Title	Modeling the Japanese financial market's volatility and its relationship with other financial markets
Sub Title	
Author	Sung, Edward Kaiyu(Sung, Edward Kaiyu) 西村, 秀和(Nishimura, Hidekazu)
Publisher	慶應義塾大学大学院システムデザイン・マネジメント研究科
Publication year	2010
Jtitle	
JaLC DOI	
Abstract	Since 1982, numerous GARCH (Generalized Autoregressive Conditional Heteroskedascity) models have been developed. Some of the more popular ones include the GJR-GARCH, and EGARCH. Even though these models are effective at mapping out the "true volatility" of an asset, they fail to capture periods of extremely high "true volatility" (0.08). When modeling the American subprime crisis from 2008 to 2009 as out-of-sample data, the maximum volatilities that the GJR-GARCH and GARCH model provided were much less than the maximum "true volatility". Therefore, these GARCH models are not quite suitable for risk management applications as they greatly underestimate the volatility during periods of extremely high volatility (0.08). In this paper, we modified the GARCH model, while using the GJR-GARCH as inspiration, and proposed a new model named GARCH-S. We used the Nikkei 225 and SP 500 as the two asset inputs required by the GARCH-S model. The GARCH-S makes use of a secondary market (SP 500) to increase the volatility forecast of the target market (Nikkei 225). This is possible when the returns of the secondary market at time t-1 is correlated (0.20) with the returns of the target market at time t. Even though we developed a MLE program to estimate the model arameters, the Gibbs Sampling method was used in general to estimate the parameters of the models. The GARCH-S is found to effective in modeling the "upper range" of "true volatility". It greatly reduces underestimate errors as compared to the GARCH, GJR-GARCH and EGARCH models.Under the several assumptions we made, the GARCH-S was found to be ranked second best in forecasting accuracy while effectively capturing the peaks of high volatility better than other GARCH type models. Therefore, the GARCH-S model is an excellent candidate for risk management purposes as it greatly reduces the underestimate errors while still providing competitive forecast accuracy against other GARCH type models.
Notes	修士学位論文. 2010年度システムエンジニアリング学 第23号
Genre	Thesis or Dissertation
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO40002001-00002010-0010

Master's Dissertation 2010

Modeling the Japanese Financial Market's Volatility and Its Relationship with Other Financial Markets

Edward Kaiyu Sung (B.A., B.S., M.A.)

(Student ID Number: 80833286)

Supervisor: Professor Hidekazu Nishimura

March 2011

Graduate School of System Design and Management,
Keio University
Major in System Design and Management

A thesis submitted to the
Graduate School of System Design and Management, Keio University
in partial fulfillment of the requirements for the degree of
Masters in Systems Engineering

SUMMARY OF MASTER THESIS

Student Identification Number	80833286	Name	Edward Kaiyu Sung (B.A., B.S., M.A.)

Title

Modeling the Japanese Financial Market's Volatility
and Its Relationship with Other Financial Markets

Abstract

Since 1982, numerous GARCH (Generalized Autoregressive Conditional Heteroskedascity) models have been developed. Some of the more popular ones include the GJR-GARCH, and EGARCH. Even though these models are effective at mapping out the "true volatility" of an asset, they fail to capture periods of extremely high "true volatility" (>0.08). When modeling the American subprime crisis from 2008 to 2009 as out-of-sample data, the maximum volatilities that the GJR-GARCH and GARCH model provided were much less than the maximum "true volatility". Therefore, these GARCH models are not quite suitable for risk management applications as they greatly underestimate the volatility during periods of extremely high volatility (>0.08).

In this paper, we modified the GARCH model, while using the GJR-GARCH as inspiration, and proposed a new model named GARCH-S. We used the Nikkei 225 and S&P 500 as the two asset inputs required by the GARCH-S model. The GARCH-S makes use of a secondary market (S&P 500) to increase the volatility forecast of the target market (Nikkei 225). This is possible when the returns of the secondary market at time $t-1$ is correlated (>0.20) with the returns of the target market at time t. Even though we developed a MLE program to estimate the model parameters, the Gibbs Sampling method was used in general to estimate the parameters of the models.

The GARCH-S is found to effective in modeling the "upper range" of "true volatility". It greatly reduces underestimate errors as compared to the GARCH, GJR-GARCH and EGARCH models. Under the several assumptions we made, the GARCH-S was found to be ranked second best in forecasting accuracy while effectively capturing the peaks of high volatility better than other GARCH type models. Therefore, the GARCH-S model is an excellent candidate for risk management purposes as it greatly reduces the underestimate errors while still providing competitive forecast accuracy against other GARCH type models.

Key Word(5 words)

GARCH, Nikkei 225, S&P 500, Volatility, Modeling

Acknowledgments

I would like to thank Professor Nishimura Hidekazu for the support and direction that he has given me. He has granted me much freedom in choosing and developing this thesis. I have gained a tremendous amount of knowledge from this thesis. I am also grateful for the time I had spent in Nishimura Laboratory and at the Graduate School of System Design and Management. It will forever be an imperative part of my life. Special thanks to Professor Takano Kenichi, and Professor Hibiya Taketoshi for continuously providing me with valuable inputs.

Lastly, I would like to thank my mom and my family, for the sacrifices they have made to put me through graduate school. None of these would be possible without them.

Contents

1 Introduction

Financial volatility modeling has become an increasingly active research area. After the financial crisis of 2008 and 2009, people around the world are becoming more aware of the financial markets. The Japanese financial market has certainly been through some tough time, and financial volatility modeling would be an excellent tool that we can use to understand the Japanese financial market and even apply it to risk management. However, modeling the volatility of the financial markets is not a simple task.

1.1 Background of Research

Before we even begin to research about financial market volatility, we need to understand volatility. Different people often have different ways of interpreting volatility. But what is volatility exactly? Can a statistical model describe its features?

1.1.1 What is Volatility?

In the field of volatility modeling, there is a general consensus of defining volatility. In the financial markets, we are concerned with the spread of the returns of the asset. Mathematically, volatility can be measured as standard deviation,

$$\hat{\sigma} = \sqrt{\frac{1}{T-1}\sum_{t=1}^{T}(r_t - \mu)^2}\ , \tag{1.1}$$

where r_t is defined as the return on the asset at time t, and μ is the average return over the time period T.

We could also use the variance, σ^2, as a measure of volatility. However, using variance would be much less stable for computer simulations and forecast evaluations. Comparatively, the standard deviation does not fluctuate as much as variance. Another advantage of using the standard deviation is that it has the same units as the mean. Thus, it would be easier to interpret our simulations and forecast using standard deviation.

Volatility can be seen as a measure of the financial market stability. When the volatility is high, the financial market becomes unstable, and when volatility is low, the financial market becomes stable. However, it is important to note that although volatility is related to risk, it is not the same as risk. This is because risk is associated with undesirable outcomes but volatility is uncertainty that could potentially result in a positive outcome. Huge volatilities could cause the financial markets to move either upwards or downwards. This fact is often overlooked. Take the Sharpe ratio [Sharpe (1996)] for example. The Sharpe ratio measures the performance of an investment by comparing the mean return in relation to its 'risk' proxy by its volatility.

$$\text{Sharpe Ratio} = \frac{\begin{pmatrix}\text{Average} \\ \text{return, } \mu\end{pmatrix} - \begin{pmatrix}\text{Risk-free interest} \\ \text{rate, e.g. U.S. Treasury Rate}\end{pmatrix}}{\text{Standard deviation of returns, } \sigma}. \tag{1.2}$$

However, it is interesting to note that periods of high volatility in the financial markets usually corresponded with periods of greater risk in the financial markets. This is due to the intrinsic nature of the financial markets where you only see the financial market "crash down" and not "crash up". This is one of the reasons volatility can be used as a tool to manage financial market risk, especially during periods of extremely high volatility such as the financial crisis of 2008 and 2009.

Although volatility is not the only determinant of asset return distribution, it is an imperative input to many different financial applications. Understanding financial market volatility will ultimately lead to us to better understand the returns of the financial markets, and potentially better manage financial market risk.

The main purpose of financial volatility modeling is to be able to apply the derived models on the financial markets. Financial markets are entities that allow people from all over the world to trade financial securities, commodities, and other various goods. There is a plethora of financial markets around the world. For our research, we are mainly concerned with the Japanese financial stock market, and its relationship with other financial stock markets around the world. For the first part of research, we are concerned with the Nikkei 225 index and the S&P 500 index.

1.1.2 Nikkei 225

The Nikkei 225 is a stock market index based on prominent Japanese companies that are publicly traded on the Tokyo Stock Exchange (TSE). These companies include Toyota Corporation, Nomura Holding and Mitsubishi Corporation. It is designed to reflect the overall trend of the Japanese stock market. Therefore, it could be used to reflect the state of the Japanese economy. The only downside of the using the Nikkei 225 index to reflect the state of the Japanese economy is that it might not reflect the economical situation of small and medium sized companies. From figure 1.1, we can see that the Japanese financial market is really quite unique and different when compared to other developed countries' financial markets. Just before the infamous Japanese asset bubble "burst", the Nikkei 225 reached its peak of around 38512 on 20th December

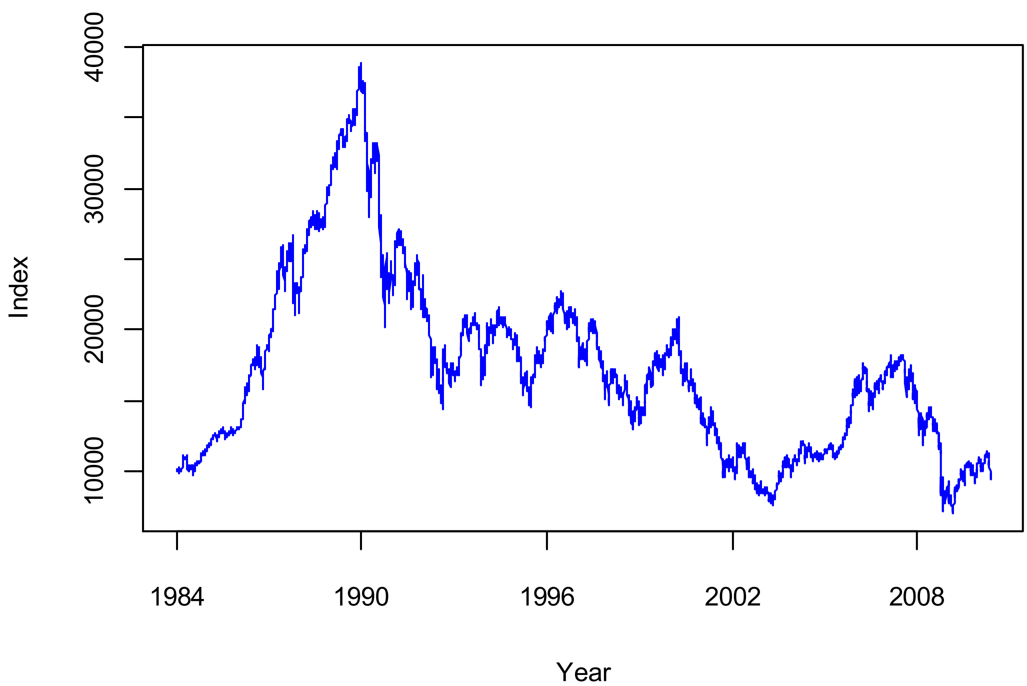

Figure 1.1 Nikkei 225 Index from 4th January 1984 to 28th May 2010

1989. Thereafter, the Nikkei 225 and the Japanese economy went drastically downhill, losing more than half its value in just a period of 3 years. The Nikkei 225 ended up fluctuating around 14309 on 18th August 1992. The period after the bubble "burst" corresponded to a period of extremely high volatility.

Other periods of extremely high volatility occurred between 2000 and 2002 as a result of the Information Technology (IT) bubble "burst"; and between 2008 and 2009 as a result of the American subprime mortgage crisis. The Japanese financial market has faced much turbulence in the past and it would definitely be an interesting financial market to research on.

Volatility describes the spread of an asset. In other words, it describes how an asset changes over time. By just using the Nikkei 225 index, we are unable to extract any volatility information. In order to model the volatility of the financial markets, we need to use some form of information that is able to describe the spread of an asset. In financial volatility modeling, the general consensus is to use the logarithm return of an asset. The logarithm allows us to assume that the distribution of the asset is continuously compounded, and it also allows us to model the returns with time additive property. The logarithm return is defined as,

$$y_t = \ln\left(\frac{x_t}{x_{t-1}}\right), \tag{1.3}$$

where y_t is defined as the return at time t, and x_t is defined as the index value at time t.

Using equation 1.3, we could then model the return of the Nikkei 225 index. From figure 1.2, we can observe that the historical daily returns fluctuate greatly around 1987, and the year 2008, resulting in extremely volatile periods. The volatile period in 1987 was caused by the financial event known as the "Black Friday". The 2008 volatile period was caused by the American subprime mortgage crisis. By observing figure 1.2, we can roughly get an image of what the volatility graph would look like for the Nikkei 225.

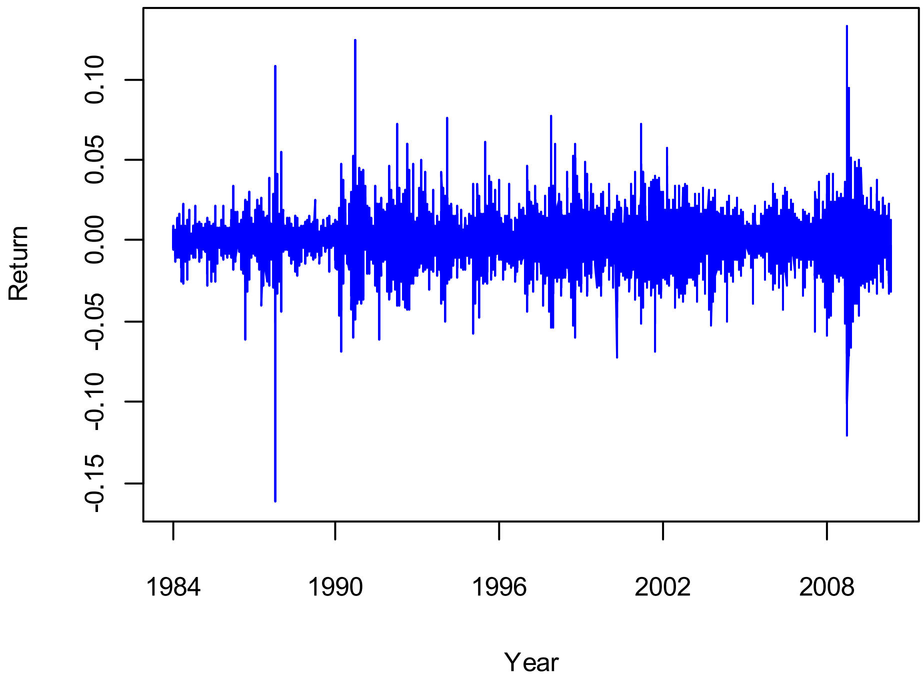

Figure 1.2 Historical daily returns for the Nikkei 225 from 5th Jan 1984 to 28th May 2010

1.1.3 Standard and Poor 500

The Standard and Poor 500 (S&P 500) is the American counterpart of the Nikkei 225. The S&P 500 is a stock market index that is based on prominent American companies that are publicly traded either on the New York Stock Exchange or the NASDAQ. Some of these companies include Berkshire Hathaway, ExxonMobil, and Microsoft Corporation. The S&P 500 is designed to reflect the overall trend of the American stock market. Again, the downside of using the S&P 500 to reflect the state of the American economy is that it might not reflect the economical situation of small and middle sized companies.

When comparing the Nikkei 225 index to the S&P 500 index, we immediately notice a huge difference between them. During 1984 to 2010, the Nikkei 225 generally displayed

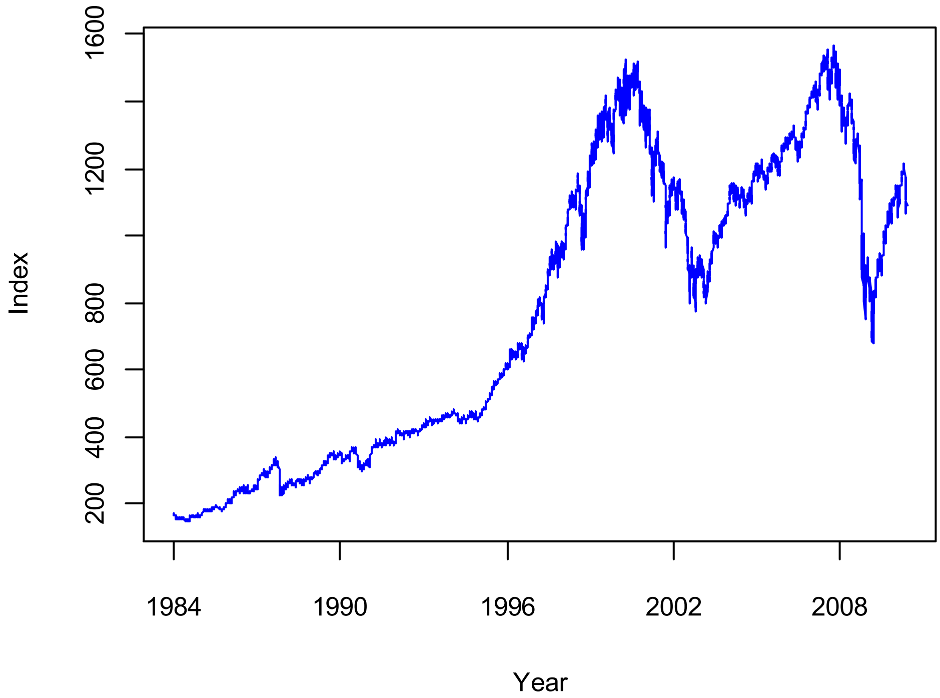

Figure 1.3 Standard and Poor 500 Index from 4th January 1984 to 28th May 2010

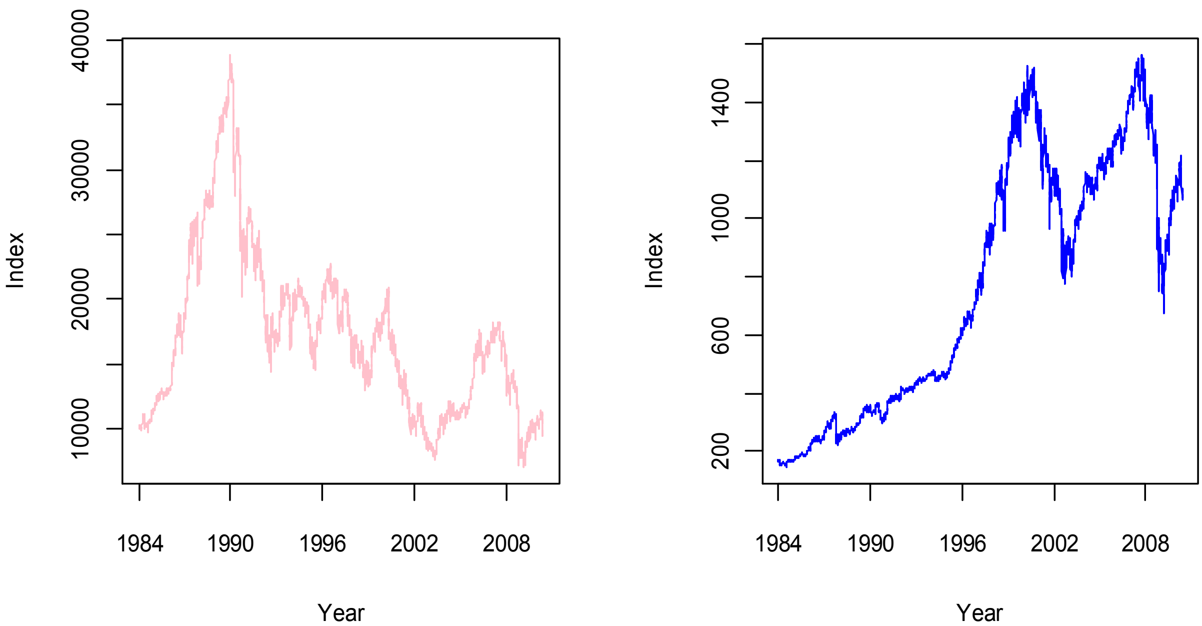

Figure 1.4 Comparing the Nikkei 225 (left, pink) and the S&P 500 (right, blue) from January 1984 to May 2008

a downward trend. However, the S&P 500 generally displayed an upward trend. This is mainly due to the Japanese asset bubble "burst". The term "lost decade" is used to symbolize the 1990s period when the economy of Japan went stagnant.

Although the Nikkei 225 and the S&P 500 index had displayed very different trends; recently, they have begun to display rather similar trends with each other. For example, take the time period from 2000 onwards, the trend is almost similar with one another. They both experienced the double dip from the IT bubble "burst" and the American subprime mortgage crisis. From figure 1.4, we can see that in recent years, the Nikkei 225 and the S&P 500 index are rather strongly correlated with one another.

However, as stated in section 1.1.2, we need to use the returns of the S&P 500 and not the index in order to model volatility. The graph of the historical daily returns for the S&P 500 is displayed in figure 1.5, and a side by side comparison is displayed in figure 1.6.

When comparing the correlation of the Nikkei 225 and the S&P 500 daily returns, we cannot be sure without a numerical value. However, by looking at the graphs of the daily returns, we can see that there is definitely a correlation between the Nikkei 225 and the S&P 500 daily returns, especially during the 2000s (recent times). By observing the historical daily returns of the S&P 500, we can roughly image what the volatility will look like.

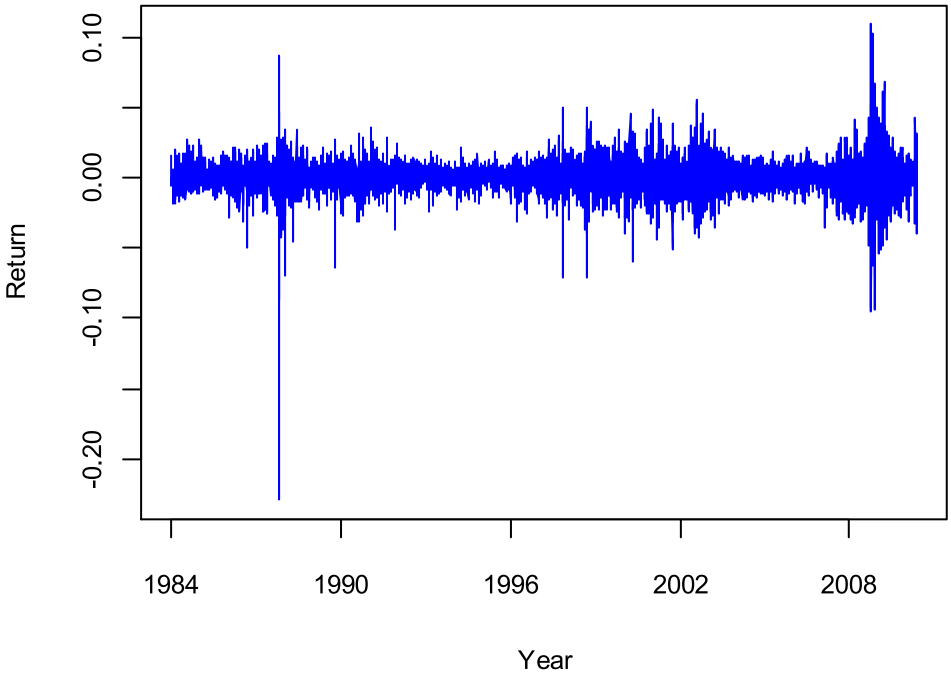

Figure 1.5 Historical daily returns for the S&P 500 from 5th Jan 1984 to 28th May 2010

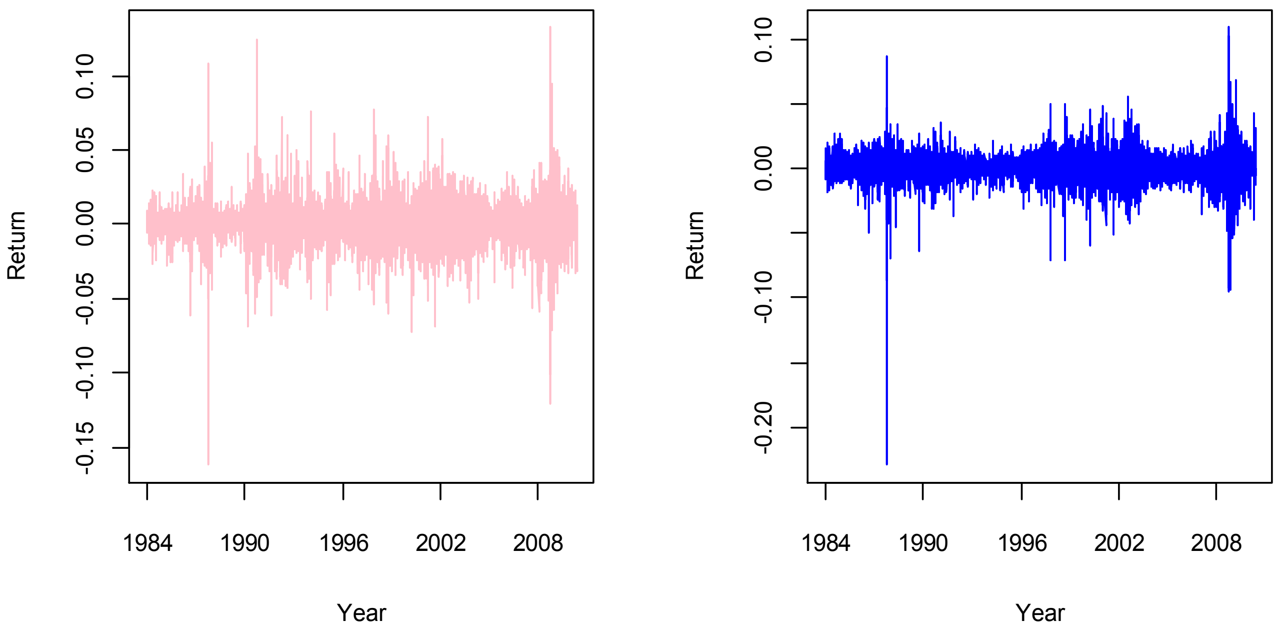

Figure 1.6 Comparing the historical daily returns for the Nikkei 225 (left, pink), and the S&P 500 (right, blue) from 5th Jan 1984 to 28th May 2010

1.1.4 Facts about Financial Markets

Understanding the probability density distributions is imperative in any statistical volatility modeling. This is because for both frequentist and Bayesian statistics, we need to assume a probability distribution model for the return values in order to model the parameters. Depending on which probability distribution models we choose for modeling the returns, our accuracy of the estimated parameters might be different. When the Autoregressive Conditional Heteroskedascity (ARCH) models were first derived, the probability distributions of the returns were commonly based on the normal distribution. However, researchers began to discover that other probability distributions produce better estimates. For example, the stable Paretian innovations might provide better estimated parameters than the Normal distribution in financial volatility modeling [Mittnk, Rachev (1999)]. Another example is that the stable Paretian innovations might fit the US, German and Portuguese main stock indexes better than the Normal distribution, and slightly better than the t-distribution [Curto, Pinto (2007)].

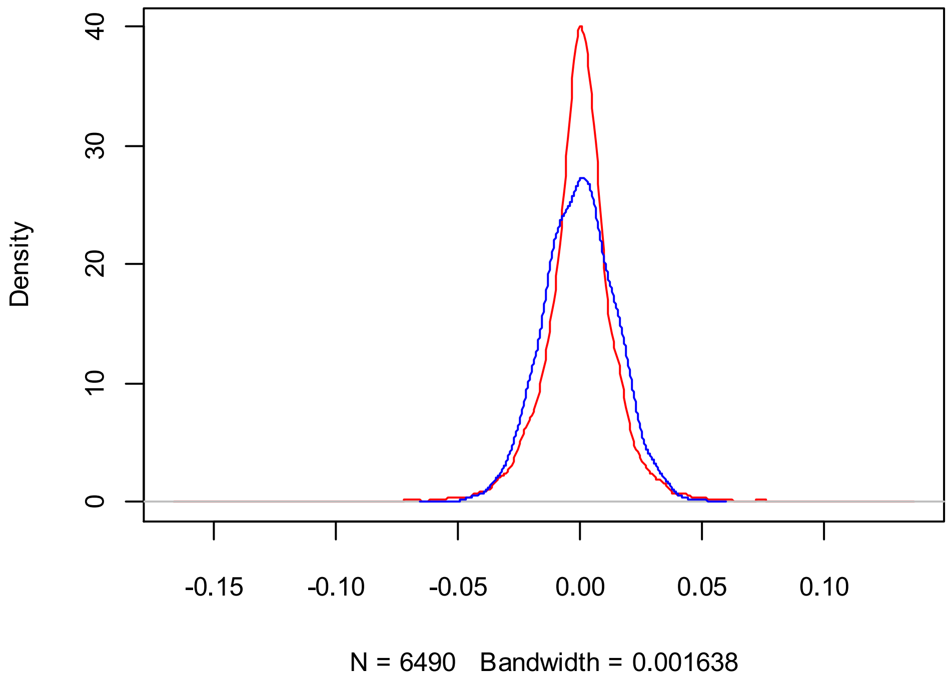

Figure 1.7 The returns distribution of the Nikkei 225 (red) as compared to the normal random variable simulation (blue)

To see the reasoning behind this, let us take a look at the probability distribution patterns for both the Nikkei 225 and the S&P 500. Using data from January 1984 to May 2010, we plotted the returns distribution of the Nikkei 225 against a simulated Normal distribution and then plotted the returns distribution of the S&P 500 against a simulated Normal distribution.

From the figures 1.7 and 1.8, we can see that although the index values are very different, the density distributions of both the Nikkei 225 and the S&P 500 generally display the same pattern. This pattern is also observed from other financial markets [Poon (2005)].

The observation made from the financial returns distribution is that the returns distributions have longer "tails" and much sharper peaks than the normal distribution. Also, the tails of the financial returns distributions are actually heavier (larger) than the normal distribution.

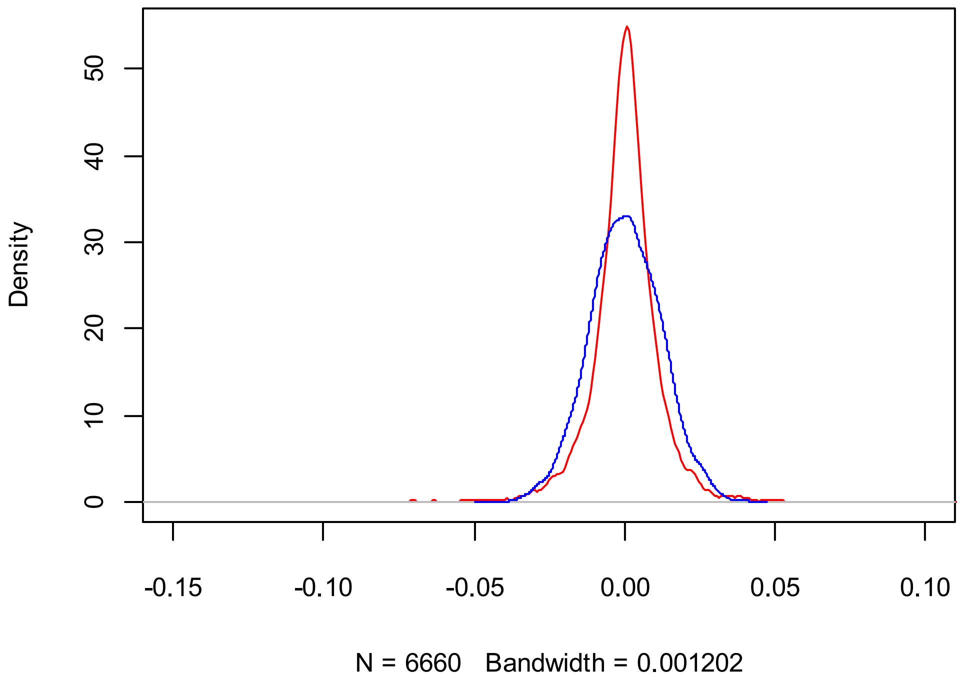

Figure 1.8 The returns distribution of the S&P 500 (red) as compared to the normal random variable simulation (blue)

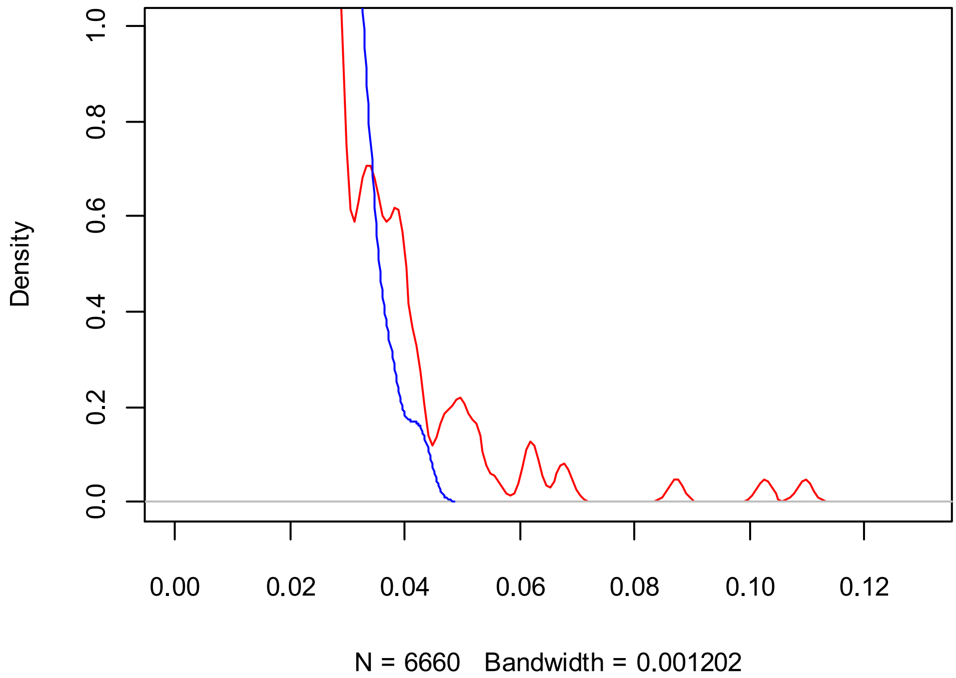

Figure 1.9 The enlarged version of the tail of the returns distribution of the S&P 500 (red) as compared to the normal random variable simulation (blue)

These observations prove that the returns of financial assets are not normal. They are in fact closer to the t-distribution. One of the ongoing research topics in financial volatility modeling is researching on which distributions fit better with which financial asset returns. However, using the Normal distribution is certainly not wrong. It is much easier and quicker to implement as compared to other distributions. The only drawback is that it might be a little less accurate. In thesis, only the normal distribution will be used, as our focus is not on the model's forecast accuracy. Therefore, we are not extremely concerned with errors.

There is another important implicit meaning from figures 1.7 and 1.8. This implicit meaning is that the Japanese and American stock markets will usually fluctuate within a smaller range than the normal distribution. However, the financial market is also extremely capable of fluctuating eccentrically beyond the normal distribution. In fact, it could fluctuate to an unimaginable extent, as it has happened during the subprime mortgage crisis.

For the time period from January 1984 to May 2010, the skewness of the Nikkei 225 is -0.22, and the kurtosis is 8.39. While the skewness of the S&P 500 is -1.35 and the kurtosis is 29.96. [Refer to the Appendix for more information on Skewness and Kurtosis.] This means that the return values of the S&P 500 are more likely to be distributed above the mean as compared to the Nikkei 225, a larger growth pattern. Also, the peak of the S&P 500 is much higher than the Nikkei 225. This can be interpreted that the IT bubble asset "burst" and the American subprime mortgage had a stronger effect on the American financial system as compared to the Japanese financial system. These skewness and kurtosis statistics are very sensitive to outliers. If we remove the extreme outliers, such as the 1987 October crash, we would be able to greatly reduce the value of the kurtosis [Poon (2005)].

It is important to note that one general statistical model or returns distribution does not apply to all the financial markets. Different financial markets require different statistical models and returns distributions. What might work well for a certain financial market might not work as well for another financial market. It is up to the researcher to decide which combination of models and returns distribution will work well for their research.

1.1.5 Problem Statement

Ever since Robert Engle introduced the Autoregressive Conditional Heteroskedascity (ARCH) model in 1982, financial time series volatility modeling has become an increasingly important research area [Engle (1982)]. This importance was recognized by the world when Robert Engle was awarded the 2003 Memorial Nobel Prize in Economics.

Robert Engle's model was designed to capture and make use of the natural characteristic of financial market volatility. The natural characteristic is that volatility tends to cluster together. This means that periods of high volatility and low volatility frequently occur in clusters separating one another. We shall discuss more about these time series models in chapter 3.

Robert Engle's ARCH model was the beginning of a large class of time series models. Ever since 1982, numerous other ARCH classes of models were developed. One of the most famous models is the Generalized Autoregressive Conditional Heteroskedascity (GARCH) model [Bollerslev (1986)]. This model was derived in 1986 by Tim Bollerslev, who was a student of Robert Engle.

Following Tim Bollerslev's GARCH model, the GJR-GARCH (Glosten-Jagnnathan -Runkle) model was derived in 1993 [Glosten, Jagnnathan, Runkle (1993)], and several other ARCH class models were also developed in the 1990s and 2000s such as the Exponential GARCH (EGARCH) model. However, none of these univariate models made use of the fact that different financial market volatilities are correlated with one another. For example, if the American financial market was to experience a period of high volatility, one can assume that the Japanese financial market will also experience a period of high volatility. Even though multivariate GARCH models use the innovation of a secondary market input to forecast the volatility of the first input, multivariate GARCH models focus on studying the correlation between financial markets, such as spillover effects. The GARCH-S model will be more focused on reducing underestimated errors. It is also much simpler to implement than a multivariate GARCH model.

Figure 1.10 shows that the Nikkei 225 and S&P 500 display similar trends from 2 March 2009 to 28 May 2010. It is this similarity that we want to exploit. We believe that by using the Nikkei 225 and S&P 500's correlation, we could provide a better representation of the volatility for risk management purposes as the secondary market acts like a foreign term. This foreign term will add information, and possibly predictive power that does not exists in the primary market. This is an abnormal idea since we are actually incorporating a secondary model's returns into the volatility model in order to help provide more information in modeling the primary model's volatility.

So far, all GARCH type models tend to underestimate periods of high volatility (>0.06)

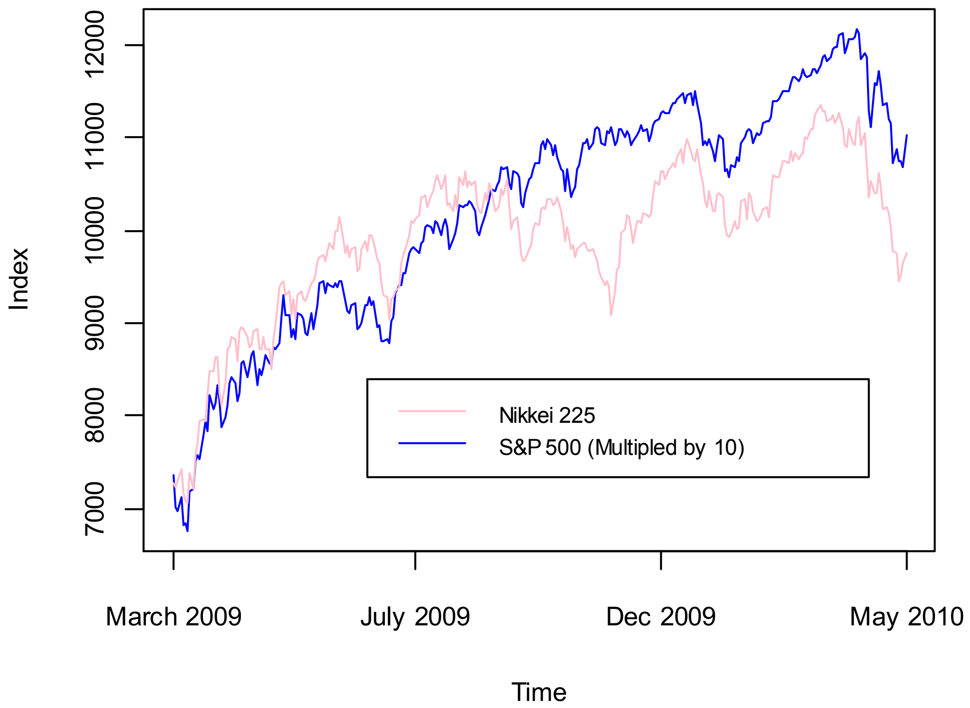

Figure 1.10 Comparing the Nikkei 225 and the S&P 500 Indexes

by a significant amount. In risk management and most applications, a user would be much more concerned about underestimating volatility rather than overestimating volatility. This is especially true during periods of high volatility. This is because we underestimate financial volatility, the consequences could be devastating such as an unexpected lose amount. However, when we overestimate financial volatility, the consequences would not be as devastating as the lose amount is within our expectations.

Our proposed GARCH-S volatility model shall be tailored towards the risk management applications by providing smaller underestimation errors, especially during periods of high volatility (>0.06). The consequence is that the GARCH-S model might overestimate errors. Since the "true volatility" usually occurs at a low range, there are much more overestimate errors than underestimation errors. Therefore, we must find a balance between reducing underestimation errors while still providing competitive forecast accuracies.

The second part of our research will be to focus on practical applications of the GARCH-S. We will use the Vector Autoregressive Model (VAR) to study the relationship of the returns between the Nikkei 225 and the FTSE (Financial Times Stock Exchange) China 25. Even though previous research papers have provided literature on the relationship of the Nikkei 225 as compared to other markets [Beikros, Georgoutsos (2006)], this thesis will focus on the effects of a simulated financial crisis on the Nikkei 225. For example, what if China has an asset bubble that "burst"? To help manage the risk associated with a financial crisis, we will implement the GARCH-S model into the Value-at-Risk (VaR).

1.2 Goal of Research

There are two primary goals of this research:

1. The first goal is to successfully derive and implement the proposed GARCH-S model. The proposed GARCH-S model would make use of a strongly correlated (>0.70 correlation in returns) secondary market's innovations to improve volatility modeling the Nikkei 225 as compared to the GARCH, GJR-GARCH and EGARCH models for risk management purposes. The GARCH-S would reduce overall underestimate errors especially during periods of high volatility.
2. The second goal is to demonstrate a practical application of the GARCH-S. We will simulate the effect of a financial crisis on the Nikkei 225 and use the GARCH-S together with the VaR as part of risk management methods.

The GARCH-S model is still in development stages and its methodology might seem to be primitive and naive at times. The GARCH-S model also cannot be applied to markets that have weak or no correlation with one another. Therefore, this model cannot be generalized to all markets around the world. However, the GARCH-S model is applicable when using the S&P 500's volatility at time t to model the Nikkei 225's volatility at time $t+1$. This is because the S&P 500 and Nikkei 225 have a strong correlation, especially during and after the American subprime mortgage crisis (0.89 for the time period from 2 March 2009 to 28 May 2010).

It is important to note that for different financial markets, different GARCH models might work better. Not one GARCH model dominates the field of time series volatility modeling [Poon (2005)]. As such, the focus of this thesis is not about forecasting accuracy, but rather on whether or not the GARCH-S model provides new information on the target market's volatility as compared to other GARCH type models. In our thesis, we are mostly concerned with the Japanese financial market, more specifically with the Nikkei 225.

The secondary goal of this research is:

1. To estimate the parameters of the GJR-GARCH, EGARCH, and the GARCH-S model through Markov Chain Monte Carlo (MCMC) Bayesian Statistics and the classical Maximum Likelihood Estimation (MLE) method. More specifically, the Gibbs Sampling method will be implemented.

In Statistics, there are two schools of thought, classical inference and Bayesian inference. In Japan, there is not so much literature focusing on Bayesian Inference. Therefore, we believe this secondary goal of using the Gibbs Sampling method to estimate the parameters of the models will be an important academia contribution to Japan.

The main reason we chose to use the Gibbs Sampling method is that it is extremely simple to implement and estimate parameters for any models including GJR-GARCH, GARCH-S, and EGARCH. Even though we have written the function for the MLE estimation method, this paper will be based on our estimates from the Gibbs Sampling method. Another advantage is that the Deviance Information Criterion (DIC) [section 2.1.1] could be easily calculated from using the Gibbs Sampling method. Since the development of modern computers, an increasing number of researchers, professors have begun to prefer using Bayesian statistics.

Financial market volatility research is still growing in all aspects. Recent practical applications include the Chicago Board of Exchange (CBOE) futures trading on a volatility index, introduced in 2004. Another practical application that had been used for awhile is the volatility swap contract.

1.3 Structure of Thesis

In chapter 2, we shall discuss model selection, and forecasting evaluation techniques. The Deviance Information Criterion (DIC) method is introduced and would be used for model selection. Error statistics techniques are introduced and would be used in evaluating our proposed GARCH-S model. Subsequently, the Diebold and Mariano (DM) test is introduced and would be used to check if these error statistics are significant or not.

Chapter 3 focuses on developing our proposed GARCH-S time series volatility model. Data preparation is an extremely important part of our research and it is also explained in chapter 3. In order to compare the GARCH-S against other GARCH type models, parameters for other ARCH/ GARCH type models are also estimated. We will write programs for both the Gibbs Sampling method and the Maximum Likelihood Estimation (MLE) method to estimate our parameters. However, our focus will be on the Gibbs Sampling method, the MLE method is mainly for comparison purposes. Subsequently, the accuracy of our proposed GARCH-S model will be compared and evaluated against other GARCH type models (GARCH, GJR-GARCH, EGARCH) using techniques described in chapter 2.

Chapter 4 focuses on evaluating our proposed GARCH-S model as compared to other GARCH type models. For the first part of chapter 4, we will focus on using the techniques developed in chapter 2 to analyze and compare our models. The second part of chapter 4 will focus on evaluating the GARCH-S model's effectiveness of providing new information in modeling the Nikkei 225's volatility as compared to other GARCH type models.

In chapter 5, we will study the Japanese financial market's returns and its relationship to the Chinese financial market. To study the returns' relationship, we will use the VAR model. Subsequently, we will create a hypothetical financial crisis in the Chinese financial market. This hypothetical financial crisis will be simulated using a stochastic drift equation. The corresponding hypothetical return values for the Nikkei 225 will be simulated using the VAR model and a stochastic drift equation. Lastly, a practical application of the GARCH-S will be demonstrated using it as an input in the VaR equation.

2 Volatility Modeling Evaluation

When modifying parameters to a statistical model, we first need to confirm whether our modifications are justifiable. These means that we need to deduce whether these parameters actually made a worthwhile contribute to the model or are these parameters just "garbage parameters" that would provide no new information in augmenting the statistical model. The parameter evaluation techniques we used to evaluate our models will be discussed in section 2.1.

After we have obtained a satisfactory model, we then need to compare and evaluate the model against other similar models. To evaluate our model, it is important that we not only evaluate the in-sample forecasting performance but also the more realistic out-of-sample forecasting performance. The model evaluation techniques we used to evaluate our models will be discussed in section 2.2.

2.1 Parameter Evaluation Techniques

Parameter evaluation techniques allow us to measure and judge the extent of the goodness of fit for the model. By plotting the autocorrelation functions of the residuals for an ARCH/ GARCH fit, we can judge whether or not the model was sufficient enough to remove any disruptions such as seasonal effects. If the residuals of the ARCH/ GARCH fit do not correspond to a white noise effect, this means that our ARCH/ GARCH model selection might not be a sufficiently suitable. [Autocorrelation functions are described in section 3.3.3] If the ARCH/ GARCH model selection is not sufficiently suitable, we might need to use a higher order GARCH model.

For most financial volatility cases, including our research, the GARCH(1,1) is sufficient. In our research, we made additional parameter modifications to our GARCH(1,1) model. After we obtained these estimated parameter values, we should check if they are statistically significant or not. If they are not statistically significant, we must then conclude that these parameters add no new information to the model are all not needed. If they are statistically significant, we could then carry on to the next step of evaluating our parameters. To check whether parameters are statistically significant or not, we need to use the t-values. In our paper, the t-values are defined as,

$$t\text{ value}=\frac{\text{parameter value}}{\text{standard deviation}}. \tag{2.1}$$

These t-values correspond to the null hypothesis H_0: the parameter has a value of 0, and the alternative hypothesis H_a: the parameter has a non-zero value. A high t-value will correspond to a low p-value. Typically, we reject the null hypothesis if we obtain a p-value of less than 0.05, which correspond to a t-value of more than 1.96 in our case. Therefore, if the t-value is more than 1.96, there is evidence to reject the null hypothesis and the parameter is statistically significant. Otherwise, we have no evidence to reject the null hypothesis that the parameter has a value of 0, meaning the parameter is statistically insignificant.

Besides checking the statistical significance of the parameters, we also used the Deviance Information Criterion (DIC) to justify our additional parameters. The DIC is a hierarchical model generalized from the Akaike Information Criterion (AIC), and the Schwarz-Bayesian Information Criterion (BIC). The idea is that if our additional parameters do make a worthwhile contribution to the model, then the DIC should decrease.

2.1.1 Deviance Information Criterion (DIC)

Although DIC is the main selection criteria we used for our research, its' predecessors the AIC and BIC are also important model selection tools we could use. The AIC [Akaike (1973)] is defined as,

$$\text{AIC}=\frac{-2}{T}\ln(\text{likelihood})+\frac{2}{T}\times(\text{number of parameters}), \tag{2.2}$$

where T is the sample size. The likelihood function is evaluated at the maximum-likelihood estimates. The first term of equation 2.2 measures the goodness of fit of the model. The better the fit of the model, the more negative the first term becomes. The second is called the penalty function of the AIC because it penalizes a candidate model by the number of parameters being used. Therefore, when we add parameters to an existing model, the AIC of the new model should be lower than the existing model's AIC. Another commonly used criterion function to evaluate model selection is the BIC, which is defined as,

$$\text{BIC} = \frac{-2}{T}\ln(\text{likelihood}) + \frac{\text{number of parameters} \times \ln(T)}{T}. \tag{2.3}$$

The BIC is the same as the AIC except that the penalty for each additional parameter used is $\ln(T)$ instead of 2. Therefore, in general, the BIC penalizes each additional parameter more severely than the AIC. The larger the sample size, the greater the penalty for the BIC.

[Dempster (1974)] first proposed to examine the posterior distribution of the classical deviance defined by,

$$D(\theta) = -2\ln f(y\,|\,\theta) + A, \tag{2.4}$$

where y is the data, θ are the unknown parameters of the model and $f(y\,|\,\theta)$ is the likelihood function, which is the conditional joint probability density function of the observations given the unknown parameters. A is a constant that will cancel out in calculations when comparing different models. Therefore, this term does not need to be known. Based on the posterior distribution of $D(\theta)$, the DIC consists of two terms,

$$\text{DIC} = \bar{D} + p_D. \tag{2.5}$$

In equation 2.5, the first term is a measure of the goodness of fit, and the second is a penalty term for increasing model complexity. This means that the more parameters that are added to the model, the larger will p_D become. The first term in equation 2.5, a Bayesian measure of model fit, is defined as the posterior expectation of equation 2.4,

$$\bar{D} = E_{\theta|y}[D(\theta)] = E_{\theta|y}[-2\ln f(y\,|\,\theta)]. \tag{2.6}$$

The better the goodness of the fit, the larger the value of $f(y\,|\,\theta)$ becomes. Since $\bar{D}$ is defined as the posterior expectation of -2× $f(y\,|\,\theta)$, the better the goodness of the fit, the smaller will $\bar{D}$ become.

The second term of the DIC in equation 2.5, p_D, measures the complexity of the model by the effective number of parameters. This term is defined as the difference between the posterior mean of the deviance and the deviance evaluated at the posterior mean $\bar{\theta}$ of the parameters,

$$\begin{aligned} p_D &= \bar{D} - D(\bar{\theta}) = E_{\theta|y}[D(\theta)] - D(E_{\theta|y}[\theta]) \\ &= E_{\theta|y}[-2\ln f(y\,|\,\theta)] + 2\ln f(y\,|\,\bar{\theta}). \end{aligned} \tag{2.7}$$

By substituting equation 2.7 into equation 2.5, we could also redefine equation 2.5 as,

$$\mathrm{DIC} = D(\bar{\theta}) + 2p_D. \tag{2.8}$$

Therefore, the DIC could be interpreted as the classical deviance measure of fit using posterior parameters plus a measure of complexity. When the number of observations grows with respect to the number of parameters p, and the prior is nonhierarchical and completely specified, then the AIC will be $D(\hat{\theta}) + 2p$, where $\hat{\theta}$ denotes the maximum likelihood estimate. This is the same as equation 2.8 except the posterior mean $\bar{\theta}$ is substituted by the ML estimate $\hat{\theta}$. Therefore, the DIC can be seen as a generalization of the AIC, and could also be compared to the BIC of equation 2.3. The idea behind DIC is that models with lower DIC should be preferred over models with higher DIC. It is important to note that the absolute value of DIC does not matter. What is important in model comparison is the difference in DIC values. Usually, a difference of more than 10 signifies that the model with lower DIC is definitely better.

One of the advantages that DIC has over AIC, and BIC is that it can be easily calculated from the samples generated by a Markov Chain Monte Carlo Simulation (MCMC). In our paper, we used the Gibbs Sampling method, which is a Bayesian MCMC, to estimate the parameters of our models. The estimate of $\bar{D}$ can be calculated from the MCMC output by monitoring $D(\theta)$ and then taking the sample mean of the simulated values of $D(\theta)$. The effective number of parameters p_D can be obtained by evaluating $D(\theta)$ at the sample average of the simulated values of θ and subtracting this plug-in estimate of the deviance from the estimate of $\bar{D}$.

For our research, we implemented the WinBUGS (Bayesian inference Using Gibbs Sampling) and JAGS (Just Another Gibbs Sampler) statistical packages through the R statistical program. The WinBUGS and JAGS statistical packages provide the DIC values with the results of the estimated parameters. Therefore, these Gibbs Sampling statistical packages are rather convenient to use.

In summary, we modify the GARCH model by adding in new parameters and then check whether the DIC values decreased or remained the same. If the DIC values did indeed decrease by a substantial amount, then we could justify that our new modifications to the model was worthwhile.

2.2 Model Evaluation Techniques

After simulating the parameters of our modified GARCH, GARCH-S models, we would then need to compare the forecasting results of these models with other GARCH type models. The models we compared against each other were the GARCH(1,1), GJR-GARCH(1,1), our modified GARCH(1,1), our proposed GARCH-S(1,1), and the EGARCH(1,1) model. Our proposed GARCH-S(1,1) is a descendant of the GARCH(1,1) model and the GJR-GARCH(1,1) model. Therefore a comparison between these models would be necessary to show that the GARCH-S(1,1) model does provide an intriguing, and new perspective to the GARCH class of volatility modeling.

Comparing forecasting performance of competing models is one of the most important aspects of any forecasting exercise. However, in contrasts to the efforts made in the construction of volatility models and forecasts, little attention has been paid to forecast evaluation. The forecast error can be defined as,

$$\varepsilon_t = \hat{X}_t - X_t, \tag{2.9}$$

where $\hat{X}_t$ is defined as the predicted variable by using our models, and X_t is the actual outcome. X_t should be σ_t, and not σ_t^2. Once a shock has entered the system, the merit of the volatility model depends on how well it captures these effects in predicting the volatility of subsequent days. Conditional variance σ_t^2 will give too much weight to errors caused by these shocks, especially the large ones, distorting the less extreme forecasts where the models are to be assessed. While using $\ln \sigma_t$ as X_t to rescale the size of the forecast errors might be one step too far. This is because the magnitude of the error directly impacts option pricing, risk management, and investment decisions. Distorting the magnitude of the errors might cause a decision maker to misjudge since a decision maker is more likely to be more risk-averse towards larger errors. Therefore, we should use σ_t as X_t, and not σ_t^2 nor $\ln \sigma_t^2$.

2.2.1 Error Statistics

In practice, often simple evaluation measures suggested by statisticians are used. The evaluation measures used in this paper include the Mean Square Error (MSE),

$$\frac{1}{N}\sum_{t=1}^{N}\varepsilon_t = \frac{1}{N}\sum_{t=1}^{N}(\hat{\sigma}_t - \sigma_t)^2 , \tag{2.10}$$

and the Root Mean Square Error (RMSE),

$$\sqrt{\frac{1}{N}\sum_{t=1}^{N}\varepsilon_t} = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(\hat{\sigma}_t - \sigma_t)^2} , \tag{2.11}$$

and the Mean Absolute Error (MAE),

$$\frac{1}{N}\sum_{t=1}^{N}|\varepsilon_t| = \frac{1}{N}\sum_{t=1}^{N}|\hat{\sigma}_t - \sigma_t|, \tag{2.12}$$

and the Mean Absolute Percent Error (MAPE),

$$\frac{1}{N}\sum_{t=1}^{N}\frac{|\varepsilon_t|}{\sigma_t} = \frac{1}{N}\sum_{t=1}^{N}\frac{|\hat{\sigma}_t - \sigma_t|}{\sigma_t}. \tag{2.13}$$

These are the four error statistics we will use in our paper. Nowadays, it is standard to use out-of-sample forecasting results to compare the errors of the models. In our paper, we will compare both the in-sample and out-of-sample errors, but our main focus would be on the out-of-sample errors.

2.2.2 Linex loss function

As stated in section 1.1.5, for risk management and other purposes, it is better to overestimate volatility rather than underestimate the volatility. In order to evaluate which models give smaller underestimate errors, we need to implement an asymmetric loss function. The linex asymmetric loss function in which positive errors are weighted differently from negative errors is defined as,

$$\text{linex} = \frac{1}{N}\sum_{t=1}^{N}[\exp\{-a(\hat{\sigma}_t - \sigma_t)\} + a(\hat{\sigma}_t - \sigma_t) - 1]. \tag{2.14}$$

The choice of parameter a is subjective. The larger positive a becomes, the larger the penalty for underestimating errors. This is because if $a > 0$, the lose function is approximately linear for overprediction and exponential for underprediction.

Given that most investors would treat gains and losses the differently, the use of an asymmetric lose function would be favored. However, they are not commonly used in literature.

2.2.3 Diebold and Mariano (DM) Test for Predictive Accuracy

If the forecasting error distribution of one model dominates the forecasting error distribution of another model, then the comparison is straightforward [Granger (1999)]. However, this is rarely the case in practice. In practice, most of the comparisons are based on the error statistics described in section 2.2.1. It is important to note that these error statistics are themselves subjected to error and noise. Therefore, if the difference is not significant enough, we cannot conclude that model B is better than model A based on the conclusion that the error statistic of model A is higher than model B. In order to make an accurate conclusion, we need to perform tests of significance.

Usually when evaluating out-of-sample forecasting, the rolling scheme method is used. However, since our time series models intrinsically do not put much weight on earlier forecasts, it is not necessary to use the rolling scheme method. A good test to compare the errors for both the recursive scheme and the rolling scheme along with small samples inefficiencies would be the Diebold and Mariano (DM) test for predictive accuracy.

Let $\{\hat{X}_{it}\}_{t=1}^{T}$ and $\{\hat{X}_{jt}\}_{t=1}^{T}$ be the two competing sets of forecasts from models i and j respectively for the time series $\{X_t\}_{t=1}^{T}$. Subsequently, let us define the associated forecast errors as,

$$\begin{aligned}\{e_{it}\}_{t=1}^{T} &= \{X_t\}_{t=1}^{T} - \{\hat{X}_{it}\}_{t=1}^{T}\\ \{e_{jt}\}_{t=1}^{T} &= \{X_t\}_{t=1}^{T} - \{\hat{X}_{jt}\}_{t=1}^{T},\end{aligned} \tag{2.15}$$

where $\{e_{it}\}_{t=1}^{T}$ are the error terms for model i and $\{e_{jt}\}_{t=1}^{T}$ are the error terms for model j. The accuracy of the forecast is measured by a particular loss function (Error Statistics in section 2.2.1) $g(\cdot)$ such that,

$$\begin{aligned}g(X_t, \hat{X}_{it}) &= g(e_{it})\\ g(X_t, \hat{X}_{jt}) &= g(e_{jt}).\end{aligned} \tag{2.16}$$

To determine if there is indeed a statistically significant difference between the error statistics of the two models, we set to null hypothesis and alternative hypothesis to be,

$$\begin{aligned}H_0 &: E\left[g(X_t, \hat{X}_{it})\right] = E\left[g(X_t, \hat{X}_{jt})\right]\\ H_a &: E\left[g(X_t, \hat{X}_{it})\right] \neq E\left[g(X_t, \hat{X}_{jt})\right].\end{aligned} \tag{2.17}$$

If we were to reject the null hypothesis, that means that there is a statistically significance difference between the error statistics of model i and model j. Therefore, the model with the lesser error statistics will be the better model. If we cannot reject the null hypothesis, this means that there is no statistically difference between the two models' error statistics, and we cannot come to a conclusive conclusion. Let us define the loss differential as,

$$d_t \equiv g(e_{it}) - g(e_{jt}). \tag{2.18}$$

Therefore, the mean of the loss differential is defined as,

$$\bar{d} = \frac{1}{T}\sum_{t=1}^{T}\left|g(e_{it}) - g(e_{jt})\right|, \tag{2.19}$$

and the test statistic is defined as,

$$S = \frac{\bar{d}}{\sqrt{\frac{1}{T}\hat{f}_{\bar{d}}}} \quad \text{where } S \sim N(0,\, 1), \tag{2.20}$$

where,

$$\hat{f}_{\bar{d}} = \gamma_0 + 2\sum_{j=1}^{\infty}\gamma_j,\ \gamma_j = \operatorname{cov}(d_t,\, d_{t-j}), \tag{2.21}$$

and $\hat{f}_{\bar{d}}$ is the consistent estimate of the asymptotic variance of $\sqrt{T}\bar{d}$. The asymptotic variance is used in the statistic to reduce serial correlations in the loss differentials. Since S is normally distributed, we could reject the null hypothesis at a 5% level if $|S| > 1.96$.

When forecasting evaluations, it is important to distinguish between in-sample and out-of-sample forecasts. A good forecasting model is one which can withstand the robustness of an out-of-sample test. It is also important to note that different forecasting methods are better suited to different financial markets, trading environment and economic conditions.

3 ARCH

As stated in Chapter 1, the financial markets returns distributions are far from being stable and constant. Volatility in these markets tends to cluster together, causing the financial markets to either fluctuate erratically sometimes, or remain relatively calm during other times. Another characteristic of financial market volatility is the time varying nature of the returns. However, all these features are important as it allows us to extract information and model it.

The first person to discover and implement a new class of time series was Robert Engle. He was awarded the 2003 Memorial Nobel Prize in Economics for this contribution. He named this new class of time series models the Autoregressive Conditional Heteroskedascity (ARCH) model. This chapter will focus on subsequent modifications of the ARCH class of models, and also introduce our proposed GARCH-S model. The GARCH-S model was derived from our own proposed Modified-GARCH model. Ever since Robert Engle proposed his ARCH model, several other ARCH type models have been developed. The popular ones include GARCH, GJR-GARCH and EGARCH.

In section 3.7, we shall propose our own GARCH-S model, and evaluate the GARCH-S model against other ARCH type models in section 3.9. In this chapter, we will use the Nikkei 225 daily values from 5th January 2006 to 28th May 2010. The corresponding Nikkei 225 log returns will be from 6th January to 28th May 2010.

3.1 Data Preparation

Before we can start modeling, we need suitable data. For our proposed modified-GARCH and GARCH-S model, we need to make use of both the Nikkei 225 data and the S&P 500 data as input variables for the model. Our proposed models make use of time $t-1$ S&P 500 data together with time $t-1$ Nikkei 225 data as inputs to produce a time t output.

Unfortunately, Japan and America have different public holidays. This means that, for a few days a year, the Nikkei 225 and S&P 500 operate on different days. This difference can cause a major problem for us. Even if just one day is off, it will mean that the remaining data will be off. From figure 3.1, we can see that just a difference of one day will create a chain of undesirable effects. It could potentially everything to be wrong.

Therefore, it is necessary to prepare the data before using. There are two scenarios: either the Nikkei 225 is on holiday and the S&P 500 is not, or the Nikkei 225 is not on holiday but the S&P 500 is. For scenario one, for the Nikkei 225 holiday at time t, we need to delete the previous one-day S&P 500 index value at time $t-1$. Then we need to move back the S&P 500 data by one time unit.

In the case of scenario two, for the S&P 500 holiday at time t, all we need to do is delete the corresponding Nikkei 225 index value at time t, and move back the Nikkei 225 index by one time unit. After this rather long and mundane process of preparing our data, we can finally proceed to modeling our volatility models. Figures 3.1, 3.2, and 3.3, provide a graphical representation of what we have just described.

Date (Nikkei)	Date (S&P 500)
19	19
20	20
~~22~~	21
~~23~~	22
~~24~~	23
~~25~~	24

Figure 3.1 Problem caused by the Nikkei 225 and the S&P 500 operating on different days (red indicating mismatch)

Date (Nikkei)	Date (S&P 500)
19	19
20	~~20~~ 21
22	~~21~~ 22
23	~~22~~ 23
24	~~23~~ 24
25	~~24~~ 25

Figure 3.2 Data preparation for scenario one where the Nikkei 225 index falls on a holiday but not the S&P 500. Blue Arrow indicates the direction in which the data have been moved forward.

Date (Nikkei)	Date (S&P 500)
19	19
20	20
21	22
~~22~~ 23	23
~~23~~ 24	24
~~24~~ 25	25

Figure 3.3 Data preparation for scenario two where the S&P 500 index falls on a holiday but not the Nikkei 225. Blue Arrow indicates the direction in which the data have been moved forward.

3.2 ARCH

In 1982, Robert Engle wrote a groundbreaking research paper titled, "Autoregressive Conditional Heteroscedasticity with Estimates of the Variation of United Kingdom Inflation". The paper introduced the ARCH class of time series statistical models. These class of models made use of the natural time-varying clustering effects of volatility to model volatility. It was a simple, yet powerful method to model volatility. The ARCH(q) model is defined as,

$$
\begin{aligned}
& r_t = \mu + \varepsilon_t, \qquad \varepsilon_t \sim N(0, \sqrt{h_t}) \\
& \varepsilon_t = z_t \sqrt{h_t}, \\
& h_t = \omega + \sum_{j=1}^{q} \alpha_j \varepsilon_{t-j}^2 \\
& \omega > 0,\ \alpha_j \geq 0 \quad \text{such that } h_t > 0,
\end{aligned}
\tag{3.1}
$$

where r_t is defined as the return of the asset at time t, μ is the conditional mean, ε_t is the error, or innovation, at time t, and $z_t \sim N(0,1)$. ω and α_j are the coefficients of the ARCH(q) model, and h_t is the conditional variance of the asset returns. The ARCH model is based on the conditional variance of asset returns, h_t, which is derived from the previous day's error, ε_{t-1}.

The ARCH model intrinsically provides us with a one-day volatility forecast. The multi-step-ahead forecast can be formulated by using $E[\varepsilon_{t+\tau}^2] = h_{t+\tau}$.

3.2.1 Weaknesses of the ARCH Model

It is important to note that there are a few weaknesses of the ARCH model. These weaknesses are mainly [Tsay (2010)],

1. The ARCH model assumes that positive and negative shocks to the volatility have the same effect. However, this is not the case in the ARCH model [McAleer, Verhoeven (2003)].

2. The ARCH model is rather restrictive in its parameters. This limits the ability of an ARCH model with Gaussian innovations to capture kurtosis.
3. The ARCH model does not provide any new insight for understanding the source of variations of the financial time series. It gives a way to describe the behavior of the conditional variance but no indication as to what causes.
4. In general, the ARCH model is likely to overpredict the volatility because it responds slowly to large isolated shocks to the returns.

Since the ARCH models have these weaknesses, why do we still continue to use them? This is because the ARCH class of models is simple to implement and quite effective to a certain degree. Nothing is perfect in finance or the social sciences. Emanuel Derman, a famous financial engineer and physicist, once said, “In Physics there may one day be a Theory of Everything; in finance and the social sciences, you’re lucky if there is a usable theory of anything.” What he meant is that in Physics, you can model Physics equations and apply it to the real world with perfect accuracy. However, it impossible to model financial equations and apply it to real world stock markets with perfect accuracy. In fact, we are nowhere close to the true real world values. In financial volatility modeling, even if we have the best fit model for a specific financial market, it is still impossible to model the “true volatility” to a degree of high accuracy. In fact, there is not even a fixed definition of “true volatility”. The researcher could choose what his/her “true volatility” will be. However, in general, squared daily returns, realized volatility are the most commonly used “true volatility”.

Therefore, we should be content with what models we have right now, while continuing to develop better models. Even though time series volatility modeling will never duplicate the “true volatility”, time series volatility modeling research is still an important topic. So much so that Robert Engle and Clive W.J. Granger were awarded the 2003 Memorial Nobel Prize in Economics for their contribution to time series volatility modeling.

3.3 Generalized ARCH (GARCH) Model

As stated in section 3.2, the ARCH class is the most basic model that started everything. In 1986, Tim Bollerslev, a bright student of Robert Engle, developed the Generalized Autoregressive Conditional Heteroskedascity (GARCH) model. Since 1986, other more complex models have been created but they all have the roots based on the GARCH model.

3.3.1 Introduction to GARCH

The GARCH model is basically the same as the ARCH model except that the conditional variance of asset returns, h_t, is not only derived from the previous day's error, ε_{t-1} but also from the previous day's conditional variance of the asset return. The GARCH model is defined as,

$$\begin{aligned} &r_t = \mu + \varepsilon_t, \qquad \varepsilon_t \sim N(0, \sqrt{h_t}) \\ &\varepsilon_t = z_t \sqrt{h_t}, \\ &h_t = \omega + \sum_{j=1}^{q} \alpha_j \varepsilon_{t-j}^2 + \sum_{i=1}^{p} \beta_i h_{t-i} \\ &\omega > 0,\ \alpha_j \geq 0,\ \beta_i \geq 0 \text{ such that } \sum_{i=1}^{\max(q,p)} \left(\alpha_i + \beta_i \right) < 1, \end{aligned} \tag{3.2}$$

where ω, α_j, and β_i are the coefficients of the GARCH(p, q) model. For the GARCH(1,1), the constraints $\alpha_1 \geq 0$, $\beta_1 \geq 0$ are needed to make sure that $h_t > 0$. The constraints on $\alpha_i + \beta_i$ implies that the unconditional variance of α_i is finite, while its conditional variance h_t evolves over time.

The unconditional variance for the GARCH(p, q) is defined as,

$$h = \frac{\omega}{1 - \sum_{j=1}^{q} \alpha_j - \sum_{i=1}^{p} \beta_i}. \tag{3.3}$$

From equation 3.3, the GARCH(p, q) is covariance stationary if and only if,

$$\left[\sum_{i=1}^{p}\beta_i+\sum_{j=1}^{q}\alpha_j\right]<1. \tag{3.4}$$

For most financial practical applications, it is sufficient enough to use the GARCH(1,1) model. There is no need for higher order models. Besides, the GARCH(1,1) model is the easiest to understand and implement, but yet is still effective enough to become widely used around the world.

3.3.2 GARCH(1,1)

In financial market volatility, usually, the GARCH(1,1) model is sufficient enough. The GARCH(1,1) model is defined as,

$$h_t=\omega+\alpha_1\varepsilon_{t-1}^2+\beta_1 h_{t-1},\qquad 0\le\alpha_1,\ \beta_1\le 1,\ \left(\alpha_1+\beta_1\right)<1. \tag{3.5}$$

Firstly, a large ε_{t-1}^2 or h_{t-1} gives rise to a large h_t. This means that a large h_{t-1} tends to be followed by another large h_t, generating, again the well-known behavior of volatility clustering in financial time series.

Secondly, it can be shown that if $1-2\alpha_1^2-\left(\alpha_1+\beta_1\right)^2>0$, then

$$\frac{E(\varepsilon_t^4)}{\left[E(\varepsilon_t^2)\right]^2}=\frac{3\left[1-\left(\alpha_1+\beta_1\right)^2\right]}{1-\left(\alpha_1+\beta_1\right)^2-2\alpha_1^2}>3. \tag{3.6}$$

[Refer to the appendix for the proof of equation 3.6] Consequently, similar to ARCH models, the tail distribution of a GARCH(1,1) process is heavier than that of a Normal distribution. This means that large changes are more often to occur as compared to a Normal distribution.

The one-step ahead forecast for a GARCH(1,1) model is,

$$\hat{h}_{t+1}=\omega+\alpha_1\varepsilon_t^2+\beta_1 h_t. \tag{3.7}$$

Subsequently, the two-day forecast will be,

$$\begin{aligned}\hat{h}_{t+2} &= \omega + \alpha_1 \varepsilon_{t+1}^2 + \beta_1 \hat{h}_{t+1} \\ &= \omega + (\alpha_1 + \beta_1)\hat{h}_{t+1}.\end{aligned} \tag{3.8}$$

[Please refer to Appendix for more information on proof.] By repeat substitution in equation 3.8, we can derive a general formula for forecasting the volatility of GARCH(1,1),

$$\hat{h}_{t+\tau} = \frac{\omega}{1-(\alpha_1+\beta_1)} + (\alpha_1+\beta_1)^{\tau}\left[\alpha_1\varepsilon_t^2 + \beta_1 h_t\right]. \tag{3.9}$$

From equation 3.9, we are able to see that long-term forecast by the GARCH(1,1) model converges to the unconditional variance of $\omega/[1-(\alpha_1+\beta_1)]$. This is because as τ goes to infinity, the term on the right hand side of the general equation disappears. This is also known as "mean reversion", which is another property of the GARCH models. Even though the GARCH model deviates around the mean, the GARCH model converges to a mean in the long run.

3.3.3 Problems with Using Daily Squared Returns

The GARCH models also share the same weaknesses as the ARCH models, as stated in section 3.2. For this chapter, we shall use both high-frequency data and daily returns to model volatility. However, using daily squared returns to proxy daily volatility might provide an imprecise estimator of the volatility, h_t. From equation 3.2,

$$r_t = \mu + \varepsilon_t, \qquad \varepsilon_t = z_t\sqrt{h_t}, \tag{3.10}$$

and $z_t \sim N(0,1)$. Then

$$E\left[\varepsilon_t^2 \mid \Phi_{t-1}\right] = h_t E\left[z_t^2 \mid \Phi_{t-1}\right] = h_t, \tag{3.11}$$

where Φ_{t-1} is all information in the set up to time $t-1$. Equation 3.11 is derived as a result of the assumption that $z_t^2 \sim \chi_{(1)}^2$, where the Normal distribution squared is a special case of the Chi-squared distribution. However, since the median of a $\chi_{(1)}^2$ distribution is 0.46, ε_t^2 would be less than $0.5h_t$ more than 50% of the time. In fact,

$$\Pr\left(\varepsilon_t^2 \in \left[\frac{1}{2}\sigma_t^2, \frac{3}{2}\sigma_t^2\right]\right) = \Pr\left(z_t^2 \in \left[\frac{1}{2}, \frac{3}{2}\right]\right) = 0.26. \tag{3.12}$$

This means that ε_t^2 is 50% greater than or smaller than σ_t^2 nearly 75% of the time. Therefore, using ε_t^2 as a volatility proxy will lead to low R^2 values and undermine the forecast inference's accuracy. However, for the one-day forecast, the R^2 value will increase by three to four folds when intraday 5 minutes squared returns are used to proxy the actual volatility instead of daily squared returns [Blair, Poon, Taylor (2001)].

When comparing forecast evaluations, using the daily squared returns as a proxy for the "true volatility" could lead to imprecise evaluations as the squared returns are an extremely noisy proxy of the "true volatility" [Andersen, Bollerslev (1997)]. In financial volatility research, it is now common to use intraday returns as a measure to proxy the "true volatility". However, just by using the intraday returns does not take into account the overnight gain/loss of the asset price. Therefore, a definition for "true volatility" has still not been clearly established.

In our thesis, we used the squared daily returns to proxy "true volatility", as the volatility derived from using daily returns not did differ too much from the volatility derived by using intraday returns for our given time period. Furthermore, as stated in section 1.2, we are not extremely concerned with the overall forecasting accuracy but rather will the GARCH-S provide any new information in forecasting as compared to other GARCH models.

3.3.4 Autocorrelation Function

Before we can start using time series to model volatility, we need to understand the autocorrelation function (acf). The acf is important because it is important tool used in choosing the right model to fit our data. If we have some serial correlation left in the residuals of our models, it means that our model might be not the best fit for the data.

The acf is defined as,

$$\begin{aligned} c_k &= \frac{1}{n}\sum_{t=1}^{n-k}(x_t-\overline{x})(x_{t+k}-\overline{x}) \\ a_k &= \frac{c_k}{c_0}, \end{aligned} \tag{3.13}$$

where x_t is defined as a random variable x at time t, and a_k is the autocorrelation function value at lag k. The acf is a tool that allows us to calculate the cross-correlation of the data within itself.

A property of the autocorrelation function is that autocorrelation of the powers of an absolute return are highest at power one: $corr(|r_t|,|r_{t-1}|) > corr(r_t^d, r_{t-1}^d),\ d \neq 1$. [Ding, Granger (1995)] call this property the Taylor effect, following [Taylor (1986)].

Another extremely popular tool used to check a model's fit is the Box-Ljung test, which is a type of portmanteau test. The Box-Ljung is defined as,

$$Q(h) = T(T+2)\sum_{k=1}^{h}\frac{\hat{\rho}_k^2}{T-k}, \tag{3.14}$$

where T is the sample size, $\hat{\rho}_k^2$ is the sample autocorrelation at lag k, and h is the number of lags being tested. The null hypothesis is $H_0 : \rho_1 = \ldots = \rho_k = 0$, and the alternative hypothesis is $H_a : \rho \neq 0$. The null hypothesis will be rejected if $Q(h) > \chi_\alpha^2$, where χ_α^2 denotes the $100(1-\alpha)$th percentile of a chi-squared distribution with m degrees of freedom. We will perform the Box-Ljung test on the residuals of our fitted model.

3.3.5 Modeling the Nikkei 225

The general guidelines for building a volatility model for a time series return are,

1.) State a mean equation by testing for serial dependency in the data. To eliminate serial dependency in the returns, we could use an ARMA model.
2.) Use the residuals of the mean equation to test for ARCH effects.
3.) State a volatility model if the ARCH effects are statistically significant.
4.) Check the fitted model and revise if necessary.

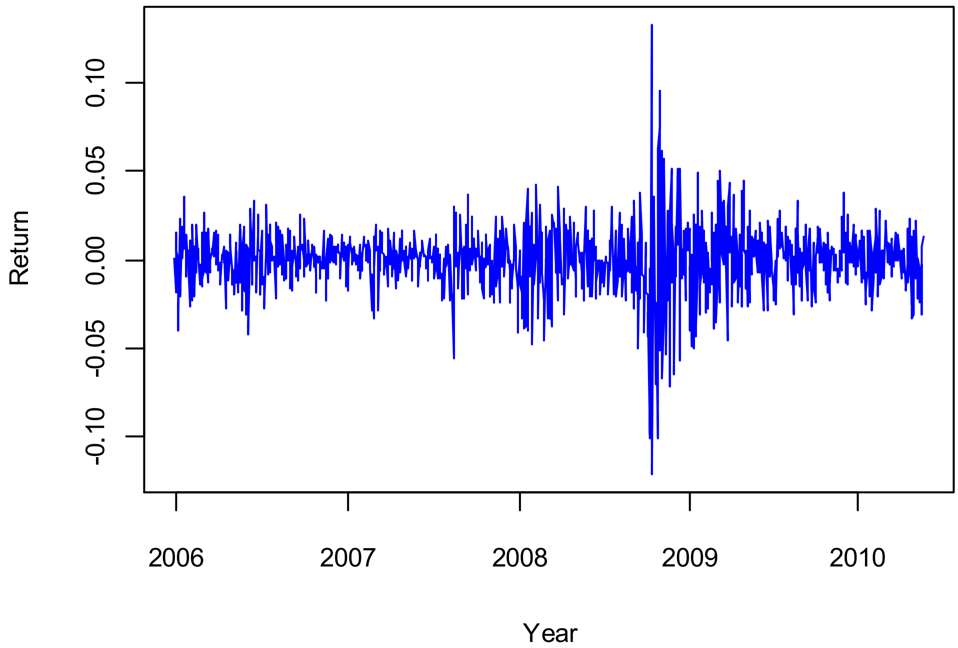

Figure 3.4 Historical daily returns for the Nikkei 225

Let us first start by modeling the Nikkei 225 using the basic GARCH(1,1) model. The returns of the Nikkei 225 from 5th January 2006 to 28th May 2010 are displayed in figure 3.4, and the corresponding acf plot for the Nikkei 225 returns from 6th January 2006 to 28th May 2010 is displayed in figure 3.5.

From figure 3.5, even though there are some statistically significant lags at the 0.05 level, we believe that there is no need to apply an ARMA model on the returns. The acf suggests that the returns are quite clearly not correlated with one another. Subsequently, let us take a look at figure 3.6 which is the Partial Autocorrelation Function (pacf) plot of the Nikkei 225 returns 5th January 2006 to 28th May 2010. The pacf is a function of the acf itself.

The pacf plot has some statistically significant lags showing evidence that ARCH effects are present. We can now proceed to applying the GARCH(1,1) model to the Nikkei 225 data.

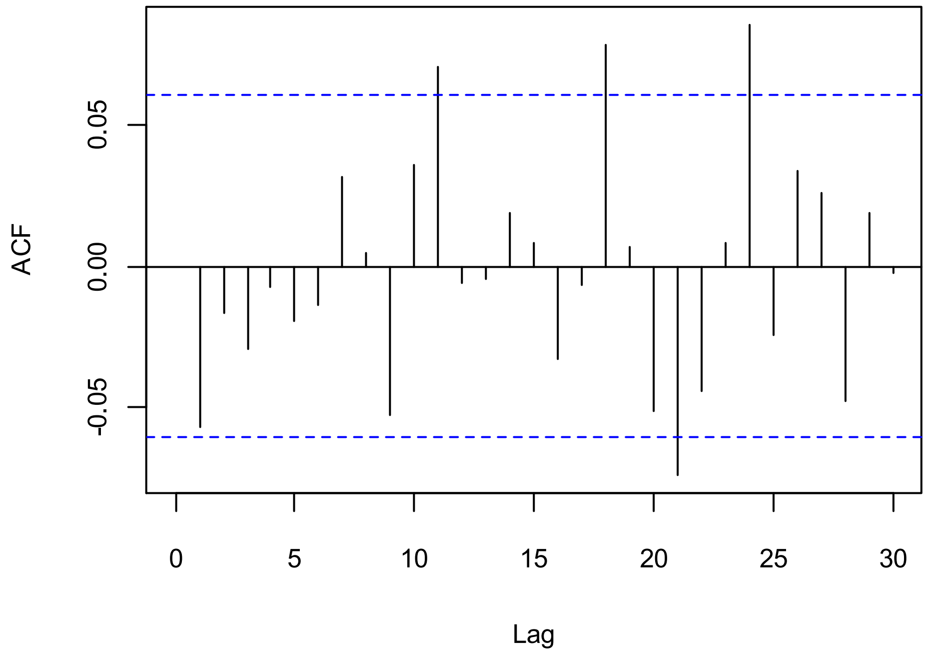

Figure 3.5 acf plot for the Nikkei 225 returns from 6th January 2006 to 28th May 2010

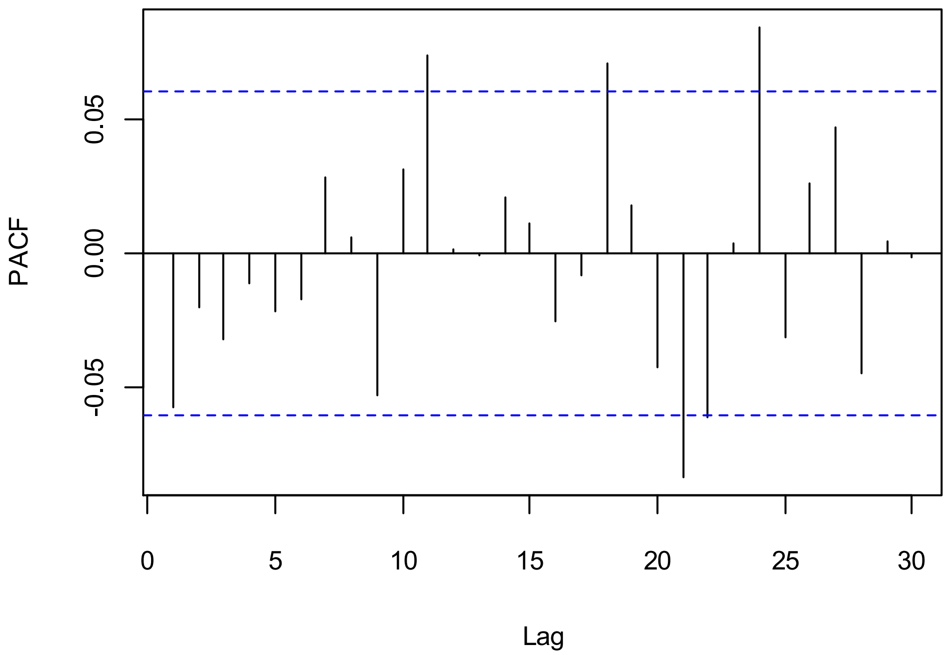

Figure 3.6 pacf plot for the Nikkei 225 returns from 6th January 2006 to 28th May 2010

Using the Gibbs Sampling method through R and JAGS [Please refer to the Appendix for more information on the program and type of Bayesian Markov Chain Monte Carlo (MCMC) method used in this thesis.] The GARCH(1,1) model summary obtained is,

Model:
GARCH(1,1)

Parameters	Estimate	Std. Error	t value	Pr(>|t|)
$>\omega$	5.33e-06	1.91e-06	2.79	0.0053
$>\alpha_1$	1.15e-01	2.07e-02	5.56	2.70e-08
$>\beta_1$	8.69e-01	2.16e-02	40.23	< 2.22e-16

The p-values for all the parameters were much smaller than 0.05. Therefore, all parameters obtained were statistically significant to the model. From the obtained parameters, our unconditional variance was estimated to be 0.00033. This results in a standard deviation of 0.018.

The DIC value obtained for the GARCH(1,1) was -5788.5. This DIC value will be used a base comparison against other models' DIC. Using the Box-Ljung test, all p-values obtained were much greater than 0.05, meaning we could not reject the null hypothesis that the data is random. And this is extremely important, as it proves that the residuals of the GARCH(1,1) are not autocorrelated. As an example the $Q(1)$ obtained had a p-value of 0.57.

We also plotted the acf and pacf of the residuals of the GARCH(1,1) model. Figures 3.7 and 3.8 are respectively the acf and pacf plots for the residuals of the GARCH(1,1) model fit on the Nikkei 225 for the time period from 6th January 2006 to 28th May 2010.

From figures 3.7 and 3.8, the residuals are clearly not statistically significant with one another. This is evidence agrees with the evidence provided from the Box-Ljung statistics test that the GARCH(1,1) model provides a good fit for the Nikkei 225 index.

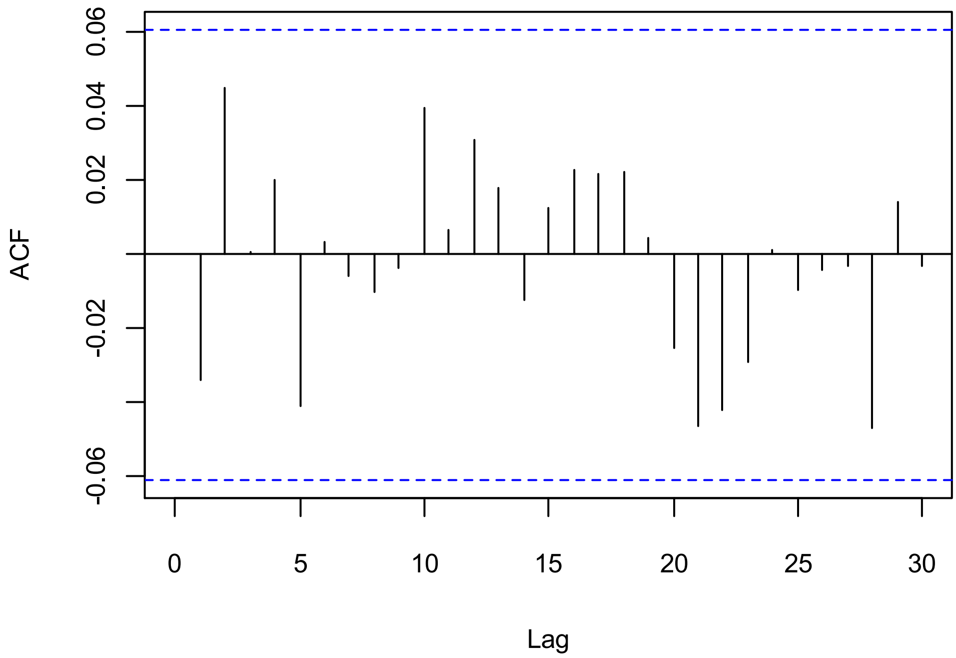

Figure 3.7 acf plot for the GARCH(1,1) model residuals of the Nikkei 225

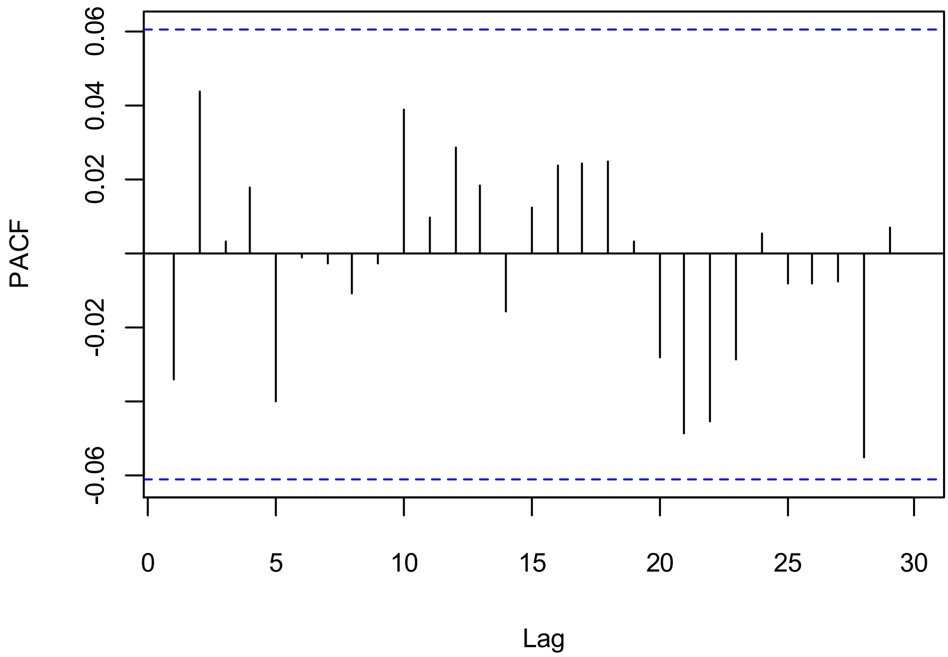

Figure 3.8 pacf plot for the GARCH(1,1) model residuals of the Nikkei 225

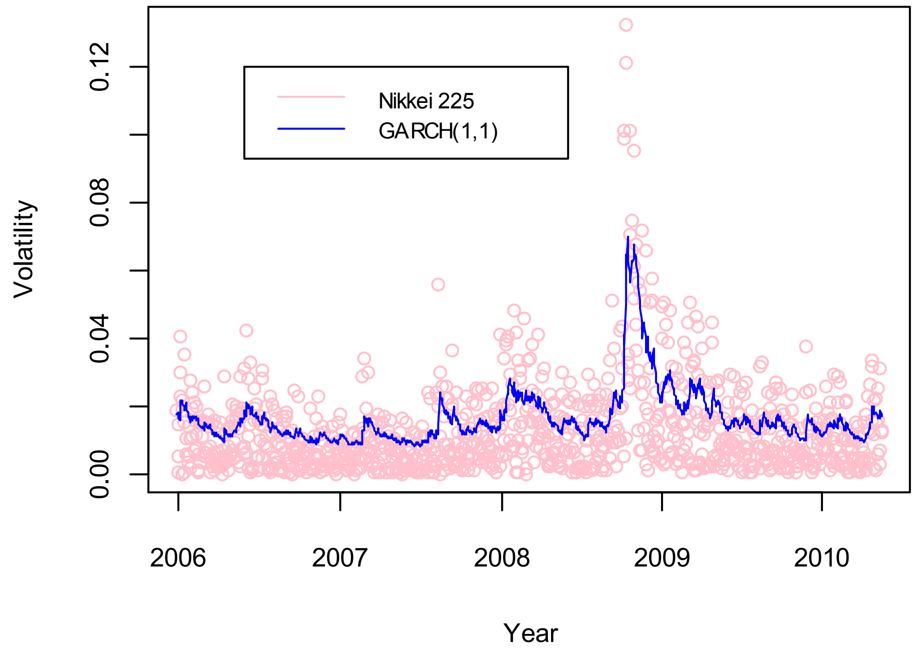

Figure 3.9 Plot of the GARCH(1,1) volatility against Nikkei 225 daily squared returns volatility from 6th January 2006 to 28th May 2010

Finally, we plotted the volatility obtained from the GARCH(1,1) against the daily squared returns volatility proxy. As described in section 3.3.3, please note that it will be quite noisy using the daily squared returns as a volatility proxy. From the graphical representation of figure 3.9, we can see that the GARCH(1,1) model provides a rather satisfactory fit of the Nikkei 225 squared returns volatility.

3.5 GJR-GARCH

The GARCH is a basic model. Since 1986, numerous other modified GARCH models have been created. One of the more commonly used models is the Glosten-Jagannathan-Runkle GARCH (GJR-GARCH) model [Glosten, Jagannathan, Runkle (1993)]. This model is an asymmetrical model in which more weight is given to negative innovations (errors) as compared to positive innovations (errors). As stated in section 3.2, one of the weaknesses of the ARCH class of models is that it assumes positive and negative shocks result in equal volatility. But in practice, negative shocks often produce higher volatility as compared to a positive shock of the same magnitude. This is also known as the leverage effect, which first appeared in [Black (1976)]. Therefore an asymmetrical model such as the GJR-GARCH would help in overcoming such a weakness.

The GJR-GARCH (p, q) model is defined as,

$$h_t = \omega + \sum_{i=1}^{p} \beta_i h_{t-i} + \sum_{j=1}^{q} \left(\alpha_j \varepsilon_{t-j}^2 + \lambda_j \varepsilon_{t-j}^2 I_{j,t-1} \right)$$
$$I_{t-1} = \begin{cases} 1 & \text{if } \varepsilon_{t-1} < 0 \\ 0 & \text{if } \varepsilon_{t-1} \geq 0 \end{cases}, \tag{3.15}$$

where ω, α_j, λ_j and β_i are the coefficients of the GJR-GARCH (p, q) model.

The difference between the GJR-GARCH model and the GARCH model is that there is an additional λ_j coefficient term. This additional term allows the model to be asymmetrical. For the GJR-GARCH model to be covariance stationary, we need the condition that

$$\sum_{i=1}^{p} \beta_i + \sum_{j=1}^{q} \left(\alpha_j + \frac{1}{2} \lambda_j \right) < 1. \tag{3.16}$$

In financial volatility modeling, the GJR-GARCH(1,1) model can provide a satisfactory fit for most modeling purposes. The GJR-GARCH(1,1) is defined as,

$$h_t = \omega + \beta_1 h_{t-1} + \alpha_1 \varepsilon_{t-1}^2 + \lambda_1 \varepsilon_{t-1}^2 I_{1,t-1}$$
$$I_{t-1} = \begin{cases} 1 & \text{if } \varepsilon_{t-1} < 0 \\ 0 & \text{if } \varepsilon_{t-1} \geq 0 \end{cases}, \tag{3.17}$$

and the one-step ahead forecast for the GJR-GARCH(1,1) model is,

$$\hat{h}_{t+1} = \omega + \beta_1 h_t + \alpha_1 \varepsilon_t^2 + \lambda_1 \varepsilon_t^2 I_t. \tag{3.18}$$

The additional λ_1 cause the one-day forecast volatility to be greater if the current day's innovation (error) is negative. This property agrees with the observed fact that negative shocks result in higher volatility than positive shocks. Using R's garchFit function, we obtained the following parameters for the GJR-GARCH(1,1) model,

Model:

GJR-GARCH(1,1) ****Using garchFit function in R****

Parameters	Estimate	Std. Error	t value	Pr(>|t|)
> ω	5.01e-06	1.47e-06	3.40	0.00068
> α_1	5.01e-02	3.02e-02	1.65	0.098
> β_1	9.00e-01	1.74e-02	51.61	< 2.22e-16
> λ_1	7.60e-01	4.93e-01	1.54	0.12

Even though the parameters were all statistically significant, the obtained GJR-GARCH(1,1) model is not covariance stationary. This is because the estimated parameters from the garchFit model do not satisfy equation 3.16. λ_1 has an unusually large weight of 0.76. This unusually large weight will exert too much influence on the fitted model and this will completely overlap any volatility momentum or mean reversion effects that we are trying to calculate.

However, we could use Bayesian statistical methods to estimate the parameters. Using R and JAGS (Just Another Gibbs Sampling) program, we obtain the following parameters,

Model:

GJR-GARCH(1,1)

Parameters	Estimate	Std. Error	t value	Pr(>|t|)
> ω	6.03e-06	1.70e-06	3.55	0.00039
> α_1	1.16e-02	0.017	0.69	0.49
> β_1	8.85e-01	0.021	43.04	<2.22e-16
> λ_1	1.60e-01	0.033	4.94	7.81-07

The t-value for the α_1 parameter is quite low, showing evidence that the α_1 parameter might be statistically insignificant. This means that we could ignore the α_1 term in the equation, as it is no different than 0. However, since the α_1 parameter makes only a rather small contribution to the volatility model as compared to other parameters, there is no need for any major concern, we still continued to use it in our model. The parameters estimated from using R and JAGS now satisfy equation 3.16, and is covariance stationary. The DIC obtained was -5820.3.

Using the Box-Ljung test, all p-values found for lags up to 20 could not provide enough evidence to reject the null hypothesis. Therefore, there are no autocorrelations in the standardized residuals. We are now able to proceed to modeling the GJR-GARCH(1,1) model's volatility using these parameters.

Figure 3.10 displays the GARCH(1,1), GJR-GARCH(1,1) volatility plotted against the Nikkei 225 historical volatility. From figure 3.10, we can see a graphical comparison between the volatility models and that the GJR-GARCH(1,1) is a slightly better model at modeling periods of high volatility as compared to the GARCH(1,1). Numerical evaluations and comparison will be worked out in chapter 4.

For easier comparison purposes, figure 3.11 provides a zoomed-in representation of figure 3.10 at the time of the American subprime mortgage crisis.

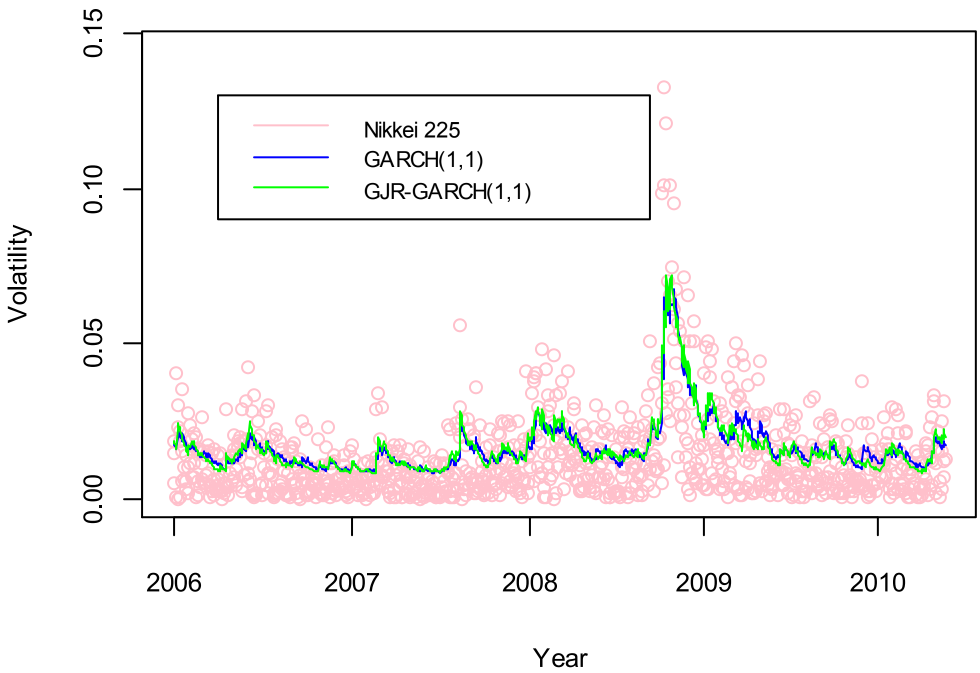

Figure 3.10 Plot of GARCH(1,1), GJR-GARCH(1,1) volatility against Nikkei 225 daily squared returns volatility

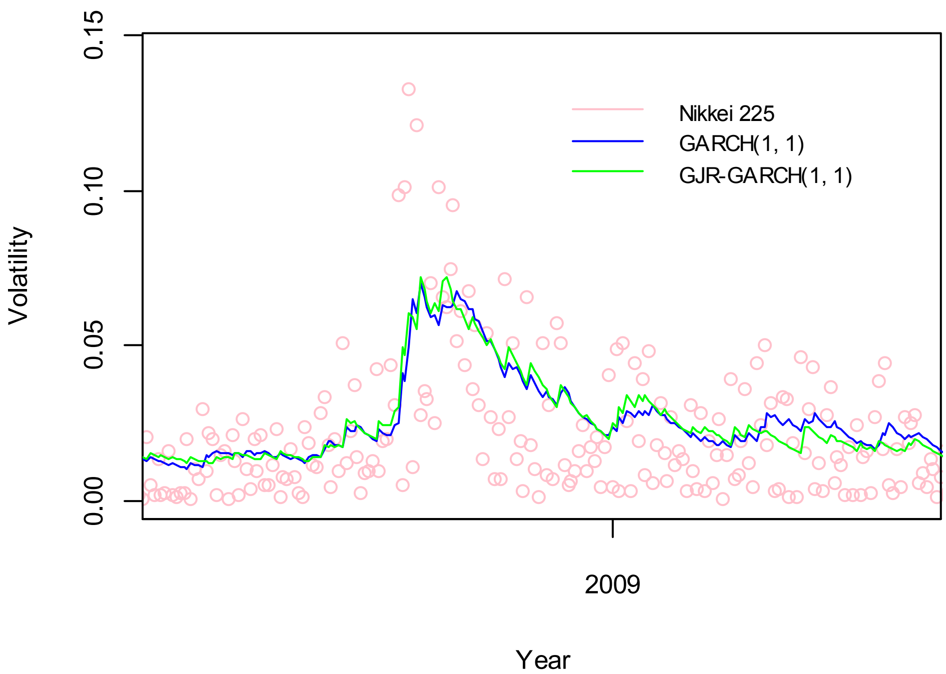

Figure 3.11 Enlarged plot of figure 3.10 focusing on the subprime mortgage crisis

3.6 Modified-GARCH

Before we proposed the GARCH-S model, we would like to begin with an earlier model we developed. We named this model the modified-GARCH model. Our goal of this model was to provide a better fit of the Nikkei 225 index while improving the one-day forecast accuracy by using S&P 500's as a secondary market to augment the target. While the overall forecast did not improve [see chapter 4], we did gained an important concept: that is that the modified-GARCH model was able to capture extremely high periods of volatility well. We then further this concept and developed our GARCH-S model.

The idea behind our proposed modified-GARCH model is to make use of the strong correlation between Nikkei 225 and the S&P 500 returns, especially in recent years. From 6th January 2006 to 28th May 2010, the correlation between the time t Nikkei 225 returns and time $t-1$ S&P 500 returns was 0.61.

The S&P 500 index is an index designed to reflect the state of the largest financial market in the world, and as such has a huge influence on other financial markets. Therefore, we expect that the Nikkei 225's volatility could potentially increase if the S&P 500's volatility increases or correspondingly, a higher absolute return of the S&P 500 at time $t-1$ would likely result in Nikkei 225 having higher volatility at time t. If we could extract information from the S&P 500, it could be potentially used to better model and forecast the Nikkei 225 volatility.

3.6.1 Programming the Modified-GARCH Model

In developing our modified-GARCH model, we used the GJR-GARCH(1,1) model for inspiration. Our proposed modified-GARCH (p, q) model is defined as,

$$
\begin{aligned}
r_t &= \mu + \varepsilon_t, \qquad \varepsilon_t \sim N(0, \sqrt{h_t}) \\
\varepsilon_t &= z_t \sqrt{h_t}, \\
h_t &= \omega + \sum_{i=1}^{p} \beta_i h_{t-i} + \sum_{j=1}^{q} \left(\alpha_j \varepsilon_{t-j}^2 + \gamma_j \varepsilon_{t-j}^2 I_{j,t-1} \right) \\
I_{t-1} &= \begin{cases} 0 & \text{if } y_{t-1} < \delta \\ 1 & \text{if } y_{t-1} \geq \delta \end{cases},
\end{aligned}
\tag{3.19}
$$

where y_{t-1} is the S&P 500 (correlated secondary market) returns at time $t-1$, and δ is an arbitrary parameter that the user can set according to the data. It is extremely important that y_{t-1} (secondary market) is chosen from an influential market. The $t-1$ returns of the influential market (secondary market) should be correlated to the t returns of the influenced market. The influenced market in this case is the Nikkei 225 (primary market). The δ chosen should correspond to a relatively large number on the magnitude scale of the absolute returns of the S&P 500 (secondary market). Typically, only the top 22% of the absolute returns should occur above δ in magnitude. The user could use a different value other than 22%. But the typical range is from 15% to 30%. If the number is used is too small, the modified GARCH would not be able to capture the peaks of high volatility, and be no different than a GARCH model. If the number used is too large, then the modified GARCH model would drift away from the mean "true volatility" rendering it rather useless.

The modified GARCH model drastically differs from other GARCH models because it introduces a foreign term, the secondary market. This is important because a foreign term allows the modified GARCH model to grasp the peaks of high volatility. If the primary market were to replace the secondary market in the modified GARCH model, the model would not be able to grasp the peaks of high volatility and be not much different as compared to the GARCH or GJR-GARCH models. We need to model the primary Nikkei 225 market while adding a pinch of information from the secondary S&P 500 market to grasp the peaks of high volatility.

From figure 3.12, we obtain a graphical representation of the absolute returns of the S&P 500. The δ value was chosen to be 0.015. On figure 3.12, the δ value of 0.015 is

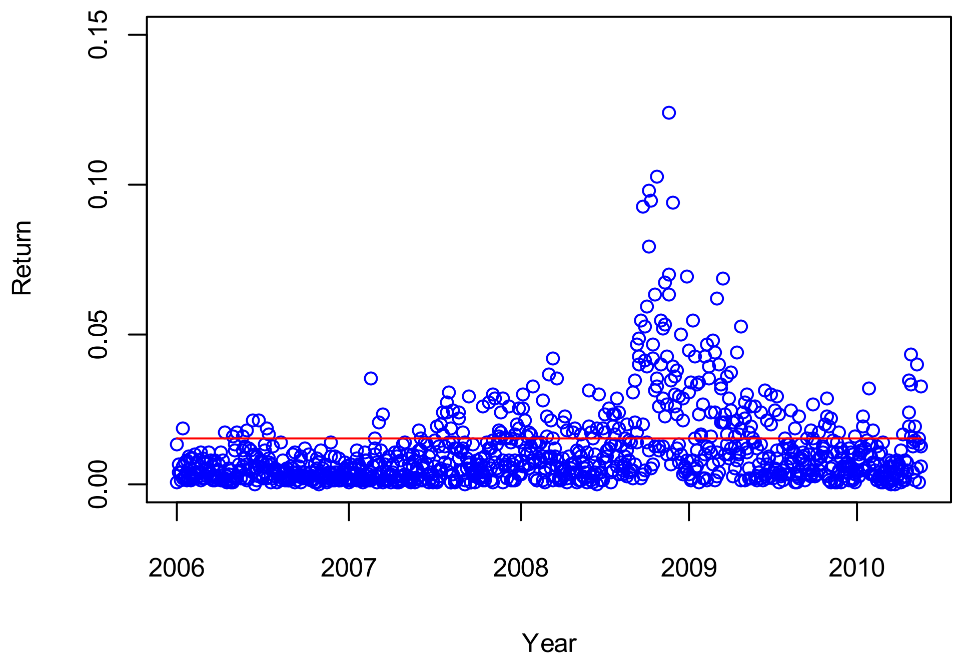

Figure 3.12 Absolute returns of S&P 500 from 5th January 2006 to 27th May 2010 and the straight line (red) with a value of 0.015 on the y-axis.

represented by a red line. From figure 3.12, it is clear that only the absolute returns that are higher in magnitude are above the red line.

Since this is an original model, there are no statistical programs or packages that could model this modified-GARCH model. Therefore, we have to create and program our model. There are 5 main steps that we must undertake when creating any GARCH type models using the Maximum Likelihood Method (MLE) are,

1. The initialization of the time series, model parameters and boundaries.
2. Setting the conditional distribution function (Normal Distribution) and composing the composition of the log-likelihood function.
3. Optimizing the model parameters and computing the numerical Hessian if using Newton-type optimization methods.
4. The summary of the optimization results.
5. Testing and verification.

[Please refer to the Appendix for a detailed description of our program and code.] For our program, we used the modified-GARCH(1,1) model and choose 0.015 for δ,

$$h_t = \omega + \beta_1 h_{t-1} + \alpha_1 \varepsilon_{t-1}^2 + \gamma_1 h_{t-1} I_{1,t-1}$$
$$I_{t-1} = \begin{cases} 0 \;\text{ if }\; y_{t-1} < 0.015 \\ 1 \;\text{ if }\; y_{t-1} \geq 0.015 \end{cases}. \tag{3.20}$$

Using our modified-GARCH program, we obtained the following estimated parameters,

Model:

Modified-GARCH(1,1) ****Using Maximum Likelihood Estimation****

Parameters	Estimate	Std. Error	t value	Pr(>|t|)
$>\omega$	7.25e-06	2.19e-06	3.32	0.00091
$>\alpha_1$	6.71e-02	1.95e-02	3.45	0.00056
$>\beta_1$	8.85e-01	2.32e-02	37.64	< 2.22e-16
$>\lambda_1$	1.60e-01	3.85e-02	3.14	0.001

The t-values and p-values provide evidence that the estimated parameters were statistically significant to the model.

The above stated MLE using Newton type estimation method is one method to estimated parameters. However, our program might run into problems such as the optimization not converging to the “true” parameters.

Therefore, for stability and comparison purposes, we shall concentrate on the more dynamic and flexible Bayesian MCMC (Gibbs Sampling) method to estimate parameters. Since Gibbs Sampling and MLE using Newton type estimation methods are two different concepts, the values of the parameters obtained from each of these estimation methods will be different. In general, Bayesian Statistics and frequentist MLE estimations should be treated as two different cases. Although in some cases, they could agree with one another [Casella, Berger (1987)]. The advantage of having both methods available to us is that we can always compare our results using both methods. The values obtained through the Gibbs Sampling method are, Model:

Modified-GARCH(1,1)

Parameters	Estimate	Std. Error	t value	Pr(>|t|)
$>\omega$	9.11e-06	2.90e-06	3.19	0.0014
$>\alpha_1$	7.41e-02	0.021	3.49	0.00050
$>\beta_1$	8.58e-01	0.029	29.95	< 2.22e-16
$>\lambda_1$	1.43e-01	0.048	2.98	0.0029

All the t-values obtained were sufficiently high, showing evidence that all parameters are statistically significant. The DIC value obtained was -5798.3. When compared to the GARCH(1,1) DIC value of -5788.5, the DIC value decreased by around 10. This means that the extra term we introduced in the Modified-GARCH(1,1) was able to provide new information as compared to the GARCH(1,1). Using these evidences, we can conclude that all the parameters made a statistically significant contribution to the model.

In this case, the parameters estimated from the Gibbs Sampling method are generally within one standard deviation from the parameters estimated using the MLE method. Therefore, in this case, the estimates derived by the MLE and Gibbs Sampling method did agree to a certain extent.

Using the Box-Ljung test, we found no evidence against the null hypothesis that the residuals are random. Therefore, the modified-GARCH(1,1) was found to be a satisfactory fit for the Nikkei 225's volatility. We can now proceed to calculating and plotting volatilities from the modified-GARCH(1,1).

From figures 3.13 and 3.14 we can see that the modified-GARCH(1,1) typically depicts a more accurate representation of high volatility as compared to the GARCH(1,1) and GJR-GARCH(1,1). The GARCH(1,1) and GJR-GARCH(1,1) tends to underestimate volatility, especially during the subprime mortgage crisis period, as compared to the modified-GARCH(1,1) model. This is because the Nikkei 225's daily volatility and the S&P 500's daily volatility have a higher correlation with each other during periods of high volatility. Thus, the modified-GARCH(1,1) model is able to more accurately use the information from the S&P 500 to model the Nikkei 225's volatility during periods of high volatility. The result is a reduction in underestimated errors. However, the overestimated errors increase as a direct consequence.

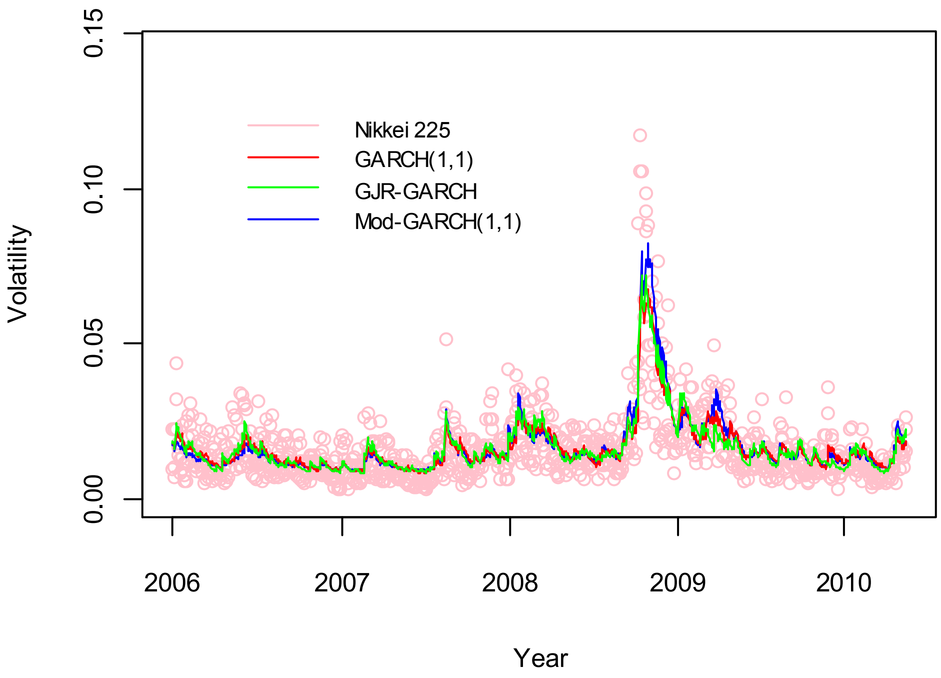

Figure 3.13 Plot of GARCH(1,1), GJR-GARCH(1,1), Mod-GARCH(1,1) volatility against the Nikkei 225 daily squared returns volatility

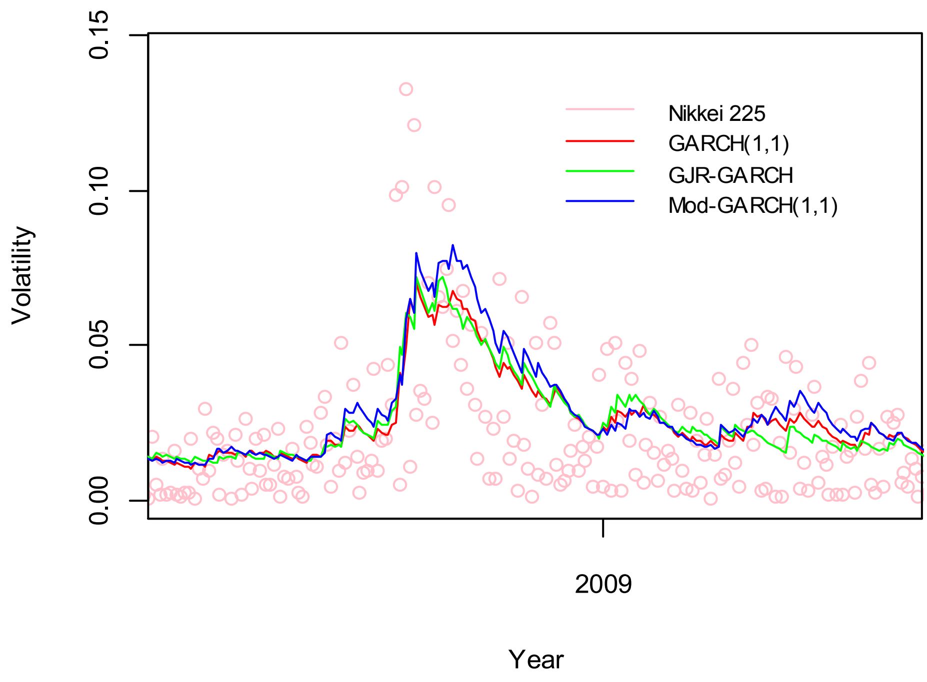

Figure 3.14 Enlarged plot of figure 3.13 focusing on the subprime mortgage crisis

3.7 GARCH-S

The GARCH-S model is the main focus of our research. Using information gained from our modified-GARCH model, we will take the modifications one step further. From section 3.6, we have shown that our additional parameter does provide new information in modeling the volatility to a certain extent. Our GARCH-S model takes this additional parameter one step further by using the S&P 500's innovations (secondary market) as an actual term in model. We believe this idea works because the absolute returns of the S&P 500 in figure 3.12 closely resembles the Nikkei 225 (primary market) daily squared returns volatility. Furthermore, it might even provide new information that is not available just by using the Nikkei 225's (primary market) daily returns.

The main idea behind the GARCH-S model is that high volatility in the S&P 500 at time $t-1$ has a strong correlation to high volatility in the Nikkei 225 at time t. Therefore if the returns for the S&P 500 at time $t-1$ were large, then the volatility for the Nikkei 225 at time t will be large. Our proposed GARCH-S (p, q) model is defined as,

$$
\begin{aligned}
& r_t = \mu + \varepsilon_t, \qquad \varepsilon_t \sim N(0, \sqrt{h_t}) \\
& \varepsilon_t = z_t \sqrt{h_t}, \\
& h_t = \omega + \sum_{i=1}^{p} \beta_i h_{t-i} + \sum_{j=1}^{q} \left(\alpha_j \varepsilon_{t-j}^2 + \gamma_j I_{j,t-1} (\varepsilon_{t-j}^2 + \phi_{t-j}^2) \right) \\
& I_{t-1} = \begin{cases} 0 & \text{if } y_{t-1} < \delta \\ 1 & \text{if } y_{t-1} \geq \delta, \end{cases}
\end{aligned}
\tag{3.21}
$$

where ϕ_t is the innovation (error) of the S&P 500 (correlated secondary model) at time t. The main difference between equation 3.20 and equation 3.18 is that equation 3.20 now has an additional ϕ_{t-j}^2 term in the equation.

For our simulation, we used the GARCH-S(1,1) model, which is defined as,

$$
\begin{aligned}
& h_t = \omega + \beta_1 h_{t-1} + \alpha_1 \varepsilon_{t-1}^2 + \gamma_1 I_{1,t-1} (\varepsilon_{t-1}^2 + \phi_{t-1}^2) \\
& I_{t-1} = \begin{cases} 0 & \text{if } y_{t-1} < 0.015 \\ 1 & \text{if } y_{t-1} \geq 0.015. \end{cases}
\end{aligned}
\tag{3.22}
$$

Before we can begin simulating the GARCH-S model, we first need to remove any serial

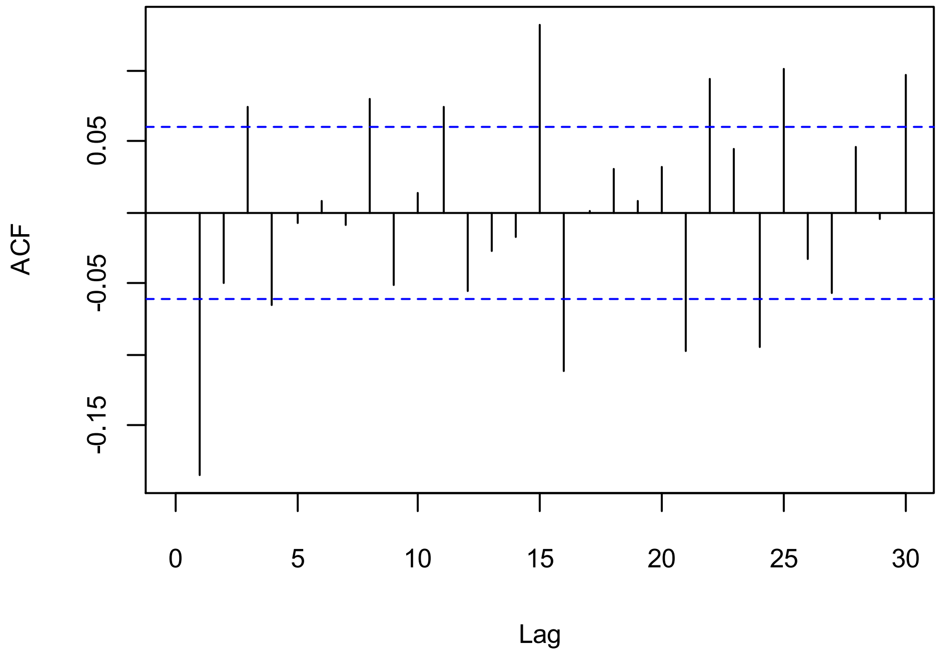

Figure 3.15 Autocorrelation function plot for the model residuals of the S&P 500

correlations in the S&P 500 returns. From figure 3.15, the acf plot of the S&P 500 returns shows evidence that there is indeed serial correlation present.

From the pacf plot of figure 3.16, we can see that lags 1 and 2 are significant lags. This hints that an ARMA (2, 0) (Autoregressive Moving Average) model might be useful in reducing the serial correlation of the S&P 500's return.

Subsequently, we proceed to implement the ARMA (2, 0). The estimated parameters for the ARMA (2, 0) fit are,

Model:

ARMA(2,0)

Parameters	Estimate	Std. Error	t value	Pr （＞｜t｜）
> ar1	-0.20	0.031	-6.51	7.64e-11
> ar2	-0.086	0.031	-2.80	0.0005
> Intercept	-0.00021	0.00052	-0.40	0.69

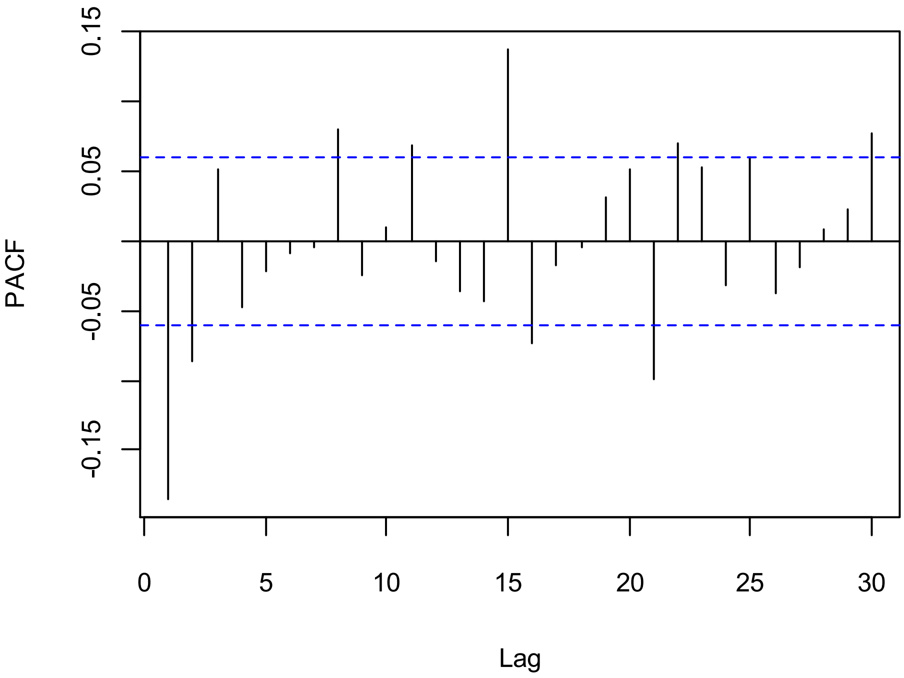

Figure 3.16 Partial autocorrelation function plot for the model residuals of the S&P 500

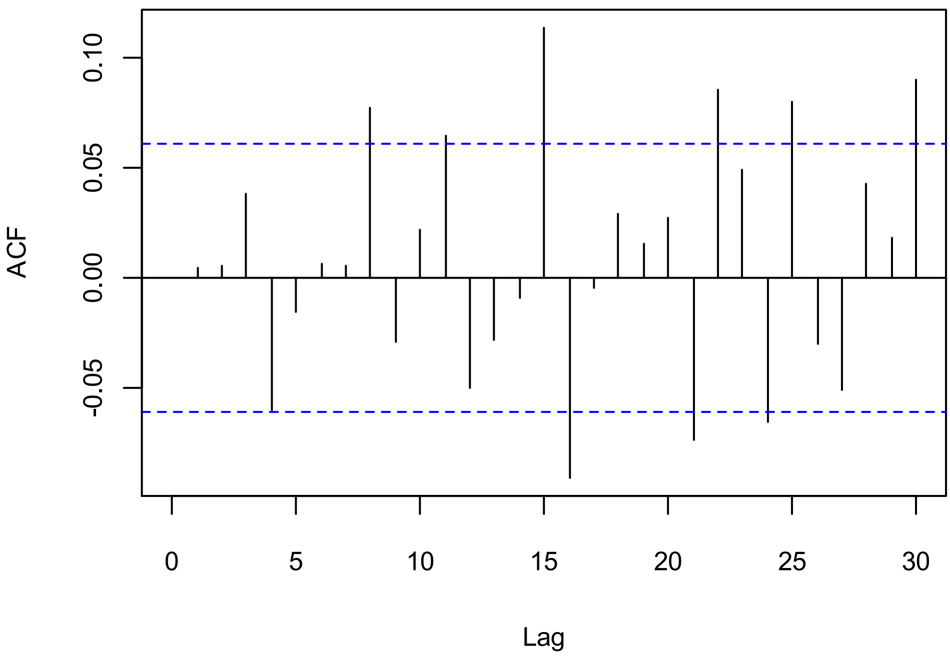

Figure 3.17 Autocorrelation function plot for the model residuals of the S&P 500 ARMA (2, 0) model from 6th January 2006 to 28th May 2010

We can see that the p-values for the ar1 and ar2 coefficients are extremely low indicating that they are statistically significant. The intercept is not statistically significant but we can ignore this as the intercept's value is extremely low and would be inconsequential when modeling.

Figure 3.17 displays the acf for the S&P 500 ARMA (2, 0) model. From figure 3.17, we can see that the ARMA (2, 0) model did indeed reduce serial correlations in the S&P 500 returns.

We can now use the ARMA (2, 0) to calculate the innovations of our S&P 500 returns,

$$y_t = -0.00021 - 0.20 y_{t-1} - 0.086 y_{t-2} + \phi_t . \tag{3.23}$$

After we have obtained our innovations for the S&P 500, we can now move on to estimating our GARCH-S(1,1) using the Gibbs Sampling method. The following parameters were obtained,

Model:

GARCH-S(1,1)

Parameters	Estimate	Std. Error	t value	Pr（>|t|）
> ω	2.08e-05	8.60e-06	2.41	0.016
> α_1	6.73e-02	0.022	3.02	0.0025
> β_1	7.47e-01	0.069	10.88	<2.22e-16
> γ_1	1.29e-01	0.040	3.22	0.0013

All the p-values obtained were sufficiently small, showing evidence that the parameters are all statistically significant. The DIC value obtained was -5833.7. This DIC value was much lower than any other DIC values we have obtained in the past. Therefore, adding the S&P 500 returns to the volatility model did provide significant new information in modeling the Nikkei 225 volatility.

Using the Box-Ljung test, we found no evidence against the null hypothesis that the residuals are random. Therefore, the GARCH-S(1,1) was found to be a satisfactory fit for the Nikkei 225's volatility. We can now proceed to calculating and plotting volatilities from the GARCH-S(1,1).

From figures 3.18 and 3.19, we can see that the GARCH-S(1,1) and modified-GARCH(1,1) models indeed have reduced the underestimated errors. From 2006 to 2007, the GARCH-S(1,1) and Mod-GARCH(1,1) were not as effective because the correlation was not as strong between the Nikkei 225 and the S&P 500. However, during the American subprime mortgage, which was caused by the USA, the correlation was rather strong. Subsequently, the GARCH-S(1,1) and modified-GARCH(1,1) models used this correlation to add new information to the volatility. Therefore, from figures 3.18 and 3.19, we can see that the GARCH-S(1,1) and modified-GARCH(1,1) differ in trend (less underestimate errors) as compared to the GARCH(1,1) and GJR-GARCH(1,1).

However, it is difficult to make any conclusions without numerical tests. In chapter 4, we will perform evaluations using numerical tests.

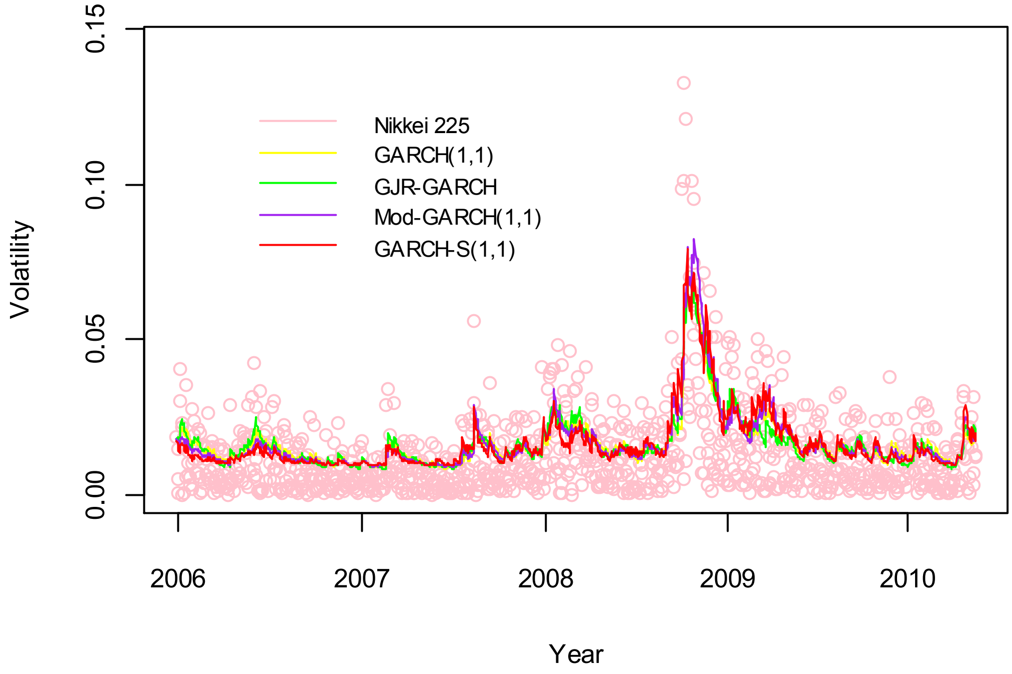

Figure 3.18 Plot of GARCH(1,1), GJR-GARCH(1,1), Mod-GARCH(1,1), GARCH-S(1,1) volatility against the Nikkei 225 daily squared returns volatility

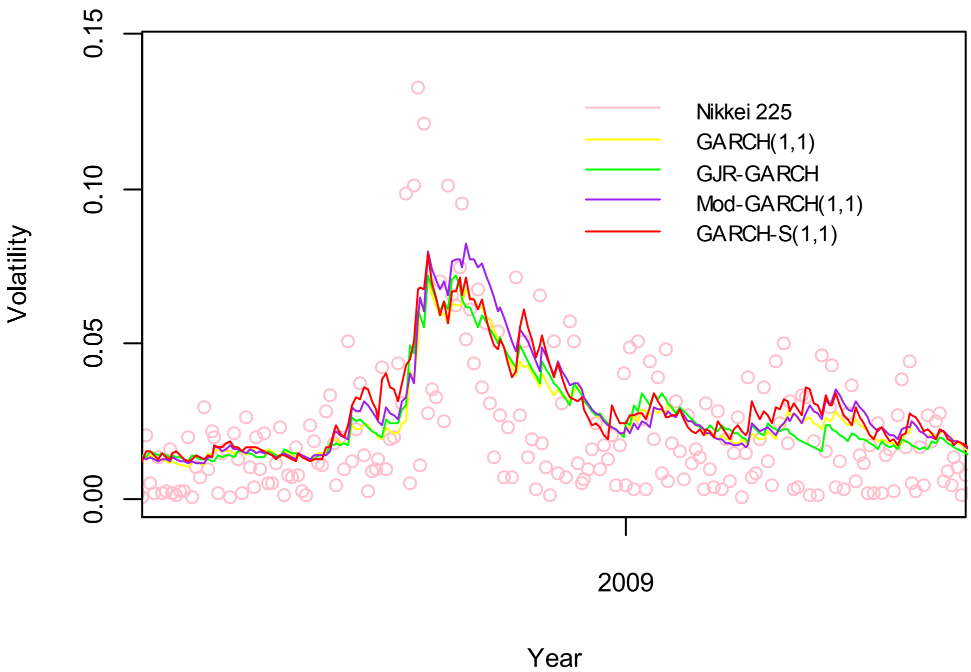

Figure 3.19 Enlarged plot of figure 3.18 focusing on then subprime mortgage crisis

3.8 EGARCH

Before we undergo any forecast evaluations, we would like to introduce the EGARCH model [Nelson (1991)]. The EGARCH model is an extremely popular and widely used model. In fact, it has been favored for the volatility of specific stock indices and exchange rates by several authors [Cao, Tsay (1992)] [Heyen, Kat (1994)] [Pagan, Stewart (1990)]. The theory behind the EGARCH model is rather different than our GARCH, GJR-GARCH, and GARCH-S models. However, we would still like to compare the effectiveness of the GARCH-S model against the popular EGARCH model.

The advantage of the EGARCH (p, q) model is that it specifies conditional variance in logarithmic form, which means that there is no need to impose an estimation constraint in order to avoid negative variance. The EGARCH (p, q) is defined as,

$$\begin{aligned}\ln h_t &= \omega + \sum_{i=1}^{q} \beta_i \ln h_{t-i} \\ &\quad + \sum_{j=1}^{p} \left[\theta_j \varphi_{t-j} + \gamma_j \left(\left| \varphi_{t-j} \right| - \sqrt{2/\pi} \right) \right] \\ \varphi_t &= \frac{\varepsilon_t}{\sqrt{h_t}}.\end{aligned} \tag{3.24}$$

The conditional variance, h_t, depends on both the size and sign of ε_t. With appropriate conditioning of the parameters, the EGARCH model will capture the stylized fact that a negative shock leads to a higher conditional variance in the subsequent period as compared to a positive shock. The EGARCH model is covariance stationary if and only if $\sum_{j=1}^{q} \beta_j < 1$.

We simulated the parameters for the EGARCH(1,1) model by using the Gibbs Sampling method. The following parameters were obtained, Model:

EGARCH(1,1)

Parameters	Estimate	Std. Error	t value	Pr (>|t|)
> ω	-0.26	0.039	-6.63	3.36e-11
> β_1	0.97	0.0046	208.42	<2.22e-16
> θ_1	-0.13	0.018	-7.32	2.48e-13
> γ_1	0.16	0.029	5.36	8.32e-08

All the p-values obtained are extremely small, showing evidence that the parameters are all statistically significant. The DIC obtained was -5823.8. And using the Box-Ljung test, we cannot reject the null hypothesis that the residuals of the fit are not random. Therefore, we can proceed to calculating and plotting volatilities using the EGARCH(1,1) model.

From figure 3.20 and 3.21, we can see that the EGARCH(1,1) model does provide a rather good fit. In fact, just by looking at the graphs, the EGARCH(1,1) model seems to be the one closest to the "true volatility" mean. However, during the American subprime mortgage crisis, it could grossly underestimate volatility in some days. Therefore, compared to the GARCH-S, it might not be as suitable for risk management as it underestimates the Nikkei 225 volatility during periods of high volatility.

From figure 3.20 and 3.21, we can see a property of the GARCH class of models. The property is that no specific GARCH model is the best for everything. Even within a specific market, some GARCH models do outperform others depending on the time period used.

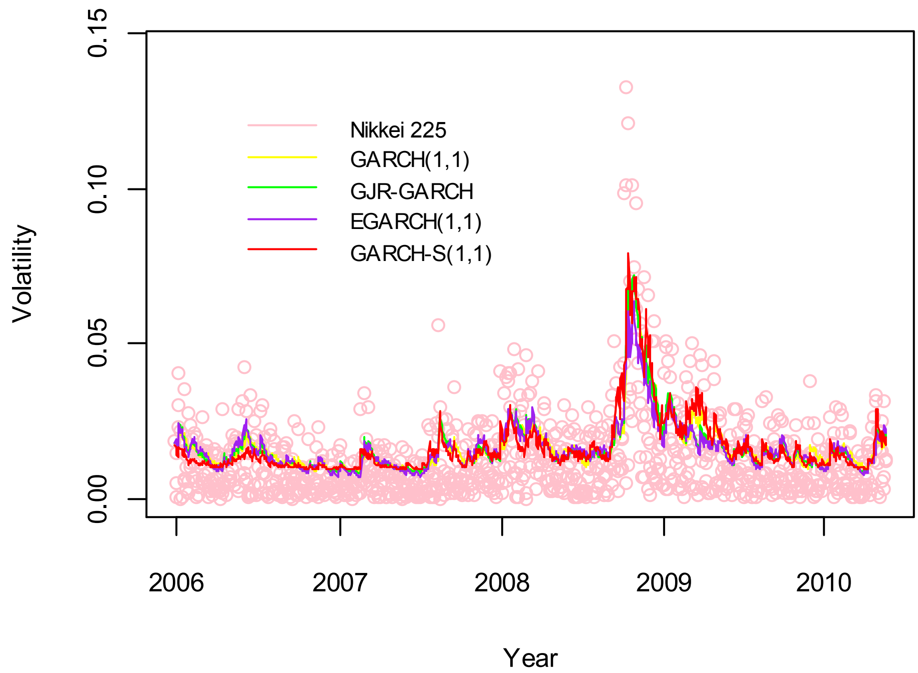

Figure 3.20 Plot of GARCH(1,1), GJR-GARCH(1,1), EGARCH(1,1), GARCH-S(1,1) volatility against daily squared returns Nikkei 225 volatility

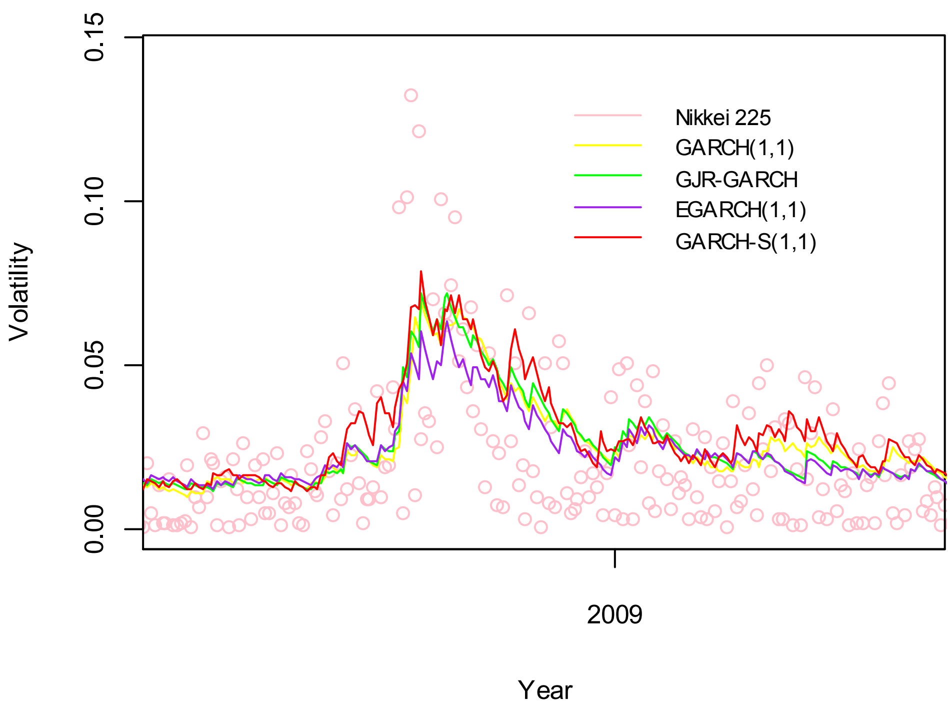

Figure 3.21 Enlarged plot of figure 3.20 focusing on the subprime mortgage crisis

4 Evaluation

In this chapter, we will evaluate and compare our models. The first part of this chapter is about using the more traditional error statistics found in section 2.2.1 to evaluate our models. Both in-sample (6th January 2006 to 28th May 2010) and out-of-sample (29th May 2010 to 28th January 2011) forecast accuracies will be evaluated.

Since we consider the period of time around the American subprime mortgage crisis to be the most interesting, we will perform another evaluation analysis in section 4.3. In this section, we will use the in-sample data from 6th January 2006 to 13th March 2008, and the out-of-sample data from 14th March 2008 to 28th May 2010. By doing so, we can classify the period of time around the American subprime mortgage crisis as out-of-sample data.

4.1 Difficulties in Forecast Evaluations

Compared to other research areas, literatures on GARCH forecast evaluations have been much less. The main reason why so little research is being on forecast evaluations is because a forecast evaluation is not a worthwhile research area. This is because there are so much numerous assumptions we have to make. For example, we do not even have a proper definition of "true volatility". What one person defines as "true volatility" might not be the "true volatility" definition for another person. There is a general consensus that realized volatility using intraday returns will provide the most accurate depiction of "true volatility". But even then, using intraday returns will not account for overnight volatility. Furthermore, intraday returns are affected by market microstructure effects including non-synchronous trading, discrete price observations, and intraday periodic volatility patterns.

Another problem with forecasting accuracy is that for different markets, different periods of time, different returns probably density function assumptions, can all lead to different results. This is due to the intrinsic nature of the financial markets. No single model can accurately describe the financial markets, and there is no exception in financial market volatility modeling. Furthermore, in this paper, we used,

1. The Normal distributions as our returns' probability density function.
2. Daily squared returns as our model inputs and proxy for volatility.
3. Bayesian MCMC (Gibbs Sampling) method to estimate the parameters.

Therefore, there is a huge amount of noise and uncertainty when we perform forecasting accuracies. However, all is not lost; a conclusive and important conclusion about the GARCH-S will be evaluated in section 4.3.

4.2 Error Statistics Evaluation

Despite the difficulties involved with forecasting accuracy evaluation, we still need to perform some tests in order to compare models. However, we need to make several assumptions before we can begin our evaluation. For our forecast evaluation purposes, the following assumptions were made,

1. The probability density function of the returns is assumed to follow the Gaussian Normal distribution.
2. The time frame for in-sample the Nikkei 225 data is from January 6th, 2006 to May 28th, 2010.
3. The out-of-sample time frame for the Nikkei 225 data is from May 28th, 2010 to January 28th, 2011.
4. For our first definition of "true volatility" we will use is the daily squared returns.
5. For our second definition of "true volatility" will we use is a hybrid mixture of the daily square returns and the high-low measure. The high-low measure [Bollen, Inder (2002)] is defined as,

$$\hat{\sigma}_t^2 = \frac{(\ln H_t - \ln L_t)^2}{4\ln 2}, \tag{4.1}$$

   where H_t is the highest price and L_t is the lowest price on day t. And our hybrid mixture volatility proxy will be defined as,

$$\hat{\sigma}_t^2 = \frac{1}{2} r_t^2 + \frac{\ln(H_t - L_t)^2}{8 \ln 2}. \tag{4.2}$$

Our hybrid volatility proxy of equation 4.2 is equal to half the daily squared returns volatility proxy plus half the high-low measure described in equation 4.1.

6. All data used have undergone the data preparation methods as described in section 3.1.
7. The methods used to evaluate forecasting accuracy are based on the methods as described in chapter 2 of this thesis.
8. Our parameter estimates were estimated using the Gibbs Sampling method which was implemented in JAGS through R.

And the methodology we used for our rankings are,

1. We will implement the MSE, RMSE, MAE, and MAPE error statistics as described in section 2.2.1.
2. Each of these error statistics will be performed two times for each model. One time using the daily squared returns, and one time using the hybrid method as our "true volatility" proxies.
3. When the error statistics for each model are compared, we will use the DM test, as described in section 2.2.3, to check if the difference between the error statistics is statistically significant. If the difference is significant, then the model with the lower error statistics will be ranked higher. If the difference is not significant, then both models will be ranked the same.
4. Models will be ranked from first place to last place for each error statistics test comparison performed. The first place model with the lowest error statistics will receive a score of 100 points, and the second place model will receive a score of 90 points, and so on. The last place model will receive a score of 60 points.
5. Error statistics comparison will be performed for both in-sample data and out-of-sample data. Therefore, a total of 16 error statistics comparison tests will be performed.
6. Both the out-of-sample and in-sample evaluations will be given the same weight.

7. The models will be ranked according to "in-sample rank", "out-of-sample rank" and "final average rank". The final average rankings for the models will be based on the total score averaged over all error statistics and all samples, including both in-sample and out-sample.

4.2.1 In-Sample Evaluation [6th January 2006 to 28th May 2010]

With these assumptions made, we can now begin our in-sample evaluation. The in-sample results obtained are displayed in table 4.1. From our results in table 4.1, the EGARCH(1,1) was clearly the best fit model, followed by the GARCH-S(1,1).This is because the EGARCH(1,1) model tends to have a lower mean on average as compared to other volatility models. The disadvantage of the EGARCH model can be seen from figure 4.1 and 4.2 where it fails to capture the "peakness" of the volatility during the subprime mortgage crisis. Therefore, the EGARCH(1,1) clearly underestimates errors during periods of extremely high volatility. This is because the EGARCH(1,1) follows its mean closer than other models.

Let us perform the same evaluation process, except instead of using MSE, RMSE, MAE,

In-sample	GARCH (1,1)	GJR-GARCH (1,1)	Mod-GARCH (1,1)	GARCH-S (1,1)	EGARCH (1,1)
MSE (squared returns)	70	100	60	100	100
RMSE (squared returns)	70	90	60	90	100
MAE (squared returns)	70	100	60	100	100
MAPE (squared returns)	80	60	90	100	70
MSE (Hybrid)	90	90	60	70	100
RMSE (Hybrid)	90	90	60	70	100
MAE (Hybrid)	90	90	60	80	100
MAPE (Hybrid)	70	90	60	80	100
Total	630	710	510	690	770
Average	78.75	88.75	63.75	86.25	96.25

Table 4.1 In-sample forecast evaluation

In-sample	GARCH (1,1)	GJR-GARCH (1,1)	Mod-GARCH (1,1)	GARCH-S (1,1)	EGARCH (1,1)
Linex (squared returns)	60	90	80	100	70
Linex (Hybrid)	80	70	90	100	60
Total	140	160	170	200	130
Average	70	80	85	100	65

Table 4.2 Results from the linex cross function

MAPE as our error statistics, we will use the linex asymmetric loss function, from section 2.2.2, as our error statistic. For this in-sample comparison, the parameter a of the linex asymmetric loss function will be set to 250. This means that compared to overestimated errors of the same magnitude, underestimation errors will be much more severely penalized. A value of 250 is quite an exaggerated value but since there are such few points of high volatility as compared to points of low volatility, the penalty for underestimating volatility should be much more severe.

From the linex loss function results displayed in table 4.2, we can see that the GARCH-S(1,1) is the model with the least underestimation errors. While, the EGARCH(1,1) and GARCH(1,1) had the most underestimation errors. The EGARCH(1,1) and GARCH(1,1) did not perform well when modeling large volatilities. Using figures 4.1, 4.2, and the results from the linex loss function, we can conclude that for modeling large periods of volatility such as the subprime mortgage crisis, the EGARCH(1,1) and GARCH(1,1) might not be as suitable as other models as it underestimates the volatility.

Therefore, in general, we still consider the GARCH-S(1,1) model to be a better model. This is because the GARCH-S(1,1) clearly is able to capture the "peakness" of the volatility during the subprime mortgage crisis while not compensating much on the overall forecast accuracy.

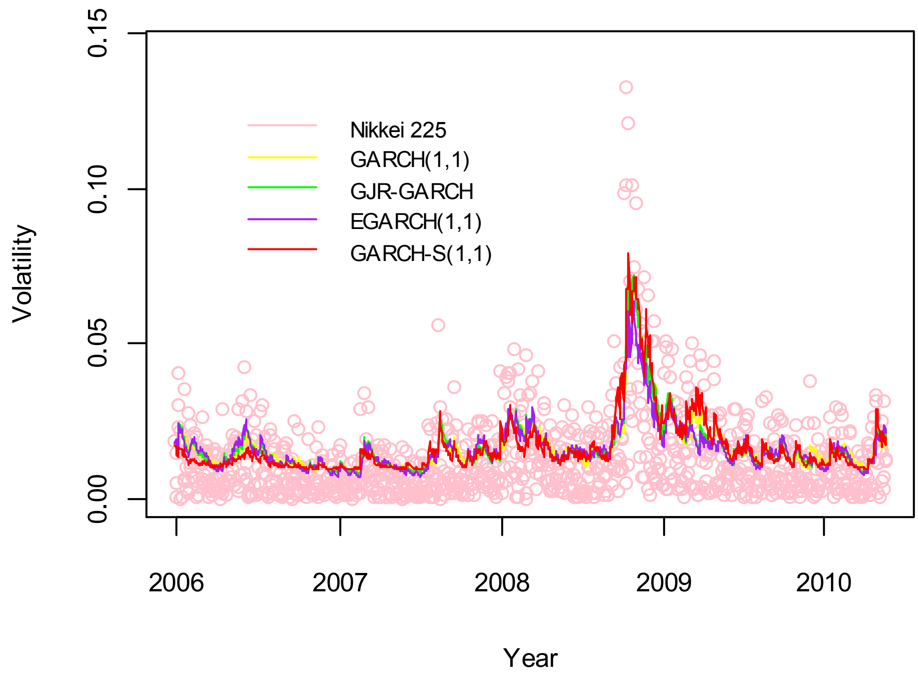

Figure 4.1 In-sample plot of GARCH(1,1), GJR-GARCH(1,1), EGARCH(1,1), GARCH-S (1,1) volatility against the daily squared returns as a measure of "true volatility"

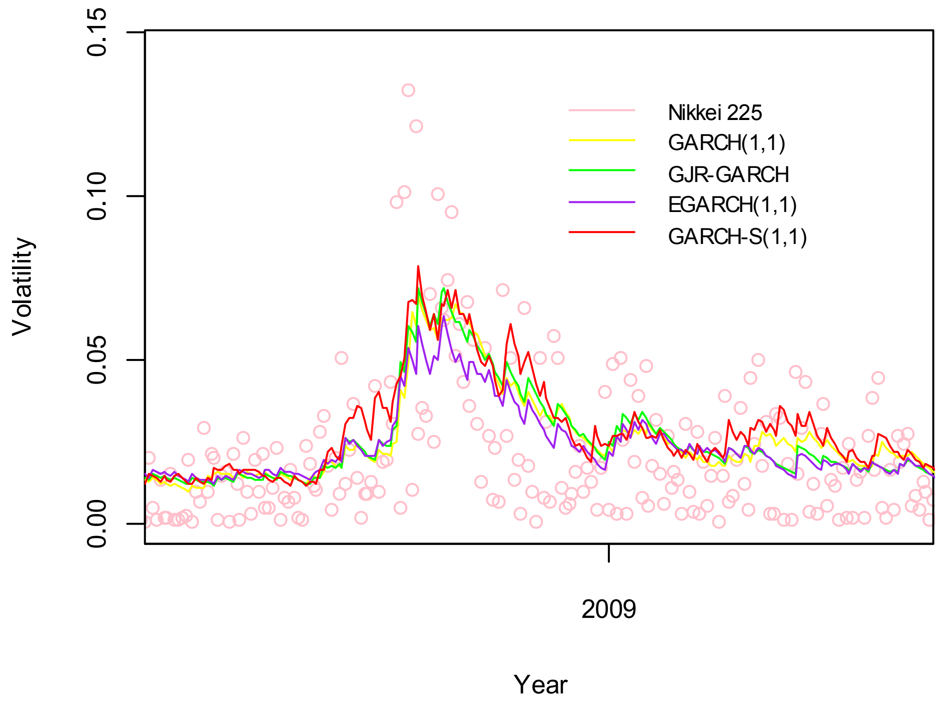

Figure 4.2 Enlarged plot of figure 4.1 focusing on the subprime mortgage crisis

4.2.2 Out-of-Sample Evaluation [29th May 2010 to 28th January 2011]

We calculated the out-of-sample volatility forecast of the models using the parameter values estimated from our in-sample data. The volatility forecasts and the squared daily returns "true volatility" were plotted in figure 4.3. The y-axis scale limit is purposely set to be the same as our previous graphs in this section. This is done to contrast the volatility levels between the figures 4.3 and figures 4.2, 4.1. From figure 4.3, we can see that this out-of-sample time period is relatively short. Furthermore, there is not much movement of volatility during this period of time. Therefore, under these conditions, even small random noises might cause a ranking of the model to go up or down. Therefore, our evaluation results should be interpreted carefully.

From the results in table 4.3, we can see that the GARCH-S(1,1) fit has the best fit for this small period of out-of-sample. From figure 4.3, we can see that there are no periods of high volatility in our out-of-sample data. In fact, all the respective volatility remains relatively the same throughout the time period. Therefore, there is no point in implementing the linex asymmetric lose function for our out-of-sample data.

From the final average results displayed in table 4.4, the first placed model was the GARCH-S(1,1) and the second place model was the GJR-GARCH(1,1). But these fore-

Out-of-Sample	GARCH (1,1)	GJR-GARCH (1,1)	Mod-GARCH (1,1)	GARCH-S (1,1)	EGARCH (1,1)
MSE (squared returns)	80	80	90	100	60
RMSE (squared returns)	80	70	90	100	60
MAE (squared returns)	60	80	90	100	60
MAPE (squared returns)	70	80	90	100	60
MSE (Hybrid)	80	80	100	100	60
RMSE (Hybrid)	80	70	90	100	60
MAE (Hybrid)	70	80	90	100	60
MAPE (Hybrid)	70	80	90	100	60
Total	590	620	730	800	480
Average	73.75	77.5	91.25	100	60

Table 4.3 Results from the in-sample evaluation

	GARCH (1,1)	GJR-GARCH (1,1)	Mod-GARCH (1,1)	GARCH-S (1,1)	EGARCH (1,1)
Total	1220	1330	1240	1490	1250
Average	76.25	83.125	77.5	93.125	78.125

Table 4.4 Final average results for both in-sample and out-of-sample evaluation

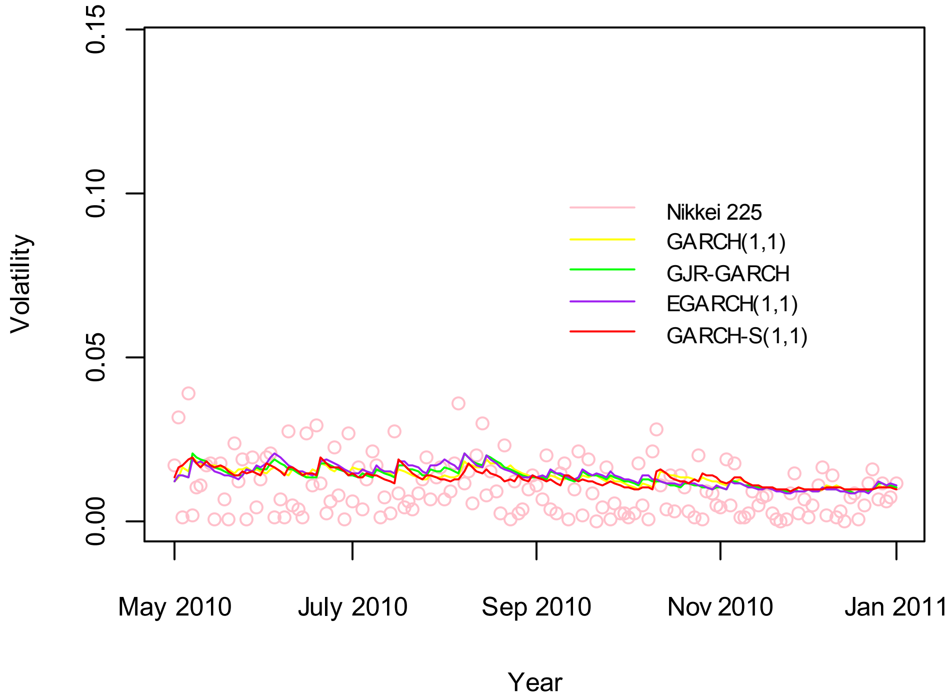

Figure 4.3 Out-of-sample plot of GARCH(1,1), GJR-GARCH(1,1), EGARCH(1,1), GARCH-S (1,1) volatility against the daily squared returns as a measure of "true volatility"

cast results can only be made after several assumptions were made. Therefore, the results cannot be generalized. Furthermore, there is a lot of noise due to our procedure used in forecast evaluations. Therefore, a conclusive conclusion about the forecasting accuracies cannot be assumed. This is rather normal in forecasting accuracy [Loudon, Watt, Yadav (2000)] [Brooks(1996)]. Besides, all forecasting accuracy literature have several assumptions made about which markets, which time frame, and which models to compare.

However, we can make two extremely important conclusions about the forecasting accuracy of the GARCH-S as compared to other models,

1. During high periods of volatility, the GARCH-S tends to capture the "peak" of the volatility more than other models. It greatly reduces underestimation errors.

Therefore, the GARCH-S model can be seen as an "upper range" time series volatility model.

2. When there are no high periods of volatility, the GARCH-S does not differ much from other models. Therefore, the GARCH-S provides a competitive accuracy when compared to other GARCH type models.

The GARCH-S reduces underestimation errors especially during a financial crisis involving the secondary market. However, due to the fact that it reduces underestimation errors, overestimation errors increase. Since most "true volatility" occur in the low range, the GARCH-S has a greater error statistics than other models. Therefore, it would be rather unfair to compare the forecasting accuracy of the GARCH-S model against other models as the GARCH-S model can been seen as an "upper range" model of the forecasting accuracy.

4.3 Evaluation from a Different Perspective

As stated in section 4.2, it would be unfair to compare to compare the GARCH-S model against other GARCH type models as the GARCH-S model can been seen as an "upper range" model of the "true volatility". Therefore, when evaluating error statistics, the GARCH-S model have a greater difference with the mean of the "true volatility" as compared to other GARCH type models.

However, for risk management and other general purposes, using the GARCH-S model might in fact provide an advantage over other GARCH type models. In this section, we will evaluate whether the GARCH-S model does indeed add in new information as compared to other GARCH type volatility models.

4.3.1 Regression-based method

Before we being, we must note that the four fundamental requirements for using linear regression models are linearity, independence, homoskedascity, and normality. The regression-based method for evaluating time series models will violate homoskedascity, but not the others.

In order to compare the informational content of forecasts made by the models, we need to use the regression-based method. The method involves regressing the actual volatility, X_t, on the forecasts made by the models, $\hat{X}_t$. The regression equation is defined by,

$$X_t = \alpha + \beta \hat{X}_t + \varepsilon_t . \tag{4.3}$$

where α is the intercept and β is the coefficient for $\hat{X}_t$, and ε_t are the residuals of the regression. The prediction is unbiased only if $\alpha = 0$, and $\beta = 1$. However, one important fundamental requirements for any regression is homoskedasticity. A violation of homoskedascity will lead to imperfect standard errors.

Since the error term, ε_t, is heteroskedastic and serially correlated when overlapping forecasts are evaluated, the standard errors of the parameter estimates are often computed on the basis of [Hansen, Hodrick (1980)]. Let Y be the row matrix of regressors including the constant term. From equation 4.3, $Y_t = [1 \;\; \hat{X}_t]$ is a 1×2 matrix. Then,

$$\begin{aligned}\hat{\psi} &= T^{-1}\sum_{t=1}^{T}\varepsilon_t^2 Y_t' Y_t \\ &+ T^{-1}\sum_{k=1}^{T}\sum_{t=k+1}^{T} Q(k,t)\varepsilon_k\varepsilon_t\left(Y_t' Y_k + Y_k' Y_t\right),\end{aligned} \tag{4.4}$$

where ε_t and ε_k are the residuals for observation t and k from the regression. The operator $Q(k,t)$ is an indicator function taking the value 1 if there is information overlap between Y_k and Y_t. The adjusted covariance matrix for the regression coefficients is then calculated as,

$$\hat{\Omega} = \left(Y'Y\right)^{-1}\hat{\psi}\left(Y'Y\right)^{-1}. \tag{4.5}$$

The standard errors of equation 2.4 would then be much closer to the correct standard deviation. However, in our research, we are more concerned with the adjusted R^2 value.

4.3.2 Out-of-Sample Evaluation [14th March 2008 to 28th May 2010]

Since the volatility for the time period around the American subprime mortgage crisis is the most interesting to us, we want to use that as our out-of-sample evaluation. Subsequently, we set our in-sample data to be from 6th January 2006 to 13th March 2008 and the out-of-sample time period to be from 14th March 2008 to 28th May 2010. We repeated the same tedious process described in chapter 3, and obtained the estimated parameters. The volatilities graphs obtained are displayed in figure 4.4.

Using the regression based method, the following rankings according to the adjusted R-squared were obtained,

1. GJR-GARCH(1,1) 0.3323
2. GARCH-S(1,1) 0.3227
3. Modified-GARCH(1,1) 0.3063
4. GARCH(1,1) 0.3025
5. EGARCH(1,1) 0.2789

This indicates that the GJR-GARCH(1,1) was able to provide the most information content for the mean of the "true volatility". The GARCH-S(1,1) was ranked second in providing information content for the "true volatility".

But what we really want to find is whether the GARCH-S(1,1) can add any more new information when added to the GJR-GARCH(1,1). When the GARCH-S(1,1) was added as a second term, the GARCH-S(1,1) did increase the adjusted R-squared value to 0.3350. In fact, the GARCH-S(1,1) did add information as a second regression term to any model. Therefore, the GARCH-S(1,1) is different from other models in providing information content that no other models could provide.

From figure 4.4, we can see that the GARCH-S(1,1) acts like an "upper range" of volatility. Therefore, the GARCH-S(1,1) model can provide this "upper range" of information that no other model can provide.

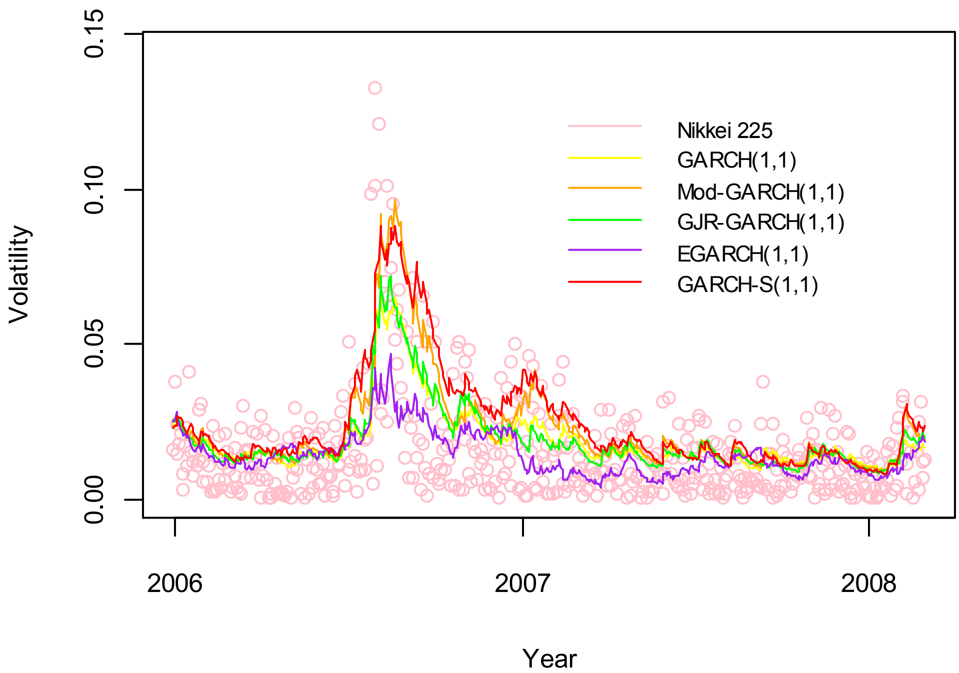

Figure 4.4 Out-of-sample evaluation for period from 14th March 2008 to 28th May 2010

4.4 Evaluation Summary

Time series forecasting is extremely difficult. Different markets, different time frames, different probability density functions might all provide different results. Even now, we still cannot be certain that any time series volatility model is comprehensively better than the GARCH (1,1) model [Lunde, Hansen (2005)].

Although finding the best forecasting accuracy model is nearly impossible since just a little bit of change could cause different results, we are able to make two important conclusions;

1. The GARCH-S model is able to capture the "peak" of high volatility, especially if there is a good correlation with the secondary market. This makes the GARCH-S model a good "upper range" volatility time series mode.
2. The GARCH-S has a comparable performance to the EGARCH, GJR-GARCH, and GARCH models when it is not capturing the "peak" of high volatility.

Therefore in summary, the GARCH-S model is an extremely good model since it could provide an "upper range" while still maintaining a decent forecast accuracy comparison against other GARCH type models. This makes the GARCH-S extremely good for risk management application purposes as it greatly reduces the number of underestimated errors. From figure 4.5, we can see that the GARCH-S clearly has the lowest number of underestimated errors (top left hand triangle, above the red line).

The weakness of the GARCH-S model is that it might not capture the mean of the "true volatility" as well as other models. This is because the GARCH-S is trying to map out the "peaks" of the "true volatility" and reverts back to the mean slower than other models. Therefore, if a user would like the "true volatility mean", the GARCH-S might not be suitable.

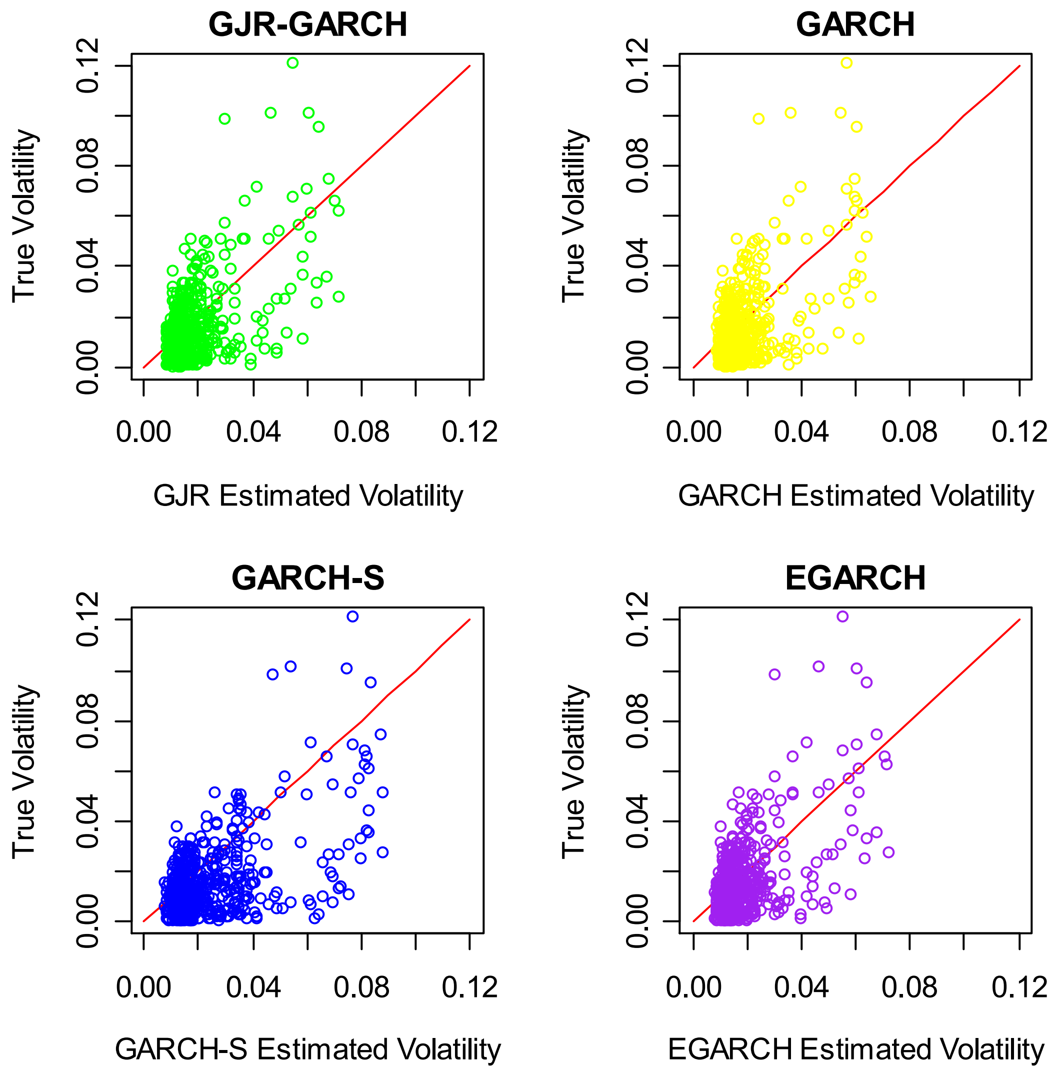

Figure 4.5 Comparison of forecasting errors. The errors above the red line indicates underestimation errors, while the errors below the red line indicates overestimation errors.

5 Financial Crisis

Financial crisis is defined as situations where financial institutions or assets suddenly lose a large part of their value. One of the largest financial crisis we have ever faced was the American subprime mortgage crisis of 2008 and 2009. Almost all developed countries experienced large negative gross domestic product (GDP) growth rates. Hundreds of banks around the world went bankrupt, numerous people lost money. Therefore, a financial crisis is devastating for everyone around the world.

This chapter will simulate a financial crisis, and demonstrate how the GARCH-S model could be used as a tool for risk management.

The period of time frame used in this chapter will be from 29th May 2008 to 1st February 2011.

5.1 Introduction to Multivariate Time Series Models

As countries around the world develop, their financial markets become more and more integrated with one another. Economic globalization and internet communication have accelerated the integration of the world's financial markets. To study the relationship between financial markets around the world, we could use multivariate time series analysis. The typical input for a multivariate time series consist of a matrix of single component series. For this chapter, we will use the Nikkei 225 and the Shanghai

A multivariate time series consists of single component series. Therefore, using matrices of these components, we could map out a relationship between the returns of the single component series involved.

5.1.1 Vector Autoregressive Models

The vector autoregressive model (VAR) is a simple vector model useful in modeling asset returns. The VAR(2) is defined as,

$$r_t = \phi_0 + \Phi r_{t-1} + \theta r_{t-2} + a_t \,, \tag{5.1}$$

where r_t is a multivariate time series r_t , ϕ_0 is a k -dimensional tor, Φ and θ are $k \times k$ matrices, and $\{a_t\}$ is a sequence of serially uncorrelated random vectors with mean zero and covariance matrix Σ. In application, the covariance matrix Σ is required to be positive definite; otherwise, the dimension of r_t can be reduced. For our research, we will simply use the bivariate case, where the vector matrix consists of two assets. The VAR(2) model can be written as,

$$\begin{aligned} r_{1t} &= \phi_{10} + \Phi_{11} r_{1,t-1} + \Phi_{12} r_{2,t-1} + \theta_{11} r_{1,t-1} + \theta_{12} r_{2,t-1} + a_{1t} \\ r_{1t} &= \phi_{20} + \Phi_{21} r_{1,t-1} + \Phi_{22} r_{2,t-1} + \theta_{21} r_{1,t-1} + \theta_{22} r_{2,t-1} + a_{2t} \,, \end{aligned} \tag{5.2}$$

where Φ_{ij} is the (i,j)th element of Φ , θ_{ij} is the (i,j)th element of θ , and ϕ_{i0} is the ith element of ϕ_0. Based on the first equation, Φ_{12} denotes the linear dependence of r_{1t} on $r_{2,t-1}$ in the presence of $r_{1,t-1}$. Therefore, Φ_{12} is the conditional effect of $r_{2,t-1}$ on r_{1t} given $r_{1,t-1}$. If $\Phi_{12} = 0$, then r_{1t} does not depend on $r_{2,t-1}$, and the model shows that r_{1t} only depends on its own past. Similarly, if $\Phi_{21} = 0$, then the second equation shows that r_{2t} does not depend on $r_{1,t-1}$ when $r_{2,t-1}$ is given. Consider the two equations jointly. If $\Phi_{12} = 0$ and $\Phi_{21} \neq 0$, then there is a unidirectional relationship from r_{1t} to r_{2t}. If $\Phi_{12} = \Phi_{21} = 0$, the r_{1t} and r_{2t} are uncoupled. If $\Phi_{12} \neq 0$ and $\Phi_{21} \neq 0$, then there is a feedback relationship between the two series.

5.1.2 Nikkei 225 and the FTSE China 25 Index Fund

China is rapidly developing and is already considered an economy giant. There is no doubt that China has an increasing influence on financial markets all over the world. Therefore, it will be interesting to study the relationship of the Chinese financial market with the Nikkei 225.

One major weakness of the GARCH-S model is that it requires the secondary market to be time lagged behind the primary market. If both markets were to have operating hours that is approximately around the same time, then the GARCH-S will not be effective. This is because whatever happens in the secondary market will affect the primary market instantly. Thus, we cannot make use of the secondary market returns at time $t-1$ to model the primary market's volatility at time t.

The Chinese financial markets' operating hours are approximately at the same time as the Japanese financial markets. Therefore, the GARCH-S model would be rather ineffective in this case. However, this is a solution to this problem. We could use a fund or index that is focused on the Chinese market but listed on the American stock markets. Such funds allow American investors to invest directly in China more safely (tighter regulations) through using the American stock markets. Such Chinese funds are excellent at depicting the health of the Chinese financial economy. This is because if investors feel the Chinese economy is losing value, the values of these Chinese funds will go down, and vice versa if the Chinese economy is gaining value.

The fund that we selected is the (Financial Times Stock Exchange) FTSE China 25 index fund. This fund basically follows the value of some of the biggest companies in China such as China life insurance.

We choose the VAR(2) because it had the lowest AIC value of all the combinations of VAR(p) models. The VAR(2) parameters were estimated using the VAR function from the vars package in R. The following estimates were obtained for the returns of Nikkei 225 (we are concerned with only the upper formulae in equation 5.2),

Model:

VAR(2) with Nikkei 225 as first input and FTSE China 25 as second input

Parameters	Estimate	Std. Error	t value	Pr（>|t|）
> Φ_{11}	-0.25	0.040	-6.20	1.01e-09
> θ_{11}	0.37	0.020	18.05	<2.22e-16
> Φ_{12}	-0.030	0.033	-0.91	0.36
> θ_{12}	0.15	0.024	6.31	5.26e-10
> ϕ_{10}	-0.00056	0.00069	-0.81	0.42

From the model p-values, we can see that the Φ_{12} and ϕ_{10} parameters are statistically insignificant. Therefore, we can ignore the terms and assume them to be 0. Therefore the VAR(2) states that,

$$\begin{aligned}\text{Nikkei return}_t &= -0.25\,\text{Nikkei return}_{t-1} + 0.37\,\text{FTSE China}_{t-1} \\ &\quad + 0.15\,\text{FTSE China}_{t-2} + a_{1t}\end{aligned} \tag{5.3}$$

5.2 Financial Crisis Simulation

We begin by modeling our GARCH-S model for our in-sample data from 29th May 2008 to 1st February 2011. The Nikkei 225 is the primary market and the FTSE China 25 is the secondary market. We repeated the tedious GARCH-S modeling process as described in chapter 3. The δ we chose is 0.04. The following estimation was obtained,

Model:

GARCH-S(1,1) Nikkei 225 (Primary market) and FTSE China 25 (Secondary market)

Parameters	Estimate	Std. Error	t value	Pr（>|t|）
> ω	4.17e-05	0.12e-04	3.39	0.00070
> α_1	2.78e-02	0.036	0.77	0.44
> β_1	7.12e-01	0.061	11.67	<2.22e-16
> γ_1	6.67e-02	0.020	3.34	0.00084

From our p-values, we can see that the α_1 parameter is statistically insignificant. Since the value of this term is so small, we are not extremely concerned about this. Therefore although we could ignore this term, it might be better to keep it.

Let us assume that a financial crisis in China happened. Although highly unlikely, the "asset bubble" of China "burst". The Chinese economy is in ruins, and the FTSE China 25 index fund, of course loses a huge amount of value. In order to simulate a decrease in the value of the FTSE China 25 index fund, we used the Box-Muller method as our random generator. The Box-Muller method is defined as,

$$z_1 = \sqrt{-2\ln x_1}\cos(2\pi x_2) \tag{5.4}$$

where x_1 and x_2 are uniformly and independently distributed between 0 and 1, then z_1 will have a normal distribution with mean $\mu = 0$ and variance $\sigma^2 = 1$.

We proceed to simulating our "financial crisis" and used a stochastic equation with a drift of around -0.15, and a multiplier of 1.2 for the stochastic random variable. This is definitely not the best way to simulate a "financial crisis". However, our goal is to model the variance. Therefore, any simulation that duplicates a "financial crisis" is acceptable. Figure 5.1 displays the FTSE index values in our simulated "financial crisis".

We now use our VAR(2) model to model the corresponding Nikkei 225 returns during this simulated "financial crisis". However, this is an extremely primitive simulation, and should not be used in proper simulations. In fact, the simulation is "wrong" because we have ignored the random variable term of a_t in equation 5.3. As stated before, our goal is not about simulation accuracy but about using the GARCH-S model to model the corresponding volatility. Using our estimated parameters from our in-sample data, we now model the out-of-sample financial crisis.

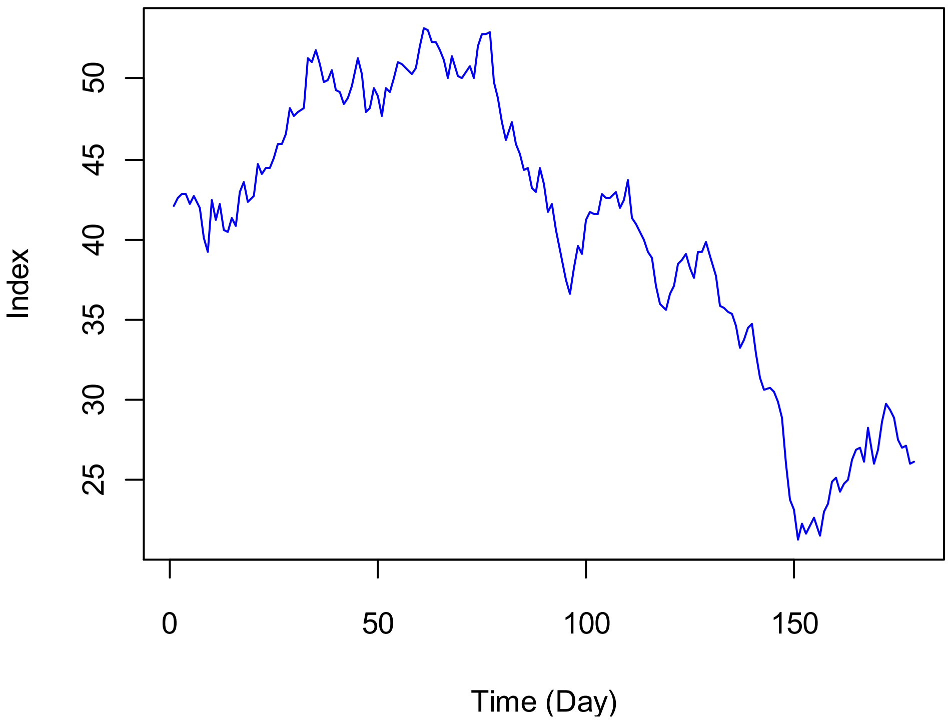

Figure 5.1 Index values for the FTSE China 25 fund in the case of a "financial crisis"

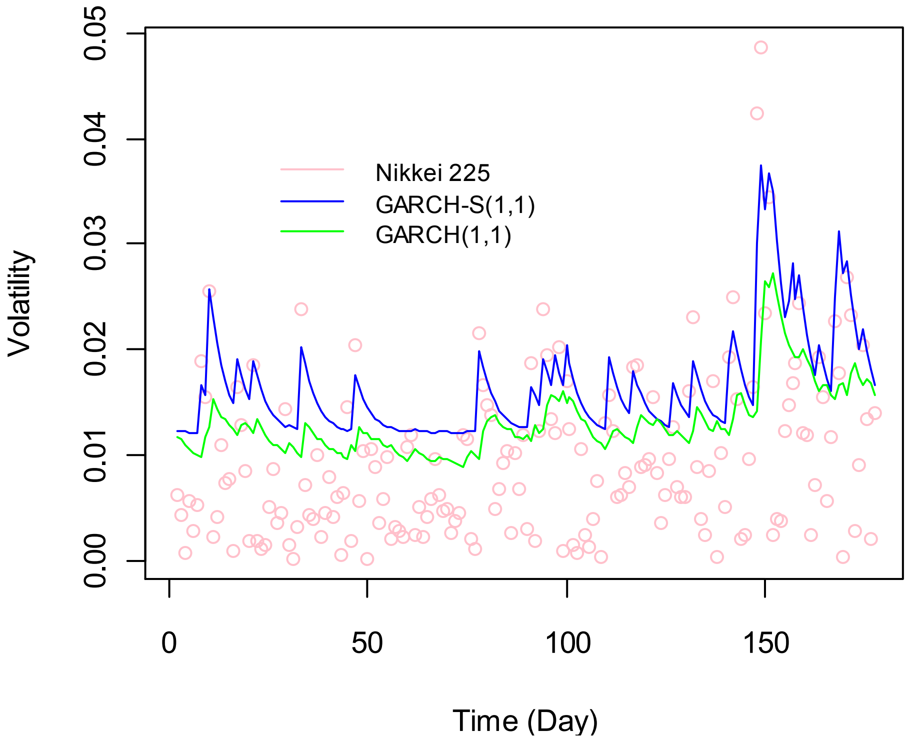

Figure 5.2 The GARCH-S(1,1) and GARCH(1,1) volatility model for the Nikkei 225 in the case of a Chinese "financial crisis"

Due to our extremely simplified simulated of the Nikkei 225 returns, figure 5.2 of the GARCH-S(1,1) volatility of the Nikkei 225 in a "financial crisis" simulation does not seem to be realistic at all. However, the important part of figure 5.2, is that we can see the GARCH-S(1,1) being used as an "upper range" volatility time series model.

5.3 Risk Management using Value at Risk

This section involves a simple demonstration of a practical application of the GARCH-S (1,1) model. Let us assume that the Nikkei 225 is an actual stock that can be bought or sold over the stock market. Let us also assume that we have a portfolio of 100000 USD worth of Nikkei 225 stock throughout the financial crisis. By using the GARCH-S(1,1) volatility, we can plot out our 5% VaR (value-at-risk) which is defined as,

$$100000\text{USD} \times (1.65 \times \text{GARCH-S}_t(1,1)\ \text{Volatility}) \tag{5.5}$$

This is a rather simplified VaR model, but our goal here is to demonstrate the practical application of the GARCH-S(1,1) model. Figure 5.3 displays the corresponding plot of the VaR given by the GARCH-S(1,1) of a 100000USD portfolio. We can expect only a 0.05 probability that our portfolio will fall below the VaR red line. By using the VaR values, banks can adjust their assets; investors can adjust their portfolios appropriately to the level of risk there are comfortable with taking. This is of course an extremely simplified example. There are more complex VaR models, and the real world stock market's movements are impossible to simulate. However, the GARCH-S(1,1) model could still be applied as an upper range time series volatility model.

In conclusion, we have successfully demonstrated a practical application of the GARCH-S(1,1).

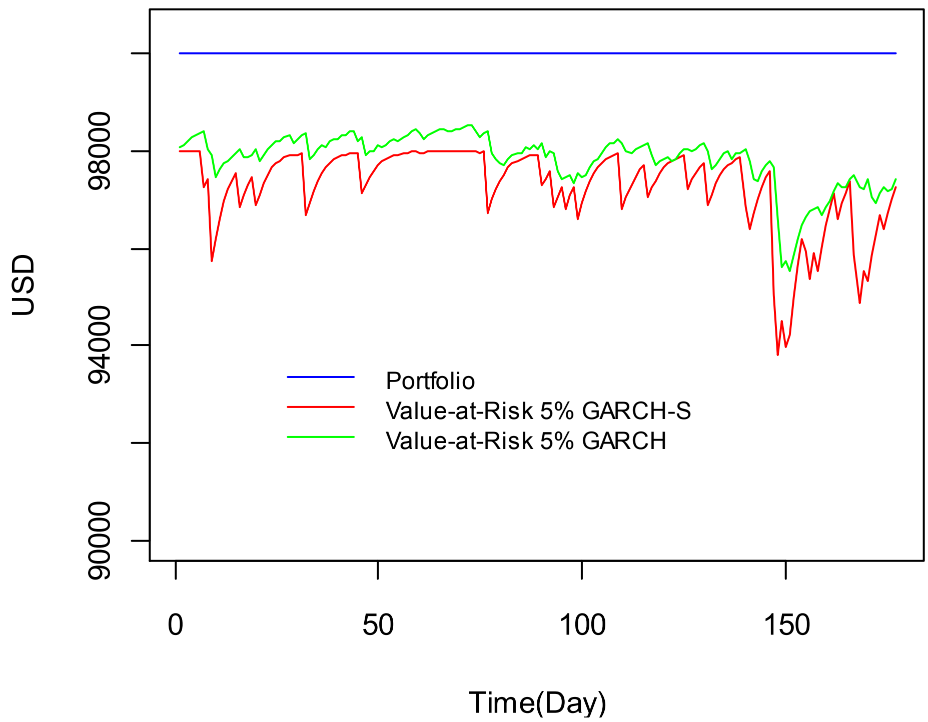

Figure 5.3 The 5% GARCH-S(1,1) VaR (red) and 5% GARCH VaR (green) for a 100000USD portfolio (blue) during a simulated "financial crisis"

6 Conclusion

In chapter 3, we derived our GARCH-S(1,1) model and applied it in chapter 4 and 5. Even though forecasting evaluation is a haphazard and difficult task due to the numerous number of assumptions and randomness (section 4.1) involved, we concluded that the GARCH-S(1,1) is generally a good volatility model to use in risk management. In chapter 5, we successfully demonstrated the GARCH-S(1,1) as an “upper range” time series volatility that could be used to calculate the VaR of a portfolio.

The advantages of the GARCH-S are,

1. The GARCH-S model acts as an “upper range” estimate of the “true volatility” as compared to other GARCH type models. It greatly reduces underestimate errors as compared to other GARCH type models (figure 4.4). It was ranked first when we ranked the models using the linex loss function with the a parameter as 250.
2. The GARCH-S model does comparably well against other GARCH type models even though it drifts away from the mean “true volatility” to model periods of high volatility. It was the second best ranked model in both our error statistics and regression-based evaluations. However, the forecasting accuracy conclusions can only be applied after several assumptions have been made (section 4.1).
3. Due to the GARCH-S acting as an “upper range” time series volatility model, it is extremely effective as a risk management tool (chapter 5).

The GARCH-S model also has several disadvantages,

1. When applying the GARCH-S model, we need two different markets that are operating in very different time zones (the American and the Japanese financial markets). The Japanese and Asian-Pacific markets are operating in approximately the same time zone, thereby decreasing the effectiveness of the GARCH-S model. To overcome this problem, we could use an index or fund that covers one of our input markets but is located in a different time zone. For example, the FTSE China 25 is a fund consisting of investment in 25 top Chinese companies, but is listed on the New

York Stock Exchange.

2. For the GARCH-S model to be truly effective, we need the secondary market (S&P 500,…) to be an influential market that influences the returns of the primary market (Nikkei 225). However, as the world economy advances, financial markets start to gain more and more influence on each other.
3. The GARCH-S is likely to be more effective if there is a strong correlation, preferably above 0.50, between the time t return of the target input market and time $t-1$ return of the secondary input market. However, during periods of high volatility, the correlation tends to be strong between markets (0.50).
4. The forecasting accuracy of the GARCH-S model is likely to be lower than other GARCH type models due to it acting as an "upper range" time series volatility model. This causes the GARCH-S model to overestimate volatility at the lower range. Since most volatility occurs at a lower range, the error for GARCH-S is comparatively larger than other GARCH type models.

In conclusion, the GARCH-S model is a new and innovative method for modeling the Nikkei 225 volatility. It is especially useful in risk management applications since it tends to map out the "upper range" of the true volatility. The underestimation errors for the GARCH-S are greatly reduced, but the overestimation errors increases. In general, it does perform comparable well against other models in forecasting accuracy. However there are still many limitations to the GARCH-S model such as not being able to generalize the GARCH-S model to all financial markets around the world.

Further research could be made on extending the univariate GARCH-S model to a multivariate time series. Another important further research could be to use the intraday returns of an asset to model realized volatility. This will increase the accuracy and greatly reduce biasness in forecasting evaluation.

Appendix

A.1 Introduction to GARCH Programming

In section 3.6.1, we proposed a modified GARCH model. In order to model this proposed modified GARCH model, we need to write our own function program. Our program is written in the R programming language. The R is an open source programming language that is becoming the de facto standard among statisticians, and graduate students for the development of statistical software. R is also free and widely available. Therefore, it has become one of the most widely used programming languages in Statistics.

The program we have written requires the use of the nlminb function from the R statistical software. The nlminb function uses optimization routines that were developed at AT&T Bell Laboratories [Gay (1990)]. The type of optimization used is a Newton type optimization, which itself is an augmented version of the Gauss-Newton algorithm. To a certain extent, it is important to choose initial values that are closer to the "true" parameters. This is because the closer the initial values are to the "true" parameters, the faster the optimization, and the better the accuracy of the estimated parameters.

A.1.1 Maximum Likelihood Estimation (MLE) Method

For our modified GARCH program, we assumed that the returns followed the Normal distribution as it is the easiest and one of the most commonly used distributions when estimating and forecasting GARCH models. Due to the problems mentioned in 1.4.3, we could certainly implement other distributions for better depiction of the returns' distribution. The probability density function (pdf) for the Normal distribution $\mathrm{N}(\mu,\sigma^2)$ is,

$$f\left(r \mid \mu,\sigma^2\right) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(r-\mu)^2}{2\sigma^2}\right). \tag{A.1}$$

Following section 3.3.1, where we define our returns and innovations as $r_t = \mu + \varepsilon_t$, and $\varepsilon_t = z_t\sqrt{h_t} = z_t\sigma_t$. Following this definition, we define the log-likelihood function for the normal distribution as,

$$\begin{aligned} \ln L(\theta \mid r_1, \ldots r_n) &= \ln f(r_1, r_2, \ldots r_n \mid \theta) \\ &= \ln \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma_t^2}} e^{-\frac{z_t^2}{2}} \\ &= -\frac{1}{2}\sum_n \left[\log(2\pi) + \log\left(\sigma_t^2\right) + z_t^2\right]. \end{aligned} \tag{A.2}$$

[Proof of this equation is in the Proofs section.] With this set up, we could now program the maximum log-likelihood function and use it to optimize the parameters of our modified GARCH model. [Please refer to the Programming section for the code.]

A.2 Advanced Parameter Estimation

In statistics, there are two main schools of thought. One is the traditional frequentist view of probability, and thus of statistical inference, is based on the idea of an experiment that can be repeated numerous times. The other is the Bayesian view of probability and of inference is based on the assessment of probability and on observations from a single performance of an experiment. The MLE method used to estimate our errors is part of the traditional frequentist statistics. In this section, we will discuss an alternative approach to the tradiaional frequentist statistics, Bayesian statistics.

A.2.1 WinBUGS and JAGS Statistical Packages

WinBUGS (Baysian estimation Using Gibbs Sampling) was developed at Cambridge University. It is an extremely simple to use program that can be called using the R statistical package, and best of all, it is free. The JAGS (Just Another Gibbs Sampling) program can be called through R using the R2jags function. The R2jags function allows WinBUGS users to direct write their code without having to adjust to the R2jags format. The WinBUGS and JAGS statistical packages are two of the most commonly used statistical software for Bayesian statistics.

A.2.2 Gibbs Sampling

When creating programs to estimate models, writing a MLE program can be tedious and time consuming. For Bayesian Statistics, all we need to do is to define our model, parameters, distribution functions, have a fast computer and we are able to estimate the parameters easily. The advantages of Bayesian estimations include,

1. Providing more intuitive and meaningful inferences. A Bayesian analysis can give a more meaning inference by stating the probability that the hypothesis is true. For the frequentist MLE method, we could only reject the null hypothesis.
2. Bayesian methods make use of all available information. This is because Bayesian estimation includes the prior information.

The Gibbs Sampling method is a Bayesian estimation method using Markov Chain Monte Carlo simulations. Simple homogenous Markov chains with a finite number K of states. Usually, we denote the state space as $S = \{1, 2, ..., K\}$. The Markov property is defined as,

$$\begin{aligned} p_{ij} &= P\{X_n = j \mid X_{n-1} = i, X_{n-2} = i_{n-2}, ..., X_1 = i_1\} \\ &= P\{X_n = j \mid X_{n-1} = i\}, \end{aligned} \tag{A.3}$$

for i, $i_{n-2}, ..., i_1$ and $j \in S$,and for $n = 1, 2, ...$ i and j are the states for the process. In simple terms the probability of making a transition from state i transition state j does not depend on previous states lagging behind state i.

The mathematics required for the Gibbs Sampling method is beyond the scope of this thesis, but in general, the Gibbs Sampling method is a computational method that uses Markov chains to approximate posterior distributions. The idea is to use available information about a prior distribution and data to construct an erogodic Markov chain whose limiting distribution is the desired posterior distribution. Simulation is done until enough steps of the chain are required to obtain a good approximation to the limiting distribution.

B Definitions

Kurtosis: Kurtosis is defined as the measure of the peak of a probability distribution. For a normal distribution, the kurtosis value is 3. If the peak is higher, then the kurtosis value would also be larger. Higher kurtosis also means that the variance of the probability distribution is a result of infrequent extreme deviations and not frequent normal deviations.

Innovation: The innovation, sometimes known as the error, is the remaining values after the mean equation values have been deducted from the return values of an asset. In equation form, it is defined as $\varepsilon_t = \mu - r_t$, where ε_t is the innovation, μ is the mean and r_t are the return values of an asset.

Out-of-sample: Out-of-sample is defined as using estimated parameters derived from a sample subset of the population to model a different sample subset of the same population.

Primary Market: The financial market which we would like to model.

Portmanteau Test: A type of statistical test which tests whether any groups of autocorrelations in a time series are different from zero. Instead of using randomness at each lag, it tests the entire randonmness according to each lag.

Secondary Market: The financial market which we would use in augmenting the volatility forecast of the primary market in the GARCH-S model.

Skewness: Skewness is defined as the measure of asymmetry in a probability distribution. A normal distribution has a skewness of 0. If the graph is skewed left, then the skewness would be a negative value, and vice versa.

Time Additive: Time additive is a property where the asset returns are spread linearly over time. For example, let us have an asset return of x for time period one and an asset return of y for time period two. The natural logarithm two time period return, which has the time additive property, will be equivalent to $x + y$.

C Proofs

Equation 3.6

To prove:

$$\frac{E(\varepsilon_t^4)}{\left[E(\varepsilon_t^2)\right]^2} = \frac{3\left[1-(\alpha_1+\beta_1)^2\right]}{1-(\alpha_1+\beta_1)^2-2\alpha_1^2} > 3$$

Proof:

$$\begin{aligned}
E(\varepsilon_t^4) &= E(z_t^4)\cdot E(h_t^2) \qquad \text{by definition} \\
&= 3\left[\alpha_0^2 + 2E(\varepsilon_t^2)\alpha_0(\alpha_1+\beta_1)\right]\cdot\left[1-\beta_1^2-2\alpha_1\beta_1-3\alpha_1^2\right]^{-1} \\
&= 3\left[\alpha_0^2 + 2\frac{\alpha_0^2}{1-\alpha_1-\beta_1}(\alpha_1+\beta_1)\right]\cdot\left[1-\beta_1^2-2\alpha_1\beta_1-3\alpha_1^2\right]^{-1} \\
&= 3\alpha_0^2\left[1+2\frac{\alpha_1+\beta_1}{1-\alpha_1-\beta_1}\right]\cdot\left[1-\beta_1^2-2\alpha_1\beta_1-3\alpha_1^2\right]^{-1} \\
&= 3\alpha_0^2(1+\alpha_1+\beta_1)\left[(1-\alpha_1-\beta_1)(1-\beta_1^2-2\alpha_1\beta_1-3\alpha_1^2)\right]^{-1}
\end{aligned}$$

Subsequently,

$$\begin{aligned}
\text{kurtosis} &= \frac{E(X_t^4)}{\left[E(X_t^2)\right]^2} \qquad \text{by definition} \\
&= \frac{3(1+\alpha_1+\beta_1)(1-\alpha_1-\beta_1)}{1-\beta_1^2-2\alpha_1\beta_1-3\alpha_1^2} \qquad \text{by substitution} \\
&= \frac{3\left[1-(\alpha_1+\beta_1)^2\right]}{1-(\alpha_1+\beta_1)^2-2\alpha_1^2}
\end{aligned}$$

Q.E.D

Equation A.2

To prove:

$$\ln\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi\sigma_t^2}}e^{-\frac{z_t^2}{2}}=-\frac{1}{2}\Sigma\left[\log\left(2\pi\right)+\log\left(\sigma_t^2\right)+z_t^2\right]$$

Proof:

$$\begin{aligned}\ln\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi\sigma_t^2}}e^{-\frac{z_t^2}{2}}&=\ln\left(\frac{1}{\sqrt{2\pi\sigma_1^2}}e^{-\frac{z_1^2}{2}}\times\frac{1}{\sqrt{2\pi\sigma_2^2}}e^{-\frac{z_2^2}{2}}\times\ldots\times\frac{1}{\sqrt{2\pi\sigma_n^2}}e^{-\frac{z_n^2}{2}}\right)\\&=\ln\left(\frac{1}{\sqrt{2\pi\sigma_1^2}}\right)+\ln\left(e^{-\frac{z_1^2}{2}}\right)+\ldots+\ln\left(\frac{1}{\sqrt{2\pi\sigma_n^2}}\right)+\ln\left(e^{-\frac{z_n^2}{2}}\right)\\&=-\frac{1}{2}\ln\left(2\pi\right)-\frac{1}{2}\ln\left(\sigma_1^2\right)-\frac{1}{2}z_1^2-\ldots-\frac{1}{2}\ln\left(2\pi\right)-\frac{1}{2}\ln\left(\sigma_n^2\right)-\frac{1}{2}z_n^2\\&=-\frac{1}{2}\sum_n\left[\log\left(2\pi\right)+\log\left(\sigma_t^2\right)+\log\left(z_t^2\right)\right]\end{aligned}$$

Q.E.D

Equation 3.8

To prove:

$$\hat{h}_{t+2} = \omega + \left(\alpha_1 + \beta_1\right)\hat{h}_{t+1}$$

Proof:

$\varepsilon_{t+1}^2 = z_{t+1}^2 h_{t+1}$ by definition

$$\begin{aligned} E(z_{t+1}^2) &= \mathrm{Cov}(z_{t+1}^2) + E\left[z_{t+1}\right]\cdot E\left[z_{t+1}\right] \quad \text{by definition of covariance} \\ &= \mathrm{Var}(z_{t+1}) \\ &= 1 \end{aligned}$$

Therefore,

$$\begin{aligned} E(\varepsilon_{t+1}^2) &= E(z_{t+1}^2 h_{t+1}) \\ &= h_{t+1} E(z_{t+1}^2) \quad \text{since variance at time } t+1 \text{ is constant} \\ &= h_{t+1} \end{aligned}$$

Therefore,

$$\begin{aligned} \hat{h}_{t+2} &= \omega + \alpha_1 \varepsilon_{t+1}^2 + \beta_1 \hat{h}_{t+1} \\ &= \omega + \left(\alpha_1 + \beta_1\right)\hat{h}_{t+1} \end{aligned}$$

Q.E.D

D Programming

Program for modified GARCH model using the log-normal MLE and the Newton type optimization,

```
##################Begin program for modified GARCH model#################

#Required package: Matrix
#Required package: stats

#Model is defined in section 3.6.1
ed.gjr.garch = function(x, y, delta)
{
# Step 1: Initialization of the time series data:
#x is the returns data that we would like to model with ed.gjr.garch model:
x <<- x;
#y is the returns data that we would like to use to augment the volatility
#model of x:
y <<- y;
#delta is the value defined in section 3.6.1 that you set for the ed.gjr.garch
#model:
delta<<- delta;

# Step 2: Initialization of the model parameters and boundaries:
#We declare T globally, where T is a variable that is used to identify number
#of loops the function will perform.
T <- length(x);
#Create an empty matrix for the errors in the ed.gjr,garch model:
e <- matrix(nrow=numeric(T) ,ncol=1);
#Creating an empty matrix for the conditional variance in the ed.gjr.garch
#model:
Var <- matrix(nrow=numeric(T),ncol=1);
```

```
#Finding mean of x:
Mean = mean(x);
# Initializing the first conditional variance term in the ed.gjr.garch
#model:
Var[1] = var(x);
#Creating an S for use in the boundary conditions:
S = 1e-6;
#Initializing the first error term in the garch.gjr.ed model
e[1]<-x[1]-Mean;

#Initializing the parameters. User can choose his/her own set of parameters.
#Important: The closer the parameters are to the "true" value, the faster
#and more accurate are the results!
params <- c(mu = Mean, omega = 0.01*Var, alpha = 0.01, beta = 0.1, gamma
=0.01);
#Set lower boundary:
ed.lowerbound <- c(mu = -10*abs(Mean), omega = S^2, alpha = S, beta = S,
gamma = S);
#Set upper boundary:
ed.upperbound <- c(mu = 10*abs(Mean), omega = 100*Var, alpha = 1-S, beta
= 1-S, gamma = 1-S);

# Step 3: Set the conditional distribution (Normal distribution) and write
#the function:
ed.garch.kaiyu = function(parm, iterate=TRUE)
{
#Setting the parameters:
mu = parm[1];
omega = parm[2];
alpha = parm[3];
beta = parm[4];
gamma = parm[5];
```

```
#Setting the likelihood function:
likelihood = 0;

#Main part of the code:
for(t in 2:T)
{
e[t] = x[t]-mu;
Var[t] =omega + alpha*e[t-1]^2 + beta*Var[t-1] + ifelse(y[t-1]>=delta,
gamma*(e[t-1]^2), 0);
#This is the log-likelihood function for a Normal distribution, the user
#could change it to #a different distribution on his/her preferences.
likelihood = likelihood - 0.5*log(2*pi*Var[t]) - 0.5*e[t]^2/Var[t];
}
#Returning the values of the code:
if(iterate) return(-likelihood)
else return(list(loglik=likelihood, sig2=Var, res=e/sqrt(Var)));
}

# Step 4: Optimizing the parameters and computing the numerical Hessian:
#Using nlminb function with trace of 3:
fit = nlminb(start = params, objective = ed.garch.kaiyu,
lower = ed.lowerbound, upper = ed.upperbound, control = list(trace=3),
hessian=TRUE)

#This is the Hessian code.
epsilon = 0.0001 * fit$par
Hessian <- matrix(0, ncol = 5, nrow = 5)
for (i in 1:5) {
for (j in 1:5) {
x1 = x2 = x3 = x4 = fit$par
x1[i] = x1[i] + epsilon[i]; x1[j] = x1[j] + epsilon[j]
x2[i] = x2[i] + epsilon[i]; x2[j] = x2[j] - epsilon[j]
```

```
x3[i] = x3[i] - epsilon[i]; x3[j] = x3[j] + epsilon[j]
x4[i] = x4[i] - epsilon[i]; x4[j] = x4[j] - epsilon[j]
Hessian[i,j]=
(ed.garch.kaiyu(x1)-ed.garch.kaiyu(x2)-ed.garch.kaiyu(x3)+ed.garch.kai
yu(x4)) /
(4*epsilon[i]*epsilon[j])
}
}

# Step 5: Printing our optimized parameters, t-values, p-values, etc…
#Solving the Hessian to find the standard errors.
se.coefficient = sqrt(diag(solve(Hessian)))

#Finding the t-values:
t.val = fit$par/se.coefficient

#Matrix to sort everything.
matcoefficient=cbind(fit$par,se.coefficient,t.val,2*
(1-pnorm(abs(t.val))))
dimnames(matcoefficient) = list(names(t.val), c(" Estimate", " Std. Error",
" t-value", "Pr(>|t|)"))
cat("¥nCoefficient(s):¥n")
printCoefmat(matcoefficient, digits = 8, signif.stars = TRUE)
}
```

###################End program for modified GARCH model#################

/Program might not work well under some conditions such as the Hessian being negative. Some initial values input might not work well with the nlminb function./

```
#############Sample code for parameters estimation using JASGS###########
/** Just the sample code for the GARCH-S mode will be written. **/

//Sample size
T=178;
//Input primary market
y=c(nikkeisp500.data$nikkeireturn[2:1790]);
//Input secondary market's absolute returns
x=c(abs(nikkeisp500.data$sp500return[2:1790]));
//Input secondary market innovations
c=c(nikkeisp500.data$sp500inn[2:1790]);

//Defining inputs
data=list("T","y","x","c")
//Defining parameters
parameters<-c("b1","b2","b3","b4")

//Initial values
inits1<-list(b1=0.000003,b2=0.08,b3=0.7,b4=0.1)
inits2<-list(b1=0.000005,b2=0.06,b3=0.8,b4=0.2)
inits<-list(inits1,inits2)

//Function to call up JAGS in R
garched.sim<-jags(data,inits,parameters,"final3.bug",n.chains=2,n.iter
=3000)


//final3.bug program
model{
//Defining initial prior distributions for parameters
b1~dnorm(0,0.01)
is.censored[1]~dinterval(b2,0)
b2~dnorm(0,0.05)
is.censored[2]~dinterval(b3,0)
b3~dnorm(0,0.2)
b4~dnorm(0,0.05)
```

```
//Initializing the first [1] volatility used in the GARCH-S
h[1]<-0.000452

//GARCH-S Model
for(t in 2:T)
{
//Returns are assumed to be Guassian distributed.
//Note: WinBUGS and JAGS assume variance to be (1/variance value entered
//in dnorm)
y[t]~dnorm(0,P[t])
//GARCH-S definition y is the Nikkei 225 return, h is the Nikkei 225
//conditional volatility, and x is the S&P 500 return, c is the S&P 500
//innovations, and delta is given a value of 0.015.
h[t]<-b1+b2*pow(y[t-1],2)+b3*h[t-1]+step(x[t-1]-0.015)*b4*(pow(y[t-1],
2)+pow(c[t-1],2))
P[t]<-1/h[t]
}
}
```

```
#################End parameters estimation using JASGS#################
```

References

Akaike, H. (1973) Information theory and an extension of the maximum likelihood principle. Proc. 2nd International Symposium on Information Theory, 267–281, Budapest.

Andersen, T.G., and Bollerslev T. (1997) Intraday Periodicity and Volatility Persistence in Financial Markets, *Journal of Empirical Finance*, 4, 115–58.

Bekiros, S.D., and Georgoutsos, D.A. (2006) Estimating the Correlation of International Equity Markets with Multivariate Extreme and GARCH Models, Universiteit van Amsterdam, Center for Nonlinear Dynamics in Economics and Finance, CeNDEF Working paper, 6–17.

Berg, A., Meyer R., Yu J. (2003) Deviance Information Criterion for Comparing Stochastic Volatility Models. *Journal of Business and Economic Statistics,* Jan 2004, 107–120.

Blair, B., Poon, S.H., and Taylor, S.H. (2001) Forecasting S&P 100 volatility: The incremental information content of implied volatilities and high frequency index returns, *Journal of Econometrics,* 105, 5–26.

Black, F. (1976) Studies in stock price volatility changes, *Proceedings of the 1976 Meeting of the Business and Economic Statistics Section, American Statistical Association*, 177–181.

Bollerslev, T. (1986) Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*, 31, 307–328.

Bollen, B., and B. Inder (2002) Estimating daily volatility in financial markets utilizing intraday data, *Journal of Empirical Finance*, 9, 551–562.

Cao, C.Q., and Tsay, R.S. (1992) Nonlinear time-series analysis of stock volatilities, *Journal of Applied Econometrics*, December, Supplement, 1, 165–185.

Casella G., Berger R.L. (1987). Reconciling Bayesian and frequentist evidence in the one-sided testing problem. *Journal of the American Statistical Association*, 82, 106–111.

Chenyao, D., Doganoglu, T., Mittnik, S., and Rachev, S.T. (1999), Maximum Likelihood Estimation of Paretian Models, *Mathematical and Computer Modeling,* 29, 235–240.

Chalabi, Y., Luksan, L., and Wurtz, D. (2006) Parameter Estimation of ARMA Models with GARCH/ APARCH Errors: An R and SPlus Software Implementation, Working paper, Institue for Theoretical Physics Swiss Federal Institute of Technology, Zurich.

Cowpertwait, P., and Metcalfe, A. (2009) Introductory Time Series with R, Springer New York.

Curto, J.D., Pinto, J.C., and Tavares, G.N. (2009) Modeling stock markets' volatility using GARCH Models with Normal, Student's t and stable paretian distributions, *Statistical Papers* 50, 311–321.

Dempster, A.P. (1997) The direct use of likelihood for significance testing (reprinted), *Sta-*

tistics and Computing, 7, 247–252.

Diebold, F.X. and Mariano, R.S. (1995) Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13, 253–263.

Ding, Z., Engle, R.F., Granger C.W.J. (1993) A long memory property of stock market returns and a new model, *Journal of Empirical Finance*, 1, 83–106

Engle, R.F. (1982), Autoregressive Conditional Heteroscedasticity with Estimates of the Variation of United Kingdom Inflation, *Econometrica* 50*,* 4*,* 987–1007.

Engle, R.F. (2001) GARCH 101: The Use of ARCH/GARCH Models in Applied Econometrics, *Journal of Economic Perspectives*,15, 4, 157–168.

Engle, R.F., and Patton, A. (2001), What good is a volatility model? *Quantitative Finance*, 1, 2, 237–245

Gay, D.M. (1990) Computing Science Technical Report No. 153, Usage Summary for Selected Optimization Routines, AT&T bell Laboratories.

Glosten, L.R., Jagannathan, R., and Runkle, D.E. (1993) On the relation between the expected value and the volatility of the nominal excess return on stocks, *Journal of Finance*, 48, 1779–1801.

Granger, C.W.J. (1999) *Empirical Modeling in Economics: Specification and Evaluation*, Cambridge University Press, New York.

Hansen, L.P., and Hodrick, R.J. (1980) Forward exchange rates as optimal predictors of future spot rates: An econometric analysis, *Journal of Political Economy*, 88, 829–853

Heyen, R.C., and Kat H.M. (1994) Volatility prediction: A comparison of stochastic volatility, GARCH (1, 1) and EGARCH (1, 1) models, *Journal of Derivatives*, 2, 50–65

Hull, J. (2002) *Options, Futures and other Derivatives*, Fifth Edition, Prentice Hall, Englewoods Cliffs, NJ

Lai, T.L., and Xing, H. (2008) Statistical Models and Methods for Financial Markets, Springer Texts in Statistics, New York.

Lunde, A., and Hansen P.R. (2005) A forecast comparison of volatility models: does anything beat a GARCH (1, 1)?, *Journal of Applied Econometrics*, 20, 7, 873–889.

McAleer, M., Verhoeven, P. (2003), Fat Tails and Asymmetry in Financial Volatility Models, Center for International Research on the Japanese Economy at the University of Tokyo.

Nelson, D.B. (1991) Conditional Heteroskedascity in Asset Returns: A new approach, *Econometrica*, 59, 2, 347–350.

Nelson, D.B., and Cao, C.Q. (1992) Inequality constraints in the univariate GARCH model, *Journal of Business and Economic Statistics*, 10, 2, 229–235.

Pagan, A.R., and Schwert, G.W. (1990) Alternative models for conditional models for conditional stock volatility, *Journal of Econometrics*, 45, 1–2, 267–290.

Poon, Ser-huang (2005) *A Practical guide to Forecasting financial market volatility*, Wiley Finance, London.

Posedel, P. (2005), Properties and Estimation of GARCH (1,1) Model, *Metodoloski Zvezki – Advances in Methodology and Statistics*, 3, 2, 267–287

Sharpe, W.F. (1994) The Sharpe Ratio, *The Journal of Portfolio Management*, Fall 1994, 49–58.

Suess, E.A., Trumbo B.E. (2010) *Introduction to Probability Simulation and Gibbs Sampling with R*, Springer, New York.

Taylor, S.J. (1986) *Modelling Financial Time Series*, John Wiley & Sons Ltd, Chichester.

Tsay, R.S. (2010). *Analysis of Financial Times Series*, Third Edition, Wiley, New Jersey.

Wang, J., Zivot, E. (2005). *Modeling Financial Time Series with S-PLUS*, Second Edition, Springer, New York.