1 and 2. In Table 1, we report the evaluation for the performance of all twelve competing forms of the GARCH models based on the AIC, the Log-likelihood value, and the sum squared of the forecasting error (SSFE). While in Table 2, we present the estimated parameters value of all twelve competing forms of the GARCH models.

From Table 1, the log-likelihood values for each form of the GARCH models are estimated. It is to be noted that each form of the GARCH models with student's  $t$  distribution generally have the bigger log-likelihood value. It indicates that any form of the GARCH models with the normal distribution are significantly worse than any forms of the GARCH models with the student's  $t$  distribution. Any competing forms of GARCH model which gives the bigger log-likelihood value will be the better fitting model. The QGARCH model with student's  $t$  distribution has the biggest log-likelihood value, which is 17,008.825.

From the AIC values for each of the forms of the GARCH models, Positive kurtosis and smaller AIC values shows the assumption of the student's t distribution is generally better than the competing assumption of the normal distribution. And the GARCH model with the less AIC values will be the best fitting model. The QGARCH with student's t distribution has the lowest AIC value, which is -34,002.45.

The value of the sum of squared of the forecasting error (SSFE) to tell us the how well the model does the forecasting. The smaller the value, the more precise is the precision. The QGARCH model with the student's t distribution gives us the lowest sum of forecasting error, which is 0.000985. The QGARCH with the student's t distribution is the most precise model for the Nikkei 225 index.

**Table 1 Goodness-of-Fit**

	AIC	Log-Likelihood value	SSFE
GARCH(1,1)~Normal Distribution	-28137.655	14,073.83	0.0012145
GARCH(1,1)~STD T Distribution	-28273.288	14,142.64	0.001196
EGARCH(1,1)~Normal Distribution	-286260.5	14,136.25	0.0012777
EGARCH(1,1)~STD T Distribution	-28710.54	14,187.81	0.0011489
IGARCH(1,1) ~Normal Distribution	-28119.532	14,063.77	0.0012146
IGARCH(1,1) W/ Integer~STD T Distribution	-28236.305	14,139.05	0.0011941
PGARCH(1,1) ~ Normal Distribution	-28263.097	14,138.55	0.0012737
PGARCH(1,1) ~ STD T Distribution	-33,524.49	16,770.25	0.0012574
TGARCH(1,1)~Normal Distribution	-28244.837	14,128.42	0.0012618
TGARCH(1,1)~STD T Distribution	-33,986.45	17,000.23	0.001078
QGARCH(1,1)~Normal Distribution	-28,273	14,142.45	0.001378
QGARCH(1,1)~STD T Distribution	<b>-34002.45</b>	<b>17008.225</b>	<b>0.000985</b>

Table 2 reports the estimated parameters of the 12 forms of GARCH models defined in Chapter 3. For all the forms of the GARCH models, except the PGARCH with the student's t distribution, the sum of  $\alpha_1$  and  $\beta_1$  coefficients are very close to one, indicating that volatility shocks are quite persistent, according to Floros (2008). The parameter  $\alpha_1$  represents the previous squared returns, and it is positive and statistically significant for most of the forms of the GARCH models, except the PGARCH with the student's t distribution. The coefficient of the previous conditional variance ( $\beta_1$ ) is positive and less than one, indicating the previous information of volatility is important. Moreover, the EGARCH models with both normal distribution and student's t

distribution show the significant and negative  $\gamma$  parameter, which proves that the leverage effect in data sets exists. From the TGARCH models, the estimated parameter  $\gamma$  is significantly greater than zero, which proves that the bad news has larger effect to the volatility. Furthermore, the parameter  $\gamma$  for QGARCH models are significantly greater than 0, which also indicates the bad news has large effects on volatility.

To sum up the Nikkei 225 index, the QGARCH with student's t distribution is the most precise model for forecasting volatility. Based on the parameter estimate results, the Nikkei 225 index has existing leverage effects. The previous volatility of this index impacts the current volatility. The bad news of the Japan stock market has largely impact on the volatility.

**Table 2 Forms of GARCH models for Volatility**

	$\omega$	$\alpha_1$	$\beta_1$	$\nu$	$\gamma$	$\lambda$
GARCH(1,1)-Normal Distribution	0.000005263***	0.0992***	0.8797***			
GARCH(1,1)-STD T Distribution	0.000003066***	0.0823***	0.9067***	0.1144***		
EGARCH(1,1)-Normal Distribution	(-)0.2542***	0.1732***	0.9699***	N/A	(-)0.5474***	
EGARCH(1,1)-STD T Distribution	(-)0.203261***	(-)0.089136***	0.976544***	9.980590***	(-)0.15149***	
IGARCH(1,1)-Normal Distribution	0.0000029458***	0.1133***	0.8867***			
IGARCH(1,1) ~STD T Distribution	0.0000018602***	0.0892***	0.9108***	0.1246***		
PGARCH(1,1) ~ Normal Distribution	0.000155***	0.0892***	0.5406***		0.8991***	0.6104***
PGARCH(1,1) ~ STD T Distribution	0.000314	0.0973	0.3566***	0.1434***	1	0.8761***
TGARCH(1,1)-Normal Distribution	0.0000051076***	0.0278***	0.8892***		0.1221***	
TGARCH(1,1)-STD T Distribution	0.000003678***	0.0229**	0.9054***	0.1029***	0.1121***	
QGARCH(1,1)-Normal Distribution	0	0.0906***	0.8804***		0.008722***	
QGARCH(1,1)-STD T Distribution	0	0.0816***	0.8942***	0.0990***	0.008213***	

Note: \*, \*\*,\*\*\* significant at 10%, 5%, and 1% significant level.



## 4.2 The S&P 500 Index

The main results are presented in Table 3 and Table 4, Table 3 report the values for all twelve competing forms of the GARCH models under the AIC, Log-likelihood value and the sum of squared of forecasting error (SSFE). In Table 4, we report the estimated parameters value of all twelve competing forms of the GARCH models.

From the Table 3, the smaller AIC and larger log-likelihood values for each forms of the GARCH models shows that any forms of the GARCH models with student's t distribution are better fitting for the S&P500 Index. The candidate value with the lower AIC value and the higher log-likelihood value will be the better model to estimate the data set. The Table 3 shows that TGARCH with student's t distribution has the lowest AIC value (-38,687.11) and highest log-likelihood value(19,358.55), This implies that TGARCH model with student's t distribution is the better model for data fitting for the S&P 500 index.

However, the value for the sum of forecasting error shows that the traditional GARCH model with student's t distribution gives the smallest SSFE (0.00053115). Thus the traditional GARCH model with the student's t distribution performs better in the presence of volatility, although the TGARCH model with the student's t distribution fits the data well.

**Table 3: Goodness-of-Fit**

	AIC	Log-Likelihood	SSFE
GARCH(1,1)~Normal Distribution	-32,599.58	16,304.79	0.00057065
GARCH(1,1)~STD T Distribution	-32,785.02	16,398.52	<b>0.00053115</b>
EGARCH(1,1)~Normal Distribution	-32,759.54	16,385.77	0.00057119
EGARCH(1,1)~Student's t Distribution	-33,689.13	16,483.52	0.00056186
IGARCH(1,1)~Normal Distribution	-32,597.27	16,302.64	0.00057141
IGARCH(1,1) W/STD T Distribution	-32,786.95	16,398.48	0.0005512
PGARCH(1,1) ~ Normal Distribution	-32,709.22	16,384.45	0.00057905
PGARCH(1,1) ~ STD T Distribution	-38,686.64	19,351.32	0.00056031
TGARCH(1,1)~Normal Distribution	-32,765.20	16,388.60	0.00057869
TGARCH(1,1)~STD T Distribution	<b>-38,687.11</b>	<b>19,358.55</b>	0.0005629
QGARCH(1,1)~Normal Distribution	-32,720.50	16,366.25	0.00056568
QGARCH(1,1)~STD T Distribution	-38,639.88	19,326.94	0.00055363

The Table 4 presents the parameter estimates of the twelve forms of the GARCH models given in the previous chapter. Firstly, the shock of volatility is quite persistent for the S&P 500 index, since the sum of the estimates of parameters  $\alpha_1$  and  $\beta_1$  is close to 1. Also, all the values of parameter  $\beta_1$  for the twelve models are significantly positive and less than one, this implies the estimate of that effect of the prior volatility is important. In addition, the statistically significant negative  $\gamma$  parameter of the EGARCH model

indicates the existence of leverage and asymmetric effect in the daily returns of the S&P index. Moreover, the results of the estimation of the PGARCH model confirm that the asymmetric effects exist because of the significant positive parameter  $\gamma$ . Furthermore, the TGARCH model shows that bad news in stock markets increases the volatility. A statistically significant positive parameter  $\gamma$  provide evidence to support this.

To sum up, the TGARCH model with student's t distribution is a better model fitting the S&P 500 index. However, the traditional GARCH model with the student's t distribution is the precise model for forecasting. Overall the results for the parameter estimation shows, the S&P 500 index has the leverage and asymmetric effect, and that volatility increases by bad news in stock market.

**Table 4 GARCH-family models for Volatility**

	$\omega$	$\alpha_1$	$\beta_1$	$\nu$	$\gamma$	$\lambda$
GARCH(1,1)-Normal Distribution	9.255e-7***	0.0785***	0.9156***			
GARCH(1,1)-STD T Distribution	5.7798e-7***	0.0708***	0.9272***	0.1427***		
EGARCH(1,1)-Normal Distribution	(-)0.1681***	0.1291***	0.9815***		(-)0.846***	
EGARCH(1,1)-Student's t Distribution	(-)0.2203***	0.1567***	0.8962***	0.4568***	(-)0.782***	
IGARCH(1,1)-Normal Distribution	7.0098e-7***	0.0829***	0.9171***			
IGARCH(1,1) W/STD T Distribution	5.15885e-7***	0.0724***	0.9276***	0.1463***		
PGARCH(1,1) - Normal Distribution	0.000402***	0.0804***	0.9231***		0.7820***	0.4224***
PGARCH(1,1) - STD T Distribution	1.5547e-6***	0.0359	0.9307***	0.0126***	0.9998	0.9290***
TGARCH(1,1)-Normal Distribution	1.3064e-6***	-0.00694	0.9239***		0.1405***	
TGARCH(1,1)-STD T Distribution	0.0000009359***	-0.1	0.9317***	0.1249***	0.1365***	
QGARCH(1,1)-Normal Distribution	0	0.0822***	0.8985***		0.005149***	
QGARCH(1,1)-STD T Distribution	0	0.0771***	0.9089***	0.1288***	0.004490***	

Note: \*, \*\*, \*\*\* significant at 10%, 5%, and 1% significant level.

### 4.3 The DAX index

The main results of the DAX index are shown in the Table 5 and 6. Table 5 presents the goodness-of-fit statistics for the twelve forms of the GARCH models under the AIC, Log-likelihood value and the sum of squared of forecasting error (SSFE). In Table 6, we report the estimated parameters value of all twelve competing forms of the GARCH models for the DAX index.

From Chapter 2, we evidenced positive kurtosis for the DAX index data. Moreover, the distribution of the conditional returns is not normal. We assume that the forms of the GARCH model will not best fit the DAX data. And the results based on the value of AIC, the Log-Likelihood, and the sum of forecasting error supports that our assumption. From the Table 5, the log-likelihood values for each forms of the GARCH models points out that that each form of the GARCH models with student's t distribution generally have the bigger log-likelihood value. It informs that any forms of the GARCH models with normal distribution are significantly worse than any forms of the GARCH models with student's t distribution. And the any candidate forms of GARCH give the bigger log-likelihood value will be better fitting model. The PGARCH model with student's t distribution has the biggest log-likelihood value, which is 18,094.96.

From the AIC values for each the forms of the GARCH models, the assumption of the student's t distribution is generally better than the competing assumption of the normal distribution. And the GARCH model with the less AIC values will be the best fitting model. The PGARCH with student's t distribution has the lowest AIC value, which is -37,173.93.

The PGARCH model with student's t distribution give us the lowest sum of forecasting error value, which is 0.002471.

The PGARCH model with student's t best compared to other forms of the GARCH models fits the DAX index data.

**Table 5 The Goodness-of-Fit for the DAX index**

	AIC	Log-Likelihood	SSFE
GARCH(1,1)~Normal Distribution	-30182.669	15,096.33	0.0024715
GARCH(1,1)~STD T Distribution	-30275.615	15,143.81	0.0024499
EGARCH(1,1)~Normal Distribution	-30287.836	15,149.92	0.0025387
EGARCH(1,1)~Student's t Distribution	-30,349.47	15,152.68	0.0024991
IGARCH(1,1)~Normal Distribution	-30,177.06	15,092.53	0.0024728
IGARCH(1,1) W/STD T Distribution	-30,276.37	15,143.18	0.0024486
PGARCH(1,1) ~ Normal Distribution	-30,303.15	15,158.58	0.0025444
PGARCH(1,1) ~ STD T Distribution	<b>-36,173.93</b>	<b>18,094.96</b>	<b>0.002471</b>
TGARCH(1,1)~Normal Distribution	-30,285.85	15,148.92	0.0025446
TGARCH(1,1)~STD T Distribution	-36,164.58	18,089.29	0.0025036
QGARCH(1,1)~Normal Distribution	-30,282.63	15,147.32	0.0025358
QGARCH(1,1)~STD T Distribution	-36163.476	18,088.74	0.0024915

Table 6 is shown the parameter estimates of twelve forms GARCH models. For all the twelve models, the sum of parameter  $\alpha_1$  and  $\beta_1$  are closely equal one, indicating that the shocks of volatility continue. Meanwhile, the parameter  $\beta_1$  which is the GARCH

coefficient are all positive, and statistically significant. It indicates that the effect of previous information on volatility is significant. Moreover, the negative and significant parameter  $\gamma$  from the EGARCH model proves that the leverage and asymmetric effects in daily returns for the case of the DAX index. The significantly positive parameter  $\gamma$  from the PGARCH models also confirms the leverage and asymmetric effects existence in the case of the DAX index. Furthermore, the significantly positive parameter  $\gamma$  of the TGARCH models shows that the bad news of stock markets increases the volatility.

To sum up, the most precise model for forecasting volatility for the DAX index is the PGARCH model with the student's t distribution. And the leverage and asymmetric effects exists in daily returns of the DAX index, and also volatility is increased by the bad news of the stock market.

**Table 6 Parameter Estimation**

	$\omega$	$\alpha 1$	$\beta 1$	$\nu$	$\gamma$	$\lambda$
GARCH(1,1)-Normal Distribution	2.1708e-8***	0.0908***	0.8996***			
GARCH(1,1)-STD T Distribution	1.4214e-6***	0.0839***	0.9115***	0.0985***		
EGARCH(1,1)-Normal Distribution	(-)0.1738***	0.1636***	0.9798***		(-)0.0418***	
EGARCH(1,1)-Student's t Distribution	(-)0.2365***	0.1245***	0.8736***	0.7621***	(-)0.0237***	
IGARCH(1,1)-Normal Distribution	1.5613e-6***	0.0994***	0.9006***			
IGARCH(1,1) W/STD T Distribution	1.1298e-6***	0.0872***	0.9128***	0.1038***		
PGARCH(1,1) ~ Normal Distribution	0.0000873*	0.0845***	0.9094***		0.5285***	0.6236***
PGARCH(1,1) ~ STD T Distribution	0.000354*	0.0846***	0.9161***	0.0867***	0.6447***	0.4605***
TGARCH(1,1)-Normal Distribution	2.6996e-6***	0.0254***	0.9018***		0.1151***	
TGARCH(1,1)-STD T Distribution	2.0346e-6***	0.0223**	0.9096***	0.0860***	0.1136***	
QGARCH(1,1)-Normal Distribution	8.06E-08	0.0006295***	0.0824***	0.9002***		
QGARCH(1,1)-STD T Distribution	0	0.0832***	0.9022***	0.0848***	0.005813***	

Note: \*, \*\*, \*\*\* significant at 10%, 5%, and 1% significant level.



## 4.5 The Bayesian Estimation

Ardia (2007) presents the R package named the bayesGARCH, which provides the Bayesian approach to estimate the traditional GARCH model with the student's t distribution. He derived the Markov Chain Monte Carlo (MCMC) method with the Metropolis-Hastings algorithm for the Bayesian Estimation with GARCH model.

Recall from chapter 3, the GARCH (1,1) model with student's t distribution can be written as

$$y_t = \varepsilon_t \sqrt{\frac{v-2}{2} \theta_t h_t} \quad t= 1, 2, 3, \dots, T$$

$\varepsilon_t$  is identically and independently distributed as the N (1,1).

$\theta_t$  is identically and independently distributed as the Inverted Gamma( $\frac{v}{2}, \frac{v}{2}$ )

$$h_t = \omega + \alpha_1 y_{t-1}^2 + \beta_1 h_{t-1}$$

Where  $\omega > 0$ ,  $\alpha_1 \geq 0$ , and  $\beta_1 \geq 0$ , and  $v > 2$ , and N represents the normal distribution.

The parameters  $\alpha, \beta, v$ , and  $\theta$  are the unknown parameters.

The posterior distribution is given by

$$p(\alpha, \beta, v, \theta | y_t) = \frac{l(\alpha, \beta, v, \theta | y_t) p(\alpha, \beta, v, \theta)}{p(y_t)}$$

Notice that v is the degree of freedom of the Inverted gamma distribution. The MCMC simulation will generate the above posterior distribution.

Table 7 records the parameter estimation for the traditional GARCH model with the student's t distribution under the Bayesian estimation, as well as the results of the Maximum Likelihood Estimation (MLE).

**Table 7 Parameter estimation based on the Bayesian Approach, compared with MLE**

	MLE	Posterior Dist		MLE	Posterior Dist		MLE	Posterior Dist	
		Mean	Std.D		Mean	Std.D		Mean	Std.D
$\alpha_0$	3.07E-06	3.05E-06	0.00218	5.78E-07	4.96E-07	4.12E-04	1.42E-06	4.99E-04	1.73E-04
$\alpha_1$	0.0823	0.09721	0.00153	0.0708	7.05E-01	0.00247	0.0839	1.38E-01	0.00318
$\beta_1$	0.9067	0.918	0.00754	0.9272	9.15E-01	0.00675	0.9115	0.918	0.00854
$\nu$	N/A	2.05E+00	0.00095	N/A	2.04E+00	0.00074	N/A	2.09E+00	0.00087

From the above table, we see that the parameter estimation based on the Bayesian approach produce similar results compared with the Maximum Likelihood Estimation. Notice that for all three data sets, the sum of parameter  $\alpha_1$  and  $\beta_1$  are close to one, which informs us that past shocks will have a continuous impact on the future volatility.

## Chapter 5 Conclusions

In this thesis paper, we examined the characteristics of daily returns of the Nikkei225 index, the S&P500 index, and the DAX index. The summary results for all these data sets show positive kurtosis, long left tails and violation of normality assumption.

Floros(2008) mentioned that the fat tails and volatility clustering are the two important characteristics for time series data and calls for the family of the GARCH models. All three data sets have these two characteristics, so the GARCH models can be used to analyze and forecast the behavior of volatility. Since the daily returns are non-normally distributed, in order to better manipulate the model, we also consider the student's t distribution as the assumption. The twelve forms of the GARCH models are used to address and forecast volatility for these three data sets. The twelve forms are: the normal GARCH, the student's t GARCH, the normal EGARCH, the student's t EGARCH, the normal IGARCH, the student's t IGARCH, the normal TGARCH, the student's t TGARCH, the normal PGARCH, the student's t PGARCH , the normal QGARCH, and the student's t QGARCH.

The Maximum Likelihood Estimation was first used to fit the models to these data. The results show that for the Nikkei 255 index, the most precise model for forecasting volatility is the QGARCH model with the student's t distribution. For the S&P 500 index, the TGARCH model with the student's t distribution is a better model for fitting the data, and the traditional GARCH with the student's t distribution is the precise model for forecasting volatility. For the DAX index, the PGARCH model with the student's t distribution is the most precise model for forecasting volatility.

Furthermore, in the summary of the results, the sum of parameter  $\alpha_1$  and  $\beta_1$  are close to one for all the three data sets. It indicates the past information has the continuous impact to the current and further conditional variance. For all the three data sets, we do find that most of the coefficients of the conditional variance equations for twelve forms of the GARCH model are significant, and it implies that large shocks for the stock markets significantly impact the future volatility. Meanwhile, the results also indicate that when the bad news happens, it will have a large impact on volatility.

Based on these finds, we suggest to investors, researchers, and analysts be prepared to deal with the potential economic crisis, because the bad shocks of the stock markets has the large effect to the volatility.

For future research, we will consider more alternative time series models to address volatility, such as the stochastic volatility models, and some multivariate time series models. In order to consider related features of financial data, we will consider for various stock exchange data.

Moreover, the three data sets we considered on this paper are the low-frequency data, which are the daily observations. Anderson, Bollerslev, Diebold, and Labys (2003)

point out that the evaluation criteria may be less natural rely on the low-frequency data. Zhou (1996) suggests that the high-frequency data, such as tick by tick data will be more effective for forecasting volatility. For our future research, we will consider the high-frequency data instead of the low-frequency data, and apply various forms of the GARCH models or other suitable methods.

## References

- Akgiray, V. Conditional Heteroscedasticity in Time Series of Stock Returns: Evidence and Forecast. *Journal of Business*, 62: pp.55-80, 1989.
- Andersen.T.G,Bollerslev.T,Diebold.F.X and Labys.P “Modeling and Forecasting Realized Volatility” *Econometrica* 71.2.pp. 579-625, 2003.
- Ardia, D.,Hoogerheide, L.F., “Bayesian Estimation of the GARCH(1,1) Model with Student-t Innovations,” *The R Journal*,2/2,pp.41-47 2010
- Bollerslev, T. “Generalized Autoregressive Conditional Heteroskedasticity,” *Journal of Econometrics*,31, pp.307–327.1986
- Borbely J., “U.S. labor market in 2008: economy in recession”, *Monthly Labor Review Online*, Vol.132.No.3. 2009
- Brandt.M.W, Jones.C.S “Forecasting With Range-Based EGARCH Models.” *Journal of Business and Economics Statistics*, 24, pp.470-486. 2004 October.
- Brooks.C. “Predicating Stock Index Volatility: Can Market Volumn Help?” *Journal of Forecasting*, 17,pp.59-80, 1998
- Busse.J.A, “Volatility Timing in Mutual Funds : Evidence from Daily Returns.” *The Review of Financial Studies* 12.5. pp 1009-1041.1999
- Clark.G.L, Mansfied.G, and Tickell. A, “University Emergent Frameworks in Global Finance: Accounting Standards and German Supplementary Pensions” *Economic Geography* 77.3.pp.250-271.2001
- CME Group .“Nikkei 225 Futures and Options” Equity Products. EQ156/0/1207 2009
- Dickey, D. A., and Fuller.W. A, “Likelihood ratio statistics for autoregressive time series with a unit root.” *Econometrica* 49.4: pp.1057–72. 1981.

- Ding, Z., Granger, C. W. J., and Engle, R. "A Long Memory Property of Stock Market Returns and a New Model." *Journal of Empirical Finance*, 1, 83–106.1993
- Edel.T,and Luecy.B.M "A Power GARCH examination of the Gold Market." *Research in International Business and Finance* 21.2 pp.316-325.2007
- Engle, R.F. "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation,"*The Econometrica*, 50, 987–1007, 1982.
- Engle, F.R GARCH 101: "The Use of ARCH/GARCH Models in Applied Econometrics." *Journal of Economic Perspectives* 15(4): 157-168, 2001.
- Engle, R. F. and Bollerslev, T. "Modeling the Persistence of Conditional Variances," *Econometric Review*, 5, 1–50.1986
- Engle, R. F. and Ng, V. K. "Measuring and Testing the Impact of News on Volatility," *Journal of Finance*, 48, 1749–1778.1993
- Floros C. "Modeling Volatility using GARCH Models: Evidence from Egypt and Israel." *Middle Eastern Finance and Economics* 2 2008.
- French.K.R,Schwert.G.W, Stambaugh.R.F, "Expected Stock Returns and Volatility." *Journal of Financial Economics* 19 pp.3-29.1987
- Glosten, L., Jaganathan, R., and Runkle, D. "Relationship between the Expected Value and Volatility of the Nominal Excess Returns on Stocks," *Journal of Finance*, 48, 1779–1802.1993
- Kim.C.J, Morley.J.C, Nelson.C.R, "Is there a Positive Relationship between Stock Market Volatility and the Equity Premium?" *Journal of Money, Credit and Banking*,36,3 pp.339-360, 2004.
- Kirchler M, Huber J., "Fat tails and volatility clustering in experimental asset", *Journal of Economic Dynamics and Control*, Vol. 31, pp. 1844-1874.2007
- Nelson, D. B. "Conditional Heteroskedasticity in Asset Returns: A New Approach," *The Econometrica*, 59, 347–370. 1991
- Ning, Cathy; Xu, Dinghai; and Wirjanto, Tony, "Modeling Asymmetric Volatility Clusters Using Copulas and High Frequency Data" *Economics Publications and Research*. 32. 2010
- Noguchi Y., "The 'Bubble' and Economic Policies in the 1980S", *Journal of Japanese Studies*, Vol.20, pp. 55-80. 1994

- Pagan.A.R, Schwert.G.W “Alternative models for conditional stock volatility” *Journal of Econometrics*, 45, 1-2, 1990
- Siegel. J. J and Schwartz. J. D “Long-Term Returns on the Original S&P 500 Companies” *Financial Analysts Journal*, Vol. 62, No. 1, pp. 18-31.2006
- Sokolinskiy O. Dijk D.V. Forecasting Volatility with Copula-Based Time Series Models. *Tinbergen Institute Discussion Paper*, 2011.
- Wilson.A.G, Ghahramani.Z “ Copula Processes” *Cornell University Library*  
arXiv:1006.1350. 2010
- Zhou.B, “High-Frequency Data and Volatility in Foreign-Exchange Rates” *Journal of Business & Economic Statistics* 14.1.pp.45-52,1996
- Zakoian, J. M. “Threshold Heteroscedastic Models,” *Journal of Economic Dynamics and Control*,18, 931–955. 1994



## Appendix I SAS Code

### The Nikkei 225

```
libname nikkei3 "C:\Users\Administrator\Desktop\revised
data";
data nikkei3.nikkei4;
    set YIHANLI4;
RUN;
data nikkei3.analysis;
set nikkei3.nikkei4;
time=_N_;
if Rt ^='';
time=_N_-1;
run;
proc print data=nikkei3.analysis(obs=100);
run;
data nikkei3.analysis;
set nikkei3.analysis;
time=_N_;
run;
proc univariate data=nikkei3.analysis;
var Close;
run;
proc sgplot data=nikkei3.analysis;
    title1 Price;
    symbol i=join h=2;
    yaxis values=(7000 to 24000 by 1700) label='Daily Closing
Price';
    xaxis values=(0 to 5000 by 500) label='time';
    series x=time y=Close;
run;
proc univariate data=nikkei1.analysis;
var Rt;
run;

proc sgplot data=nikkei3.analysis;
    title1 Returns;
    symbol i=join h=2;
    yaxis values=(-0.14 to 0.14 by 0.02) label='Daily
Returns';
    xaxis values=(0 to 5000 by 500) label='time';
    series x=time y=Rt;
run;
Data Rawdata;
```

```

set nikkei3.analysis;
x_1st_LAG          = LAG1(Close);
x_1st_DIFF         = DIF1(Close);
x_1st_DIFF_1st_LAG =DIF1(LAG1(Close));
x_1st_DIFF_2nd_LAG =DIF1(LAG2(Close));
x_1st_DIFF_3th_LAG =DIF1(LAG3(Close));
x_1st_DIFF_4th_LAG =DIF1(LAG4(Close));
x_1st_DIFF_5th_LAG =DIF1(LAG5(Close));
run;
proc REG data=Rawdata;
    Model x_1st_DIFF = x_1st_DIFF_1st_LAG
                    x_1st_DIFF_2nd_LAG
                    x_1st_DIFF_3th_LAG
                    x_1st_DIFF_4th_LAG
                    x_1st_DIFF_5th_LAG;
run;

data nikkei20;
input time Open close Adj_close Rt;

datalines;
1 . . . . .
2 . . . . .
3 . . . . .
4 . . . . .
5 . . . . .
6 . . . . .
7 . . . . .
8 . . . . .
9 . . . . .
10 . . . . .
11 . . . . .
12 . . . . .
13 . . . . .
14 . . . . .
15 . . . . .
16 . . . . .
17 . . . . .
18 . . . . .
19 . . . . .
20 . . . . .
;
run;

```

```

proc print data=nikkei20;
run;
data analysis1;
set nikkei3.analysis nikkei20;
time = _N_;
run;
proc print data=analysis1;
run;
proc autoreg data=analysis1;
  model Rt = time /garch=(q=1,p=1) maxit=50;
/* type=Exp, INTEG, THRES, POWER, QUADR, DIST=T*/
  output out=out1 cev=vhat P=PRT;
run;
data forecast1;set out1;
  length type $ 8.;
  if Rt^=. then do;
    type = 'estimate'; output;end;
  else do;
    type = 'forecast';output;end;
run;
data garch11;
set forecast1(firstobs=4920 keep =prt vhat time);
run;

Data garch11new;
set garch11;
Rt=0;
if time=4920 then Rt=Rt+0.003795965;
else if time=4921 then Rt=Rt+0.003037322;
else if time=4922 then Rt=Rt-0.007400766;
else if time=4923 then Rt=Rt+0.01355376;
else if time=4924 then Rt=Rt-0.014416325;
else if time=4925 then Rt=Rt+0.010453198;
else if time=4926 then Rt=Rt+0.009893728;
else if time=4928 then Rt=Rt+0.010366428;
else if time=4929 then Rt=Rt+0.014556124;
else if time=4930 then Rt=Rt-5.24747E-05;
else if time=4931 then Rt=Rt+0.002214091;
else if time=4932 then Rt=Rt+0.011133726;
else if time=4933 then Rt=Rt-0.003859441;
else if time=4934 then Rt=Rt-0.000932694;
else if time=4935 then Rt=Rt-0.005463239;
else if time=4936 then Rt=Rt+0.001075271;
else if time=4937 then Rt=Rt+0.000826695;

```

```
else if time=4938 then Rt=Rt+0.007579782;  
else if time=4939 then Rt=Rt-0.005069821;  
else if time=4940 then Rt=Rt+0.010953241;  
run;
```

```
data garchnormal;  
set garchllnew;  
fesq=(Rt-Prt)**2;  
run;  
proc means data=garchnormal sum;  
var fesq;  
run;
```

## The S&P 500

```
libname sp5002 "C:\Users\Administrator\Desktop\revised
data";

data sp5002.sp5003;
set YIHANLI5;
run;

data sp5002.analysis;

set sp5002.sp5003;

if Rt^='';

run;

data sp5002.analysis;

set sp5002.analysis;

time=_N_;

run;

proc print data=sp5002.analysis;
run;
proc univariate data=sp5002.analysis;
var Close;
run;
proc univariate data=sp5002.analysis;
var Rt;
run;
proc sgplot data=sp5002.analysis;

    title Prices;

    symbol i=join h=2;

    yaxis values=(350 to 1600 by 125) label='Daily Closing
Price';

    xaxis value=(0 to 5100 by 510) label='time';

    series x=time y=Close;
```

```

run;
proc sgplot data=sp5002.analysis;

    title1 Returns;

    symbol i=join h=2;

    yaxis values=(-0.10 to 0.13 by 0.01) label='Daily
Returns';

    xaxis values=(0 to 5200 by 500) label='time';

    series x=time y=Rt;
run;
/*Augmented Dickey Fuller tests*/
data rawdatasp;

    set sp5002.analysis;

    x_1st_LAG          = LAG1(Close);
    x_1st_DIFF         = DIF1(Close);
    x_1st_DIFF_1st_LAG = DIF1(LAG1(Close));
    x_1st_DIFF_2nd_LAG = DIF1(LAG2(Close));
    x_1st_DIFF_3rd_LAG = DIF1(LAG3(Close));
    x_1st_DIFF_4th_LAG = DIF1(LAG4(Close));
    x_1st_DIFF_5th_LAG = DIF1(LAG5(Close));

run;
proc REG data=rawdatasp;

    Model x_1st_DIFF = x_1st_LAG
                x_1st_DIFF_1st_LAG
                x_1st_DIFF_2nd_LAG
                x_1st_DIFF_3rd_LAG
                x_1st_DIFF_4th_LAG
                x_1st_DIFF_5th_LAG;

run;
data sp20;
input time open close Adj_close Rt;
datalines;
1 . . . . .
2 . . . . .

```

```

3 . . . . .
4 . . . . .
5 . . . . .
6 . . . . .
7 . . . . .
8 . . . . .
9 . . . . .
10 . . . . .
11 . . . . .
12 . . . . .
13 . . . . .
14 . . . . .
15 . . . . .
16 . . . . .
17 . . . . .
18 . . . . .
19 . . . . .
20 . . . . .
;
run;
proc print data=sp20;
run;
data analysis2;
set sp5002.analysis sp20;
time = _N_;
run;

proc autoreg data=analysis2;

    model Rt=time / garch=(q=1,p=1) maxit=50;
/* type=Exp, INTEG, THRES, POWER, QUADR, DIST=T*/
    output out=out1 cev=vhat;

run;
data forecast2;set out1;
    length type $ 8.;
    if Rt^=. then do;
        type = 'estimate'; output;end;
    else do;
        type = 'forecast';output;end;
run;
*proc print data=forecast2;
*run;
data garch11t;

```

```

set forecast2(firstobs=5047 keep =prt what time);
run;
*proc print data=garch11;
*run;
Data garch11tnew;
set garch11t;
Rt=0;
if time=5047 then Rt=Rt+0.00030953;
else if time=5048 then Rt=Rt+0.002333868;
else if time=5049 then Rt=Rt-0.004960178;
else if time=5050 then Rt=Rt+0.003546597;
else if time=5051 then Rt=Rt+0.011046693;
else if time=5052 then Rt=Rt+0.004926532;
else if time=5053 then Rt=Rt+0.000669232;
else if time=5054 then Rt=Rt+0.000471236;
else if time=5055 then Rt=Rt-0.001026362;
else if time=5056 then Rt=Rt+0.008641669;
else if time=5057 then Rt=Rt-0.005770505;
else if time=5058 then Rt=Rt-0.001594073;
else if time=5059 then Rt=Rt-0.00252535;
else if time=5060 then Rt=Rt-0.00045707;
else if time=5061 then Rt=Rt+0.008860289;
else if time=5062 then Rt=Rt+0.001094493;
else if time=5063 then Rt=Rt+0.014499737;
else if time=5064 then Rt=Rt-0.000423913;
else if time=5065 then Rt=Rt+0.002021269;
else if time=5066 then Rt=Rt+0.002157946;
run;
data garchnormal;
set garch11tnew;
fesq=(Rt-Prt)**2;
run;
proc means data=garchnormal sum;
var fesq;
run;

```



## The DAX Index

```
libname dax1 "C:\Users\Administrator\Desktop\revised data";
data dax1.dax2;
set YIHANLI3;
run;
PROC PRINT data=dax1.dax2;
run;
data dax1.analysis;
set dax1.dax2;
if Rt ^='';
run;
data dax1.analysis;
set dax1.analysis;
time=_N_;
run;
proc univariate data=dax1.analysis;
var Close;
run;
proc sgplot data=dax1.analysis;
title Price;
symbol i=join h=3;
yaxis values=(1340 to 8400 by 1412) label='Daily Closing
Price';
xaxis values=(0 to 5100 by 510) label='time';
series x=time y=Close;
run;
proc univariate data=dax1.analysis;
var Rt;
run;
proc sgplot data=dax1.analysis;
    title Returns;
    symbol i=join h=2;
    yaxis values=(-0.10 to 0.12 by 0.022) labe='Daily
Returns';
    xaxis values=(0 to 5100 by 510) label='time';
    series x=time y=Rt;
run;
Data Rawdata;
set dax1.analysis;
x_1st_LAG          = LAG1(Close);
x_1st_DIFF         = DIF1(Close);
x_1st_DIFF_1st_LAG =DIF1(LAG1(Close));
x_1st_DIFF_2nd_LAG =DIF1(LAG2(Close));
x_1st_DIFF_3th_LAG =DIF1(LAG3(Close));
```

```

x_1st_DIFF_4th_LAG =DIF1(LAG4(Close));
x_1st_DIFF_5th_LAG =DIF1(LAG5(Close));
run;
proc REG data=Rawdata;
    Model x_1st_DIFF = x_1st_DIFF_1st_LAG
                x_1st_DIFF_2nd_LAG
                x_1st_DIFF_3th_LAG
                x_1st_DIFF_4th_LAG
                x_1st_DIFF_5th_LAG;

run;
data dax20;
input time open close Adj_close Rt;
datalines;
1 . . . . .
2 . . . . .
3 . . . . .
4 . . . . .
5 . . . . .
6 . . . . .
7 . . . . .
8 . . . . .
9 . . . . .
10 . . . . .
11 . . . . .
12 . . . . .
13 . . . . .
14 . . . . .
15 . . . . .
16 . . . . .
17 . . . . .
18 . . . . .
19 . . . . .
20 . . . . .
;
    run;
proc print data=dax20;
run;
data analysis2;
set dax1.analysis dax20;
time = _N_;
run;

proc print data=analysis2;
run;

```

```

proc autoreg data=analysis2;

    model Rt=time / garch=(q=1,p=1) dist =t maxit=50;
    output out=out cev=vhat p=prt;
run;
data forecast2;set out;
    length type $ 8.;
    if Rt^=. then do;
        type = 'estimate'; output;end;
    else do;
        type = 'forecast';output;end;
run;
data garchllt;
set forecast2(firstobs=5073 keep =prt vhat time);
run;
Data garchlltnew;
set garchllt;
Rt=0;
if time=5073 then Rt=RT-0.006739486;
else if time=5074 then Rt=RT+0.023933406;
else if time=5075 then Rt=Rt-0.00172793;
else if time=5076 then Rt=Rt+0.004357934;
else if time=5077 then Rt=Rt-0.005864187;
else if time=5078 then Rt=Rt+0.012445269;
else if time=5079 then Rt=Rt+0.01799149;
else if time=5080 then Rt=Rt+0.003411235;
else if time=5081 then Rt=Rt+0.009661154;
else if time=5082 then Rt=Rt-0.001851701;
else if time=5083 then Rt=Rt+0.005019865;
else if time=5084 then Rt=Rt-0.002706943;
else if time=5085 then Rt=Rt+0.000409623;
else if time=5086 then Rt=Rt+0.018207991;
else if time=5087 then Rt=Rt-0.004270672;
else if time=5088 then Rt=Rt-0.010424262;
else if time=5080 then Rt=Rt+0.002241278;
else if time=5090 then Rt=Rt+0.024127115;
else if time=5091 then Rt=Rt+0.005875425;
else if time=5092 then Rt=Rt+0.016545977;
run;
data garchnormal;
set garchlltnew;
fesq=(Rt-Prt)**2;
run;
proc means data=garchnormal sum; var fesq; run;

```

## Appendix II R Code

### The Nikkei 225

```
nikkei225<-read.csv(file="D:/Spring 2013/Thesis/nikkei225.csv")
library(bayesGARCH)
#data(nikkei225)
data(nikkei225)
y<-nikkei225[,2][1:4918]
set.seed(1234)
#1.chain=5000, n.chain=2
y<-nikkei225[,1][1:4918]
set.seed(1234)
MCMC<-bayesGARCH(nikkei225[,1],control=list(1.chain=2000, n.chain=2))
gelman.diag(MCMC)
1-rejectionRate(MCMC)
autocorr.diag(MCMC)
smpl<-formSmpl(MCMC,l.bi=1000,batch.size=2)
summary(smpl)
```

### The S&P 500

```
sp<-read.csv(file="D:/Spring 2013/Thesis/sp500.csv")
library(bayesGARCH)
library(bayesGARCH)
args(bayesGARCH)
y<-sp[,2][1:5046]
set.seed(1234)
MCMC<-bayesGARCH(sp[,2],control=list(1.chain=2000, n.chain=2))
gelman.diag(MCMC)
1- rejectionRate(MCMC)
autocorr.diag(MCMC)
smpl<-formSmpl(MCMC,l.bi=1000,batch.size=2)
summary(smpl)
```

### The DAX Index

```
dax<-read.csv(file="D:/Spring 2013/Thesis/thedax.csv")
library(bayesGARCH)
args(bayesGARCH)
y<-dax[,1][1:5072]
set.seed(1234)
MCMC<-bayesGARCH(dax[,2],control=list(1.chain=2000, n.chain=2))
gelman.diag(MCMC)
1-rejectionRate(MCMC)
autocorr.diag(MCMC)
smpl<-formSmpl(MCMC,l.bi=1000,batch.size=2)
summary(smpl)
```