LUND UNIVERSITY
School of Economics and Management

Master Thesis, Spring 2009
School of Economics and Management
Lund University

Forecasting Volatility

An Empirical Investigation of Implied Volatility and Its Information Content

Authors
Linus Mattsson
Johan Möse

Supervisor
Birger Nilsson

Abstract

Title:	Forecasting Volatility - An Empirical Investigation of Implied Volatility and Its Information Content
Final seminar:	June 4, 2009
Course:	NEKM03 Master Thesis - Finance Programme
Authors:	Linus Mattsson Johan Möse
Supervisor:	Associate Professor Birger Nilsson Department of Economics, Lund University
Keywords:	Volatility forecast, implied volatility, model based forecast, predictive power, incremental information
Purpose:	The purpose of this thesis is to evaluate volatility forecasts by testing the predictive power of implied volatility vis-à-vis model based forecasts. Furthermore we test if implied volatility contains any additional information beyond that captured by the model based forecasts.
Methodology:	A number of time series models are fitted to historical data. The fitted models are then used to forecast volatility. The procedure is repeated to produce a series of forecasts. The forecast are evaluated against out-of-sample realized volatility through regression analysis. Finally we test for additional information in implied volatility through GMM and OLS estimation.
Results:	We find that volatility can be predicted to some extent. Tests indicate that implied volatility is the superior forecast of future realized volatility when compared bilaterally against time series models. Implied volatility does not contain any additional information about future realized volatility in levels when orthogonalized to all model based forecasts. There is however some incremental information regarding changes in future realized volatility.

"*If you have to forecast, forecast often*"
- Edgar R. Fiedler

Table of Contents

DEFINITIONS **6**

1 BACKGROUND **7**

1.1 Problem Specification 8
1.2 Target Audience 9
1.3 Outline and Delimitations 9

2 PREVIOUS RESEARCH **11**

3 THEORY **13**

3.1 Implied Volatility – an Introduction 13
3.1.1 A Note on Terminology: Volatility, Standard Deviation and Risk *13*
3.1.2 Black- Scholes Option Pricing Model *14*
3.1.4 The Greeks *15*
3.1.5 Implied Volatility and Its Potential Information Content *15*
3.1.6 Volatility Smiles, Skews and Surfaces *16*
3.2 Calculation of VSTOXX 18
3.3 Stylized Facts About Volatility 20
3.3.1 Volatility Clustering *20*
3.3.2 Asymmetric Reactions to Shocks *20*
3.3.3 Long Memory of Volatility *21*
3.3.4 The Structure of Volatility Over Time *21*

4 METHOD **23**

4.1 Forecasting with Parametric Models 23
4.1.1 General Forecasting Principle *23*
4.1.2 The Naïve Historical Model *24*
4.1.3 ARMA Model *25*
4.1.4 GARCH Forecasting Models *29*
4.1.5 Estimation of GARCH Using Maximum Likelihood *31*
4.1.6 Forecasting with GARCH *31*
4.2 Realized Volatility Calculations 32
4.3 Evaluation of Forecast 33
4.3.1 Forecast Performance Measures *33*
4.3.2 Predictive Power *35*
4.3.3 Information Content *36*
4.4 Test of Additional Information in Implied Volatility 36
4.4.1 Decomposition of Implied Volatility *37*
4.4.2 Construction of the Realized Volatility Vector *37*
4.4.3 Test of Additional Information in VSTOXX Using the F-test *38*
4.4.4 Test of Additional Information in VSTOXX Using GMM *39*

5 RESULT **43**

5.1 Data and Descriptive Statistics in the Estimation Period 43
5.2 Model Estimation 45
5.3 Data and Descriptive Statistic in the Forecast Period 49
5.4 Performance Measure Analysis 51
5.5 Predictive Power Analysis 52
5.6 Information Content Analysis 55
5.7 Analysis of Additional Information in VSTOXX 56
5.7.1 Test of Additional Information in VSTOXX Using the F-test *56*
5.7.1 Test of Additional Information in VSTOXX Using GMM-Estimation *59*

6 CONCLUSION **62**

6.1 PREDICTABILITY OF VOLATILITY 62
6.2 RELATIVE PERFORMANCE OF VOLATILITY FORECASTS 62
6.3 THE BIAS OF VOLATILITY FORECASTS 63
6.4 ADDITIONAL INFORMATION 63
6.5 CONCLUDING REMARKS 63

7 REFERENCES **65**

7.1 PUBLISHED REFERENCES 65
7.2 ELECTRONIC REFERENCES 67

8 APPENDIX **68**

A1 EVEIWS CODE FOR FORECASTING WITH EGARCH 68
A2 EVIEWS CODE FOR FORECASTING WITH ARIMA(1,1,1) 68
A3 EVIEWS CODE FOR FORECASTING WITH ARMA(2,1) 69
A4 EVIEWS CODE FOR CALCULATING THE ADJUSTED F-STATISTIC 69

Definitions

Volatility	Volatility can be defined as the amount of uncertainty about the size of changes in an asset's value. In this thesis we refer to volatility on a daily basis rather than annual.
True volatility	Given that the price of an asset follows some unknown process then the true volatility is the volatility of that price process. The true volatility is not observable and therefore has to be proxied.
Realized volatility	Realized volatility is the volatility actually observed in the market. The observation is just a proxy for the true underlying volatility. In this thesis we use a log range estimator to proxy true volatility.
Implied volatility	Volatility of an asset derived from the value of a derivative written on that asset. The value of the derivative implies a volatility of the underlying asset.
Volatility index	An index that is designed to capture the volatility of e.g. a stock index. The volatility index is derived from several traded index options written on the stock index.
Forecasted volatility	Forecasted volatility is the prediction of an unknown future volatility. It can be evaluated against realized volatility.

1 Background

In this section we intend to give a background to the topic of this thesis. We will also specify the purpose and the questions we intend to answer. In addition we will describe the outline and limitations of the thesis.

A derivative is a financial instrument that is linked to another (financial) asset. The value of the derivative is derived from the value of that other asset.[1] During the past 30 years the derivative market has exploded. Today there exist an almost infinite amount of derivatives covering a variety of underlying assets. Investors acting in the financial markets are provided with all thinkable and unthinkable ways to hedge against, or expose themselves to, different risks. The option is one of the most well known classes of derivatives. In short, an option offers the holder the right to sell, or buy, an underlying asset at a certain date in time at a pre-specified price.[2] Option pricing was revolutionized in the 1970s by an article published in the Journal of Political Economy written by Fischer Black and Myron Scholes. The well known Black-Scholes formula provided investors with a simple way to value European call options. All inputs needed to value option with the model are readily available in the market, all but one, the volatility of the underlying asset. Volatility is therefore a key element in the pricing of option contracts. The difficulty with volatility is that it is not directly observable and therefore has to be forecasted. The research on volatility has attracted a lot of attention within finance, both from academics and practitioners. The amount of literature on the topic is overwhelming to say the least, typing in volatility on Google Scholar yields almost a million hits. Volatility estimates are not only used as input in option pricing but is also an important factor in other financial applications such as Value-At-Risk and portfolio optimization.

The Black-Scholes formula provided investors a relationship between volatility and option value. Investors thereby could derive volatility implied from options traded in the market, this is what is known as implied volatility. Implied volatility is derived from contracts that are traded based on the market participants expectations about the future. It is therefore widely believed that implied volatility might hold some information about future realized volatility that cannot be captured by time series models fitted to historical data. For decades market

[1] Hull (2006) p.1
[2] Ibid p.6

participants and academic researchers have tried to build forecasts models in an attempt to gain more information about the future volatility. It is easy to understand the big interest among academics to find an accurate forecast of volatility to use as input in option pricing. The gain from a good model would not only be great in monetary terms but also in terms of knowledge about the option markets and its dynamics.

Taking a starting point in the above discussion we present the purpose of this thesis. The purpose of this thesis is to investigate the information content of model based volatility forecasts and the volatility index VSTOXX and their ability to predict the volatility of Dow Jones EURO STOXX 50.

1.1 Problem Specification

The thesis will focus on volatility and volatility forecasting. We start by examining how well volatility can be predicted. In order to determine the predictability of volatility we evaluate the accuracy of different forecast methods. We investigate which model that produces the most accurate forecast of future volatility. Finally we address the issue whether implied volatility contains any additional information beyond that of several model based forecasts.

Most academic research concludes that implied volatility is the best estimator of future volatility. However since the result is not entirely consistent it is important to initially investigate whether or not implied volatility actually outperforms all other models in our sample.[3] Most studies find that implied volatility dominates model based forecast. This result is usually derived by benchmarking each model based forecast against implied volatility. What is seldom done is to test the combined forecasting ability of model based forecasts. Becker et. al. addresses this issue and tests if implied volatility contains any additional information about future realized volatility beyond that forecasted by the models. Becker et. al. (2007) investigates the predictive power of the Standard & Poor 500 volatility index (VIX). In this thesis we intend to test the predictive power of the Dow Jones EURO STOXX 50 Volatility Index (VSTOXX) on the future realized volatility of Dow Jones EURO STOXX 50 (DJE50). More specifically we intend to test if there is any additional information in VSTOXX beyond that produced by a number of model based forecasts.

[3] see for example Granger and Poon (2003)

The general problem specification above can be summarized in four explicit questions:

i. *How well can volatility be predicted?*
ii. *How does model based forecast perform relative implied volatility when it comes to predicting future realized volatility?*
iii. *Is implied volatility an unbiased forecast of future realized volatility?*
iv. *Does implied volatility contain any incremental information about future realized volatility beyond that supplied by combined model based forecasts?*

1.2 Target Audience

The target audience of this thesis is people with an interest in volatility and volatility forecasting. To fully appreciate this thesis it is recommended that the reader has some basic knowledge about option pricing theory and econometrical methods. The authors of the thesis hope that academics and other people with interest in volatility forecasting will find the thesis interesting and meaningful. Furthermore we aim to give other finance students a better understanding of the concept of volatility.

1.3 Outline and Delimitations

Here we give a brief presentation of the outline of this thesis. In chapter 2 we start by presenting some of the existing research in the field of implied volatility and model based forecasts. In chapter 3 we continue with a more thorough presentation of relevant theories and empirical observations regarding volatility. We present the concept of implied volatility and problems associated with deriving it. For this purpose we describe the Black-Scholes model and its limitations when it comes to deriving implied volatility. We also outline the theoretical framework for deriving model free implied volatility. In chapter 4 we present the methodology used in this thesis. Here we present the time series models used for forecasting. We also describe how we will test the information content, predictive power and forecast errors of the different models. Furthermore we present how to test for any additional information in implied volatility beyond that supplied by the time series models. In chapter 5 we present the empirical results. Chapter 6 is devoted to conclusions and a discussion of the empirical results.

Working on this thesis we have been forced to make some delimitations. There are for example literally thousand of different time series models that could be tested in terms of predictive power. Most of these models are of course redundant and does not add much. We have tried to choose models that have been proven useful in the field of volatility forecasting and that capture as much of the empirical properties of volatility as possible.

We have also chosen to investigate only one stock index with its corresponding volatility index rather than several indices. While it could be argued that investigating more indices would add to the reliability of the results we believe it is of greater interest to pursue a more thorough investigation of one particular index. The investigated period is limited in duration due to the fact that VSTOXX has only existed since the start of 1999. The concerned reader should however not be alarmed since we have more than ten years of daily observations which is more than sufficient for the purpose at hand.

2 Previous Research

In this section we present previous research on the subject of volatility forecasting.

Researchers have struggled for over 20 years to determine which forecasting model that gives the most accurate prediction of future volatility. Granger and Poon (2003) examines over 90 studies on the subject of volatility forecasting. Based on their survey one can conclude that implied volatility in most cases outperforms other forecasting models. This result is not surprising since implied volatility potentially contains information about the market participants' expectations about the future that is not captured by models based on historical data. The implied volatility has been tested as predictor of future volatility for different asset classes. Lamoureux and Lastrapes (1993) find that implied volatility from stock options outperforms model based forecasts when forecasting the volatility of individual stocks. Blair et. al. (2001) reach the same conclusion when investigating the stock index S&P 100 with corresponding volatility index (old VIX[4]). In the foreign exchange market the implied volatility is also found to be the best predictor when investigated by Pong et. al. in 2004. Some of the earlier research reaches deviating conclusions about the predictive power of implied volatility. Canina and Figlewski (1993) conclude that implied volatility has little explanatory power of the realized volatility on S&P 100 index options. They find that implied volatility is even dominated by historical volatility. Jorion (1995) questions this result and argues that the conclusion that implied volatility is a poor forecast is driven by measuring errors rather than poor forecasting ability.

Even if most studies find that implied volatility provides the best forecast it has also often been found to be biased. Lamoureux and Lastrapes (1993) find implied volatility to be downward biased i.e. implied volatility consistently underestimates future realized volatility. The downward bias is confirmed by Blair et. al. (2001). Moreover Jorion (1995) finds implied volatility to be a biased forecast when investigating the foreign exchange market. Granger and Poon (2003) concluded in their survey that the overall result is that implied volatility is the best forecast vis-à-vis model based forecasts although a biased one.

[4] The methodology used to calculate VIX changed in 2003 after the article by Blair et. al. was published. Today VIX is calculated on S&P 500 using the same methodology used to calculate VSTOXX.

Many different models have been tested in an attempt to find a more accurate forecast method than implied volatility. The first class of models is the historical volatility models. This class of models include, but is not limited to: random walk models, historical averages, moving averages, autoregressive models and different exponential weighting schemes.[5] The Autoregressive Moving Average (ARMA) model is commonly used for forecasting. Theoretically the ARMA model should capture the persistent nature of volatility observed in the financial markets. Research conducted by Pong et. al. (2004) on the foreign exchange market show that the ARMA(2,1) model is a good forecast model for realized volatility on short time horizons.

The second class of models that is considered is the GARCH family models. GARCH models are based on conditional variance given some mean model rather than being fitted to historical realized volatility. GARCH models are motivated by their ability to capture some of the properties empirically observed in realized volatility. In the academic literature various types of GARCH models have been examined. The GJR or Threshold GARCH (TGARCH) developed by Glosten et. al. (1993) has been found to capture the asymmetric behaviour of volatility since this property is not captured in the ordinary GARCH model. The GJR model has been used by Taylor (2001) on different stock indices and has been found to outperform the ordinary GARCH(1,1) model. Another model to capture the asymmetry of volatility is the Exponential GARCH (EGARCH) model as suggested by Nelson (1991). Granger and Poon (2003) conclude that GARCH-type models dominate historical models in about half of the surveyed studies.[6]

The previous research on volatility forecasting suggests that volatility is fairly predictable and that the implied volatility provides the most accurate forecast. Furthermore implied volatility usually dominates model based forecasts in bilateral comparisons. The implied volatility has in most research papers been found to be biased.

[5] Granger and Poon (2003) p. 482*f*
[6] Ibid p. 506

3 Theory

In this section we intend to give a theoretical background of the concept of implied volatility. Furthermore we present some stylized facts about volatility. Finally we present some other theoretical frameworks underlying the rest of the thesis.

3.1 Implied Volatility – an Introduction

To understand the concept of implied volatility we take starting point in the Black-Scholes formula. We then present some of the issues associated with using the Black-Scholes option pricing formula for deriving implied volatility which motivates the methodology used to calculate VSTOXX.

3.1.1 A Note on Terminology: Volatility, Standard Deviation and Risk

There is some confusion as regarding the meaning of the word volatility, much depending on the lack of one clear definition of the word. In an option pricing context volatility is a measure of the uncertainty of the returns provided by the underlying asset. Within e.g. the Black-Scholes framework, volatility of an asset is defined as the standard deviation of continuously compounded asset returns.[7] In this thesis we always refer to volatility on a daily basis where volatility is assumed to increase with the square root of time. In other words if we were given the standard deviation on an annual basis then the volatility on a daily basis would be equal to:

$$\sigma_{daily} = \frac{Std.Dev.}{\sqrt{T}} \qquad (1)$$

Where T is the number of trading days in a year.

To further complicate the matter, the volatility of an asset is not observable and therefore has to be estimated somehow. In other words we not only have to produce a forecast of volatility but the actual (realized) volatility also has to be estimated. Below we return to the methodology used to estimate realized volatility in this thesis.

[7] Hull (2006) p. 286

Finally volatility is not risk per se, Granger and Poon (2003) argue that, unless attached to a distribution or a pricing dynamic, volatility is useless as a risk measure. Usually when standard deviation is thought of as risk this is done with the implicit or explicit assumption that returns are normally distributed.

3.1.2 Black- Scholes Option Pricing Model

The rapid expansion of the option market is mainly due to the discoveries made by Black, Scholes and Merton. Variations of the Black-Scholes option pricing model are used when traders price and hedge options on all sorts of underlying asset. There are some issues associated with using the implied volatility backed out from the Black-Scholes model. To understand these issues we give a brief background of the model.

The basic idea with the Black-Scholes model is that the value of the option can be replicated at all times using a bond and the underlying asset. Since the option can be replicated by the bond and the underlying asset it is possible to make a risk free arbitrage profit if any of the three assets is mispriced relative to the other two. If we can observe the price of the bond (the risk free interest rate) and the price of the underlying asset it is possible to determine the theoretical price of the option. Thus the Black-Scholes models is an arbitrage based model. In order to value the option an assumption has to be made about the distribution of prices for the underlying asset. In the Black-Scholes model it is assumed that asset prices are lognormally distributed. It then follows that continuously compounded returns are normally distributed.[8] If the normality assumption is violated the theoretical value of the Black-Scholes model will deviate from actual option prices. Implications of this will be discussed below.

There are five parameters needed to value an option with the Black-Scholes model. These are the strike price (K), the time to maturity (t), the price of the underlying asset today (S_0), the risk free interest rate (r) and finally the volatility (σ).[9] All these parameters are observable, all parameters except volatility. To correctly estimate volatility is therefore of great importance in option pricing.

[8] Figlewski (1997) p. 4
[9] Black and Scholes (1973) p. 640

3.1.4 The Greeks

In option pricing theory there are a number of different measures of the option values' sensitivity to the different variables affecting the value of the option. These sensitivities are known as the greeks[10] and are calculated as the partial derivative of the option value with respect to the different inputs. In order to make the following sections more understandable we define some of the most common greeks in table 1 below.[11]

Table 1 Definition of the greeks

Greek	Definition
delta	First partial derivative with respect to the price of the underlying asset
gamma	Second partial derivative with respect to the price of the underlying asset
theta	First partial derivative with respect to time to maturity
vega	First partial derivative with respect to volatility of the underlying asset
rho	First partial derivative with respect to the interest rate

3.1.5 Implied Volatility and Its Potential Information Content

In so far we have treated the value of the option as the unknown and listed the inputs needed to derive the value of the option. The Black-Scholes formula is a closed form solution into which we can simply plug in the inputs and derive a theoretical value of the option. Since options are traded in the market there are prices at which the option contracts are traded. It is then possible to invert the formula used to price options in order to derive the value of any of the inputs given that all other inputs are known. Above we concluded that all inputs except volatility are observable in the market, the inverted option pricing formula can therefore be used to derive volatility – this is what is known as implied volatility. If the Black-Scholes formula was inverted in order to derive the implied volatility there are however some implications, these implications are more thoroughly discussed below. The issues associated with using the inverted Black-Scholes model to derive implied volatility can be alleviated by using what is referred to as model free implied volatility. It is model free in the sense that it does not assume a specific model for how options are priced. We later return to the methodology used to calculate VSTOXX which is based on the concept of model free implied volatility.

[10] Vega is in fact not a letter in the Greek alphabet but is nevertheless referred to as a greek in options pricing theory since vega closely resembles the Latin letter V (for volatility)
[11] For a more thorough discussion on the greeks see for example Hull (2006) chapter 15

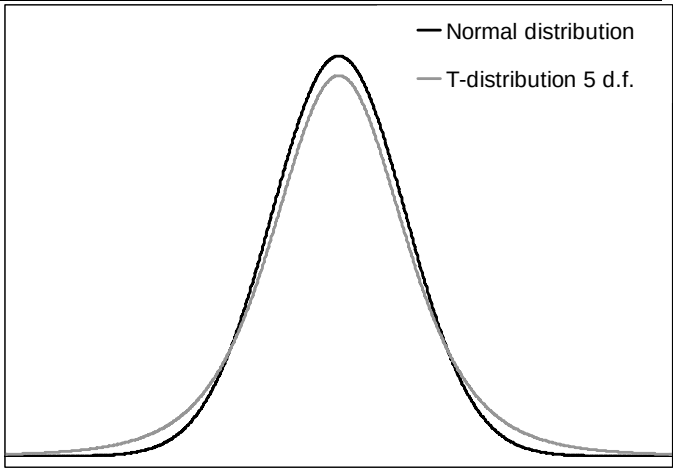
As we will see there are some problems associated with deriving implied volatility. The question therefore arises whether it is worth the effort. As we described in the previous research section above, several studies have shown that implied volatility outperforms time series models in forecasting future volatility. The argument in support of this empirical observation is that implied volatility not only incorporates information about past prices but also takes expectations about the future into account. More precisely the implied volatility of an option is usually interpreted as the markets expectation about the volatility during the life of the option.[12] The implied volatility derived from an index option written on e.g. S&P500 that expires in one month could be interpreted as the markets expectation about the average volatility of S&P500 during the coming month. Time series models only capture properties of past volatility. It therefore seems intuitive that implied volatility could hold additional information about the future.

3.1.6 Volatility Smiles, Skews and Surfaces

Recall the assumption of lognormal prices in the Black-Scholes model. Given that asset prices are lognormal, which implies that daily asset returns are normally distributed, the implied volatility derived from options should be a straight line over different strikes. In other words, if the assumptions of Black-Scholes are not violated and traders use Black-Scholes to price options, the volatility implied by options with different strike prices should be the same. However, if the distribution of asset returns has fatter tails than the assumed normal distribution then the implied volatility will be overestimated. [13]

Think of a call option that is deep out of the money, i.e. the current asset price is far below the strike price of the call. The value of the option comes only from the probability that the asset price will rise above the strike price before maturity. If the distribution of assets returns has fatter tails than the assumed normal distribution, the probability of extreme

Figure 1 Distributions



[12] Granger and Poon (2003) p. 486
[13] See for example Hull (2006) p. 377*f*

outcomes (high asset prices) is higher than in the theoretical framework of the Black-Scholes model. In figure 1 we illustrate the normal distribution and the fat-tailed t-distribution. Looking at the right tail we see that there is more probability (greater area) under the tail of the t-distribution than under the normal distribution. If the t-distribution is the true distribution of returns, the model (assuming a normal distribution) will underestimate the option price. If the Black-Scholes model then is inverted to derive implied volatility it becomes evident that the implied volatility will vary with the moneyness of the option. Extreme prices (and returns) will result in a relatively high implied volatility vis-à-vis less extreme outcomes resulting in an implied volatility that varies with strike price. The same argument could be made for deep out of the money puts which will be in the money only if the asset price falls below its strike price before maturity. [14]

Given that the empirical distribution of returns is not normal in the way described above, a plot of the implied volatilities of options (with different strikes) against their strike price would yield a convex graph. This convex pattern is popularly referred to as the volatility smile.[15] Evidence drawn from the currency market indicates that the distribution of exchange rates has fatter tails than the lognormal distribution, i.e. the probability of extreme highs and lows is higher in the empirical distribution as compared to the model.[16] In order for asset prices to have a lognormal distribution the volatility of the asset has to be constant. The price of the asset should also change smoothly without any jumps. The effect of jumps is largely dependent on the maturity of the option since jumps tend to be smoothed over time and thereby not affecting the distribution as much for options with long maturities.[17] Empirical evidence has shown that asset volatility is not constant over time. We develop this empirical observation further below.

The empirical evidence from the equity market differs somewhat from the findings regarding currencies. Equity options typically experience what is referred to as the volatility skew rather than the volatility smile. The volatility skew refers to the empirical observation that implied volatility decreases with increased strike price. The implied distribution has more probability in the left tail as compared to the lognormal distribution. There are some theories as for why the skew exists. One of these theories is related to the leverage of firms. If the stock of a firm

[14] See for example Hull (2006) p. 377*f*
[15] Ibid p.377*f*
[16] Hull (1998)
[17] Hull (2006) p. 379f

falls in value there is an increased probability of default which increases the risk of the stock. The implied volatility derived from a lower strike call option is thus greater than the implied volatility derived from a call option with the same time to maturity but a higher strike on the same stock. Interestingly the skew appeared first after the stock market crash of 1987 which suggests that the market participants after the crash started to incorporate the possibility of a future crashes when pricing stock options. The volatility skew is also present in options written on equity indices. One final empirical observation is that implied volatility varies not only with strike price but also with time to maturity, i.e. the term structure of implied volatility. When the moneyness (variation in strike price) of options is combined with the term structure we get what is referred to as the volatility surface.[18]

Since implied volatility derived from the Black-Scholes model varies with strike price, it is problematic to use the price of a single option to derive implied volatility. To alleviate this problem the concept of model free implied volatility can be utilized. This is the foundation for VSTOXX.

3.2 Calculation of VSTOXX

As we have discussed there are variations in the implied volatility depending on the moneyness of the option used to derive the implied volatility. One solution to this problem is to use some kind of weighting scheme to combine the implied volatility of a portfolio of many different options into one implied volatility measure.

We concluded above that implied volatility is positively related to the price of a call option. The intuition is that the asymmetry of the call option contract makes the contract more valuable when markets are more volatile. It then follows that volatility is positively related to the value of the option. Given this positive relationship between volatility and option value it seems reasonable to be able to track changes in volatility by tracking changes in option value.

The obvious problem when deriving the implied volatility from option values is that there are other sources than volatility affecting the value of the option. Besides volatility the main factor driving the value of the option is the value of the underlying asset. Now recall that the option's sensitivity to (small) price changes in the underlying asset is measured by the partial derivative of the option price with respect to the underlying asset's price (delta). If a portfolio

[18] Hull (2006) p. 381*f*

of options could be made delta neutral, i.e. the portfolio has a delta equal to zero, such a portfolio would not change in value even though the price of the underlying asset changed. Above we concluded that there are five factors driving option value. Since a delta neutral portfolio is insensitive to price changes of the underlying asset any change in value of the portfolio must originate from any of the other four factors. Now imagine if it was possible to make the portfolio insensitive to all but one of the other factors. Then it would be possible to indirectly track that factor by observing the value of the portfolio. For example if we intended to measure volatility of the underlying asset then it would be possible to do this by tracking changes in the value of the options in the portfolio. This is the basic idea behind the construction of volatility indices such as VIX and VSTOXX.

The construction of volatility indices such as VSTOXX is closely related to the methodology used when valuing so called volatility swaps.[19] We therefore refer to some of the volatility swap literature when discussing the theory behind VSTOXX below.

We have concluded that a delta neutral portfolio was needed in order to eliminate the changes in option value attributable to price changes of the underlying asset. When constructing VSTOXX, delta neutrality is achieved by adding futures[20] on the underlying asset to the portfolio.[21] Since implied volatility varies with the moneyness of the options, the portfolio is constructed to include different options of varying moneyness. The question is how to determine the portfolio weights for the different options. It turns out that it is possible to construct a portfolio with a constant vega (sensitivity to the volatility of the underlying) independent of the price of the underlying asset by weighting each option by the inverse of its squared strike price. This is valid as long as the price of the underlying is within the range of available strike prices and far away from the end of the range of available strikes.[22] If the portfolio has a constant vega across all included strikes, all price changes of the options have the same impact on implied volatility. It is then possible to track changes in volatility by observing several options. In order to control for the variation in implied volatility with term structure, VSTOXX has a constant time to maturity of a calendar month (approximately 21

[19] A volatility swap is an agreement between two market participants where the long (short) position of the contract makes a profit (loss) if volatility rises above a pre determined level in the future and vice versa
[20] A future is an agreement to buy/sell the underlying asset at a predetermined date at a predetermined price
[21] VSTOXX methodology http://www.stoxx.com/download/indices/rulebooks/djstoxx_indexguide.pdf p. 78
[22] Demeterfi et. al. (1999) p. 7

trading days).[23] Above we said that the implied volatility can be seen as the market's view of the average volatility of the underlying asset during the remaining life of the option. It then would follow that the observed value of VSTOXX at time t could be used as a forecast of the average volatility of the DJE50 during the subsequent 21 trading days. The methodology used to calculate VSTOXX alleviates some of the issues associated with using Black-Scholes to derive implied volatility from a single option. Since VSTOXX is a pure volatility measure it lends itself well to our purpose of volatility forecasting.

3.3 Stylized Facts About Volatility

The aim of this thesis is to model volatility and in order to choose the proper model we first have to identify some characteristics that are typical for volatility. In their survey of volatility models from 2003, Granger and Poon list a number of more or less documented properties of financial market volatility. These properties include, but are not limited to, fat tailed distributions of risky asset returns, volatility clustering and asymmetric reactions to shocks.

3.3.1 Volatility Clustering

Volatility clustering refers to the empirical observation that volatility in financial data appears to vary over time. The emerging pattern is that high (low) absolute returns are followed by more high (low) absolute returns. This pattern was first observed by Mandelbrot (1963) and has since been observed in almost all financial returns series. One possible explanation for this empirical observation is that information affecting returns does not come at evenly spaced intervals but rather comes in clusters.[24] If we for example look at stock returns it seems intuitive that stock returns are more volatile during reporting season when a lot of new information about the firm and its competitors is revealed. If return series experience volatility clustering it would be motivated to use a model of volatility that takes this feature into account. We will return to such models below.

3.3.2 Asymmetric Reactions to Shocks

Another empirical observation from financial return data is that the volatility of returns increases more following negative shocks than positive shocks of equal size. This asymmetry was pointed out by for example Black (1976, cited in Figlewski 1997). The asymmetric reaction to shocks of different sign is also called the leverage effect. The leverage effect refers

[23] VSTOXX methodology http://www.stoxx.com/download/indices/rulebooks/djstoxx_indexguide.pdf p. 75
[24] See for example Brooks (2008) p. 380

to one theory trying to explain the phenomena, the leverage hypothesis. According to the leverage hypothesis, negative shocks to returns increases the risk of the firm since the firm carries more debt relative to equity and thus makes it more risky when the stock price falls. Another theory for explaining the existence of asymmetric reactions to shocks is the volatility feedback effect. The volatility feedback effect is related to time varying risk premiums for bearing risk. If risk is priced, increased volatility would require asset prices to fall in equilibrium in order to compensate investors for the additional risk. The causality behind the feedback effect is the opposite from that of the leverage hypothesis.[25] Even though the cause of asymmetric reactions to shocks of different signs is still an issue open to debate, we can conclude that they do exist empirically and that we therefore have to consider this observed property in order to successfully model volatility.

3.3.3 Long Memory of Volatility

The persistent nature of volatility is another commonly observed property. Granger et. al. (2000) test the statistical properties of different asset classes including stock indices, interest rates and commodities. They conclude that while returns are at most linearly related at the first lag absolute returns have a very long memory. One way to test for the long memory of volatility is to fit a fractional integration model. Granger et. al. find evidence for fractional integration in the absolute return series. Furthermore they find the level of fractional integration to be time varying. Some studies even find that volatility is integrated of order one (unit root). The long memory of volatility is a reoccurring empirical observation in most studies and across asset classes. These empirical findings imply that a shock to volatility will decay very slowly.[26]

3.3.4 The Structure of Volatility Over Time

When forecasting volatility one implicitly assumes something about the underlying structure of volatility. For example the choice of a GARCH-family model is accompanied with the implicit assumption of volatility being deterministic. Others argue that volatility is best modelled as a stochastic process, see for example Hull and White (1987). The true volatility structure is of course not known and therefore we can only evaluate different models in terms of their ability to forecast out-of-sample volatility. Figlewski (1997) makes some comments on important aspects when forecasting volatility. The first issue concerns forecast horizon and

[25] Bekaert et. al. (2000) p. 1f
[26] Granger and Poon (2003) p. 482

the relation to data frequency. If the structure of volatility varies with forecasting horizon, it seems reasonable to choose a data frequency of the historical dataset that matches the structure that is going to be forecasted. When short forecast horizons are investigated daily or even intra-day data is used.[27] When the forecast horizon is increased so is the data sampling frequency in order to obtain the lowest forecast error.[28]

The second issue when forecasting volatility is related to the notion that the underlying volatility is not only unknown but that it also might be changing over time. Given that the structure of volatility is time varying it seems reasonable to recalibrate the model that is trying to capture it. If the underlying structure is changing but the parameters of e.g. a GARCH-model are not re-estimated, the GARCH-model could in fact be outperformed by less sophisticated models that do not take time varying volatility into account. So even though the GARCH-model was constructed to capture time varying volatility it could fail since it assumes a deterministic volatility structure.[29] Following this reasoning, frequent re-estimation of time series models is motivated.

[27] See for example Granger and Poon (2003) for a summary of such studies
[28] Figlewski (1997) p. 42
[29] Ibid p. 16

4 Method

In this section we describe the empirical method used in this thesis. We start by a describing the general forecasting principle we then continue by defining the models used for forecasting. Finally we describe the methodology used to evaluate the different forecasts.

4.1 Forecasting with Parametric Models

To analyze if future volatility can be predicted we estimate a number of different models to see which model that produces the best forecast of future volatility. We want to include models that capture the properties and behaviour of the volatility observed in the financial markets. As described in section 3.3, volatility has empirically been found to be time varying and subject to clustering. It is thus important that we choose models that capture these properties.

4.1.1 General Forecasting Principle

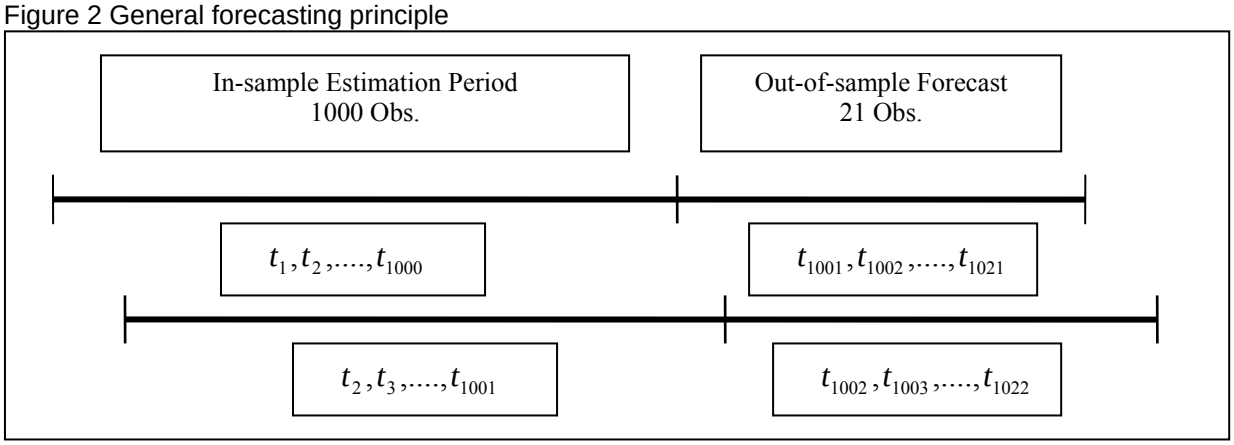
In this thesis we want to examine the ability of different models to predict the average volatility over the following 21 trading days. Different models use different approaches to facilitate this target. We use a basic principle for all forecast models where the parameters are estimated in the in-sample period. Parameters obtained from the in-sample period are used to forecast the out-of-sample period.[30] The forecasted volatility is benchmarked against the realized volatility. The ability to forecast realized volatility is then compared for the competing models.

Each model is used to forecast the daily volatility during the subsequent 21 trading days. The average of these forecasts is the forecast to be evaluated. The motivation for forecasting 21 days ahead is that VSTOXX is constructed to reflect a constant time to maturity of one calendar month which equals 21 trading days. Hence it follows that the forecast constructed is a multi-step-ahead forecast which goes $s = 1,2,...,21$ steps ahead for each of the out-of-sample volatility estimates. The total sample period follows rolling window estimation, meaning that the number of observations is constant for both the in-sample period and the out-of-sample

[30] Brooks (2008) p.245

period. Hence when the parameters in the in-sample period are re-estimated the oldest observation is dropped and replaced by a new additional observation.[31]

Figure 2 General forecasting principle



4.1.2 The Naïve Historical Model

We include a simple historical model based on historical realized volatility, i.e. a naïve forecast model that does not capture the time varying properties of volatility. We use the naïve historical model as a benchmark model since it gives a point of reference when evaluating the more sophisticated models.[32] The benchmark model will be used in a Theil-U test to evaluate the other more sophisticated models, see section 4.3 below for more details. From the historical daily high and low prices of the DJE50 index we compute daily volatility using the log range estimator. A more thorough motivation for the use of the log range estimator instead of daily squared returns is given below in section 4.2. Using the log range estimator, daily realized volatility is calculated with the following formula:

$$RV_t = \ln\left(\frac{high_t}{low_t}\right) \tag{2}$$

Where $high_t$ is the highest intraday price of the index and low_t is the lowest price level of the index during the day. For a given period T the average daily volatility can be calculated as:

$$\hat{\sigma}_T = \frac{\sum RV_t}{T} \tag{3}$$

[31] For EViews code used to calculate forecasts see appendix
[32] See for example Figlewski (1997) p. 16

The forecast $\hat{\sigma}_T$ from the naïve benchmark model is the average daily volatility from the preceding T days.

4.1.3 ARMA Model

Some academic literature suggests that an Autoregressive Moving Average (ARMA) model should be able to capture the persistent nature of volatility in the financial markets. The Moving Average process (MA) of the ARMA model is a linear combination of white noise disturbance terms[33]. In general a white noise process is defined as:

$$\begin{aligned} &E(u_t) = \mu \\ &\text{var}(u_t) = \sigma_u^2 \\ &u_{t-r} = \begin{cases} \sigma_u^2 \ \textit{if} \ t = r \\ 0 \ \textit{otherwise} \end{cases} \end{aligned} \tag{4}$$

The white noise process has constant mean μ, a constant variance σ_u^2 and zero autocovariance for all lags except lag zero. The MA(q) process for volatility would be defined as:

$$\sigma_t = \mu + \sum_{i=1}^{q} \theta_i u_{t-i} + u_t \tag{5}$$

The dependent variable σ_t depends on q white noise processes u_{t-i}. If we assume that the mean μ is equal to zero, the observed value of today is only dependent on previous error terms u_{t-i}.

If however σ_t follows an Autoregressive (AR) process then the current value depends only on previous values of the dependent variable, i.e. σ_{t-i}. The AR(p) process is defined as:

$$\sigma_t = \mu + \sum_{i=1}^{p} \phi_i \sigma_{t-i} + u_t \tag{6}$$

[33] Brooks (2008) p.211

It is important that the AR process is stationary. An AR-process is stationary if the roots of the characteristic equation lie within the unit circle, i.e. the process does not contain a unit root nor is it explosive.[34] For example an AR(2) is stationary if $|\phi_1| < 2$ and $|\phi_2| < 1$. For the AR(2) process, coefficients close to two and one implies that the modelled process decays very slowly. The coefficients can thus be seen as a measure of persistence of the modelled process, in this case volatility. If the process is non-stationary the impact of previous values never decays. A non-stationary process cannot be modelled within the ARMA framework, instead it has to be modelled with an Autoregressive Integrated Moving Average (ARIMA). It is thus important to test the stationarity condition in order to choose the proper model. In this thesis we chose to test the stationarity condition of the data with the augmented Dickey-Fuller unit root test.[35]

A combination of an AR and MA process gives the ARMA model. In the ARMA model today's observed value σ_t depends on both previous values σ_{t-i} and previous error terms u_{t-i}. The general ARMA(p,q) model is defined as:

$$\sigma_t = \mu + \sum_{i=1}^{p} \phi_i \sigma_{t-i} + \sum_{i=1}^{q} \theta_i u_{t-i} + u_t \qquad (7)$$

The first model we choose from the ARMA model family is the ARMA(2,1) model. The observed value today σ_t depends on the previous values σ_{t-1} and σ_{t-2} plus the previous error term at u_{t-1} and the mean. The use of the ARMA(2,1) model is motivated by the work made by Pong et. al. (2004). They show that the sum of two AR(1) processes can capture the persistent nature of volatility. The sum of two AR(1) processes is equivalent of an ARMA(2,1) process.[36] Other researchers favour a long memory Autoregressive Fractional Integrated Moving Average (ARFIMA) model for describing volatility[37]. The basic idea behind the ARFIMA model is that the persistence in volatility is more long lived than what an ARMA model can describe[38]. The ARFIMA model has been used by Li (2002) who concludes that the ARFIMA model gives a more accurate forecast of the volatility on longer

[34] See for example Brooks (2008) p. 217
[35] Ibid p.327*ff*
[36] See Granger and Newbold (1976)
[37] Martens and Zein (2004)
[38] Pong et. al. (2004) p.2542-2543

forecasting horizons than implied volatility. Since we only look at a fairly short forecast horizon of 21 trading days it could be argued that the long memory property of the ARFIMA model adds less to the forecasting ability. For example Pong et. al. (2004) find that the ARMA(2,1) and ARFIMA models perform equally well when the realized volatility is estimated using high frequency data. Since we use the log range estimator to capture some of the intraday properties of realized volatility we find it motivated to follow Pong et. al. and choose an ARMA(2,1) model.

A first look at the statistical properties of the realized volatility series in section 5.1 below suggest that the realized volatility suffers from unit root, i.e. is non-stationary, in several in-sample-periods. From this empirical observation we draw the conclusion that it would be proper to use an ARIMA model in order to capture the non-stationary property of our time-series. As a complement to the ARMA(2,1) model an ARIMA(1,1,1) model is estimated. The general ARIMA(p,d,q) is denoted:

$$\Delta^d \sigma_t = \mu + \sum_{i=1}^{p} \phi_i \Delta^d \sigma_{t-i} + \sum_{i=1}^{q} \theta_i u_{t-i} + u_t \qquad (8)$$

Where Δ^d denotes the number of times the series is differentiated. Since we differentiate our series once the ARIMA(p,1,q) model is defined as:

$$\Delta \sigma_t = \mu + \sum_{i=1}^{p} \phi_i \Delta \sigma_{t-i} + \sum_{i=1}^{q} \theta_i u_{t-i} + u_t \qquad (9)$$

With the intention of deciding the predictive ability of the ARMA(2,1) model and the ARIMA(1,1,1) model we need to forecast the future volatility and compare the result with the realized volatility. We let $\hat{\sigma}_{t,s}$ denote the forecast made by an ARMA(p,q) at time t for s days into the future.[39]

[39] Brooks (2008) p. 248

$$\hat{\sigma}_{t,s} = \sum_{i=1}^{p} \alpha_i \hat{\sigma}_{t,s-i} + \sum_{j=1}^{q} \beta_j u_{t+s-j}$$
$$\hat{\sigma}_{t,s} = \sigma_{t+s} \; if \; s \leq 0 \qquad (10)$$
$$u_{t+s} = \begin{cases} 0 \; if \;\; s \geq 0 \\ u_{t+s} \;\; if \;\; s < 0 \end{cases}$$

The α_i coefficient captures the autoregressive part and β_j captures the moving average part of the ARMA(p,q) model. The MA part of the process dies after lag q. Since the error term u_t in the forecast period is equal to zero, the MA part of the forecasted ARMA(2,1) model will die out two steps into the future. The AR part of the process will not die out in the forecast period, the forecasted value $\hat{\sigma}_{t,3}$ will be based on the forecasted values $\hat{\sigma}_{t,1}$ and $\hat{\sigma}_{t,2}$. Since the error term u_t is equal to zero in the forecast period the ARMA(2,1)-model forecasts volatility as an AR(2) process two steps ahead into the future.

The forecast with the ARIMA(1,1,1) model is not as straightforward as the ARMA(2,1) model. The following methodology is used to forecast with the ARIMA(1,1,1) model:

$$\hat{\sigma}_{t,1} = \sigma_t + \alpha_1(\sigma_t - \sigma_{t-1}) + \beta_1(u_t - u_{t-1})$$
$$\hat{\sigma}_{t,2} = \hat{\sigma}_{t,1} + \alpha_1(\hat{\sigma}_{t,1} - \sigma_t) \qquad (11)$$
$$\hat{\sigma}_{t,s} = \hat{\sigma}_{t,s-1} + \alpha_1(\hat{\sigma}_{t,s-1} - \hat{\sigma}_{t,s-2})$$

Where $\hat{\sigma}_{t,s}$ denotes the forecast made by the ARIMA(1,1,1) at time t for s days into the future. The coefficient α_1 denotes the autoregressive part, β_1 denotes the moving average part of the ARIMA(1,1,1) model. It is important to note that we exclude the intercept μ in the forecast even though it is estimated. The reason for this is that it would be counterintuitive to have a deterministic component of changes in volatility. As expected estimation results, presented in section 5.1, indicate that the intercept is zero. However we choose to include the intercept when estimating to assure unbiased slope coefficients.[40] The MA part of the process dies after lag 1 since the error term becomes zero in the forecast period. The ARIMA(1,1,1) is estimated to model changes in volatility. Since we are interested in forecasting levels rather

[40] The OLS estimator is guaranteed to produce unbiased parameter estimates if an intercept is included in the regression equation, see for example Brooks (2008) p. 131*f*

than differences we add the current level of volatility in t to the forecast in $t+1$. In the general case we use the forecasted value in $t+s$ as an input to the forecasted value in $t+s+1$.

4.1.4 GARCH Forecasting Models

In section 3.3 we summarized a number of stylized facts about volatility. One of these was the existence of volatility clustering, i.e. large (small) absolute returns are followed by more large (small) absolute returns. Engle (1982) suggests that this property should be modelled with an Autoregressive Conditional Heteroskedasticity (ARCH) model. The ARCH family of models consists of at least two equations. The first equation, referred to as the mean equation, is trying to model the first moment of returns. The second equation, referred to as the variance equation, is intended to capture the second moment of returns.[41] In the original ARCH(q) model the variance is conditional upon q past shocks to the mean equation.

$$\begin{aligned} R_t &= \mu + \varepsilon_t \\ \sigma_t^2 &= \omega + \sum_{i=1}^{q} \alpha_{t-i} \varepsilon_{t-i}^2 + \eta_t \\ \varepsilon_t &\sim N(0, \sigma_t^2) \end{aligned} \qquad (12)$$

The ARCH-model was generalized by Bollerslev (1986) to capture the persistent nature of variance. In order for the above mentioned ARCH-model to capture the effect of shocks far back in time, a lot of regression coefficients have to be estimated. To reduce the number of estimated parameters Bollerslev introduced the Generalized Autoregressive Conditional Heteroskedasticity (GARCH)-term into the variance equation. The GARCH-term is the conditional variance in the previous period. A GARCH(p,q) has the following general specification.

$$\begin{aligned} R_t &= \mu + \varepsilon_t \\ \sigma_t^2 &= \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 + \eta_t \\ \varepsilon_t &\sim N(0, \sigma_t^2) \end{aligned} \qquad (13)$$

[41] See for example Champbell, Lo, Mackinley p. 483

The most commonly used specification of the GARCH model is the GARCH(1,1).[42] Many existing papers utilize a variation of the standard GARCH(1,1) proposed by Glosten et. al. (1993) known as GJR or Threshold GARCH (TGARCH).[43] The GJR model is constructed to capture the asymmetric reaction to shocks described in section 3.3.2. In the model, a dummy variable is introduced that takes the value of one when the shock in the previous period is negative and the value of zero when the shock is positive. The coefficient associated with the dummy variable would then capture the additional contribution to variance from negative shocks. When it comes to parameter estimation the GJR-model shares some issues with the standard GARCH(1,1)-model. If no restrictions are imposed when estimating the parameters, estimates could be negative which implies non-stationarity of volatility. Negative parameters could also result in negative variance forecasts which would be counterintuitive.[44] In order to avoid negative variance but still capture the asymmetric behaviour of the variance we will use the Exponential GARCH model (EGARCH) proposed by Nelson (1991). There are different ways to specify the EGARCH model but in this thesis we use the following specification:[45]

$$\varepsilon_t = R_t - \mu$$
$$\ln(\sigma_t^2) = \omega + \beta \ln(\sigma_{t-1}^2) + \gamma \frac{\varepsilon_{t-1}}{\sqrt{\sigma_{t-1}^2}} + \alpha \frac{|\varepsilon_{t-1}|}{\sqrt{\sigma_{t-1}^2}} \qquad (14)$$

The left-hand side of the equation $\ln(\sigma_t^2)$ is the log of the conditional variance. The EGARCH model has two important and sought after properties. First of all the variance will always be positive even if the coefficients are negative. Secondly the EGARCH model is able to capture asymmetric shocks. The asymmetry is captured by the γ coefficient. A negative γ coefficient implies that there is a negative relationship between return and volatility. A negative γ indicates that positive shocks lead to lower volatility than negative shocks of the same magnitude which gives empirical support to the existence of a leverage effect (and volatility feedback).[46]

[42] See for example Champbell, Lo, Mackinley p. 483
[43] See for example Becker et. al (2007),
[44] Estimation of the GJR model resulted in negative coefficients. We therefore replaced GJR with EGARCH
[45] See for example Brooks (2008) p.406
[46] Brooks (2008) p.406 f

4.1.5 Estimation of GARCH Using Maximum Likelihood

The parameters of the GARCH model are estimated using Maximum Likelihood (ML) which is a suitable estimation technique for non-linear models such as GARCH. When estimating parameters using ML we have to assume a distribution for the errors and a model specification. Given that the assumed distribution is correct and that the model is correctly specified, the model estimate is found by altering the parameters so that the particular dataset best fits the assumed model specification. We define a vector of parameters (θ) whose elements are altered. When the vector contains the true parameters the errors are *IID* and follow the assumed distribution. Given that we assume that the errors follow a normal distribution, the true parameters are obtained by maximizing the following function given that we estimate the EGARCH model above.[47]

$$\ln L(\theta) = -\frac{T}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\ln(\sigma_t^2) - \frac{1}{2}\sum_{t=1}^{T}\frac{(R_t-\mu)^2}{\sigma_t^2} \quad (15)$$

The maximization is done numerically using the Marquardt algorithm. One problem when estimating GARCH-models with ML is to find the proper starting values for the optimization. The log-likelihood function has several local optima and the solution can therefore depend on the starting values. We use backcasting to derive the presample variance, i.e. the initial conditional variance. If the backcasting parameter is set to one the initial conditional variance will be equal to the unconditional variance. In this thesis we will smooth the presample variance since this usually results in better estimates.[48]

4.1.6 Forecasting with GARCH

Using the EGARCH model we get the following forecast of the conditional variance:

$$\begin{aligned}
\hat{\sigma}_{t,1}^2 &= \exp\left[\omega + \beta\ln(\sigma_t^2) + \gamma\frac{\varepsilon_t}{\sqrt{\sigma_t^2}} + \alpha\frac{|\varepsilon_t|}{\sqrt{\sigma_t^2}}\right] \\
\hat{\sigma}_{t,2}^2 &= \exp\left[\omega + \beta\ln(\sigma_{t,1}^2)\right] \\
\hat{\sigma}_{t,s}^2 &= \exp\left[\omega + \beta\ln(\sigma_{t,s-1}^2)\right]
\end{aligned} \quad (16)$$

[47] See for example Campbell et. al. (1997) p. 487*f*
[48] Eviews 6 User Guide II p. 193

The error term is zero in the forecasting period. Forecasts made more than one-step-ahead thus only takes the previous conditional variance as input.

4.2 Realized Volatility Calculations

To proper evaluate the accuracy of the forecasted volatility it is important to have a correct measure of the realized volatility. Since the realized volatility is a latent variable, not directly observable, it has to be estimated.[49] We use a logarithmic range method to calculate a proxy for the realized volatility. This method is based on the work of Alizadeh et. al. (2002). The range based estimator for calculating the realized daily volatility is computed by taking the difference between the natural logarithm of the daily high price and the natural logarithm of the daily low price.

$$RV_t = \ln(high_t) - \ln(low_t) = \ln\left(\frac{high_t}{low_t}\right) \qquad (17)$$

Where RV_t is a proxy of the realized daily volatility for day t. We then construct $\overline{RV}_{t+21}$ as the average of the 21 daily realized volatilities between t and $t+21$.

Another proxy for realized volatility is to use the square root of squared daily returns. Academic studies have criticized this approach since much information is lost about the true volatility when only closing prices are considered.[50] This is due to that squared returns only incorporates one price observation per day while the intraday high and low prices contain additional information. According to Alizadeh et. al. (2002) the range based estimator produces more efficient estimates of true volatility than squared returns. For the above reasons the authors favour the range based estimator over squared returns.

Many existing studies favour the use of higher frequency data when calculating the realized daily volatility based on work by Andersen and Bollerslev (1998) and Andersen et. al. (2003). The most common approach of calculating realized volatility with high frequency data is to collect observations of the underlying asset with an interval of 5 minutes during the day. The data points are then squared and summed for the day. The basic idea with short time intervals

[49] Granger and Poon 2003 p. 492
[50] Granger and Poon 2003 p. 492 and Alizadeh et al (2002)

is that tick data incorporates more information about the true volatility. The use of high frequency data has been found to improve the efficiency of the realized volatility estimates as compared to squared daily returns.[51] The use of high frequency data is not unproblematic, market microstructure can have a significant impact on the estimation of realized volatility. One effect of the market's microstructure is that prices bounce between bid and ask.[52] This leads to an upwardly bias of the volatility estimate when using high frequency data to estimate volatility. The range based estimator is also biased due to the bid-ask bounce. The problem is however less significant since we only have two observations per day, the high and low.[53] Less problems with market microstructure and lack of tick data observations motivates the use of high/low data rather than high frequency data.

4.3 Evaluation of Forecast

A model is not much worth if it does not work empirically thus testing the models is an essential part of this thesis.

4.3.1 Forecast Performance Measures

In the econometric literature a number of different performance measures are suggested and they all have their individual merits and drawbacks. Performance measures evaluate the accuracy of a forecast by comparing the out-of-sample forecast with the actual value in each time step and aggregate over the out-of-sample period. We choose to use three performance measures to rank and evaluate the performance of the forecast models. Since the loss functions have little to say about the performance of a forecast when the statistics are examined individually we compare the result of the different loss functions for competing forecast models.

One of the most common performance measures is the mean squared error (MSE). The MSE loss function is defined as[54]:

$$MSE = \frac{1}{n}\sum_{t=t+1}^{t+n}(\overline{RV}_{t+21} - \hat{\sigma}_t)^2 \qquad (18)$$

[51] See for example Blair et. al. (2001)
[52] For more about the bid ask-spread see Campbell et. al. (1997) Chapter 3
[53] Alizadeh et al (2002)
[54] Dunis et al (2001)

Where n is the number of forecasts to be evaluated, $\hat{\sigma}_t$ is the forecasted volatility at time t and $\overline{RV}_{t+21}$ is the realized volatility. MSE is a quadratic loss function and is best suited in situations where large forecast errors are more serious then smaller errors.[55]

The second performance measure we use is the mean absolute error MAE. This measure is more suitable than MSE when outliers are present[56]. The MAE loss function is defined as:

$$MAE = \frac{1}{n}\sum_{t=t+1_1}^{t+n} | \overline{RV}_{t+21} - \hat{\sigma}_t | \qquad (19)$$

Where n is the number of forecasts to be evaluated, $\hat{\sigma}_t$ is the forecasted volatility at time t and $\overline{RV}_{t+21}$ is the realized volatility.

Another common performance measure is the Theil-U statistic which is defined as[57]:

$$Theil - U = \frac{\sqrt{\sum_{t=1}^{n}\left(\frac{\overline{RV}_{t+21} - \hat{\sigma}_{ft}}{\overline{RV}_{t+21}}\right)^2}}{\sqrt{\sum_{t=1}^{n}\left(\frac{\overline{RV}_{t+21} - \hat{\sigma}_{bt}}{\overline{RV}_{t+21}}\right)^2}} \qquad (20)$$

Where n is the number of forecasts to be evaluated, $\hat{\sigma}_{ft}$ is the forecasted volatility at time t, $\overline{RV}_{t+21}$ is the realized volatility and $\hat{\sigma}_{bt}$ is the forecast obtained from some naïve benchmark model, in this case historical volatility. The idea behind the Theil-U statistic is to use a naïve benchmark model and see if a more complex forecasting model yields a different result. If the Theil-U statistic equals one the forecasting model and the naïve model perform equally bad. If the Theil-U statistic is less than one the more complex forecasting model is superior to the benchmark model.[58]

[55] Brooks (2008) p.251-253
[56] Dunis et al (2001)
[57] Brooks (2008) p.254
[58] See Brooks (2008) for further discussion

4.3.2 Predictive Power

With the intention of determining the predictive power of the forecasts we perform a regression analysis for each forecasting method. The basic idea behind the regression analysis is to determine the explanatory power of the forecasted volatility. We estimate the following equation:

$$\overline{RV}_{t+21} = \alpha + \beta\hat{\sigma}_t + \varepsilon_t \qquad (22)$$

Where $\hat{\sigma}_t$ is the forecast at time t, $\overline{RV}_{t+21}$ is the realized volatility and ε_t is the error term. We estimate the regression using ordinary least squares (OLS). The predictive power of the forecasted volatility is expressed by the coefficient of determination R^2. We test all our different models and rank them on their predictive power expressed by the R^2 coefficient.

To evaluate the bias of the forecasts we test some hypotheses on the OLS estimates of α and β. In order to reliably infer around the estimated parameters, the OLS estimator has to be efficient. OLS is an efficient estimator given uncorrelated residuals and homoskedasticity. The realized volatility could be heavily autocorrelated and heteroskedastic as suggested by Jorion (1995). In the presence of autocorrelation and heteroskedasticity the properties of the error term ε_t does not satisfy the OLS-assumptions. In order to correct the t-statistics we use the correction of the standard errors proposed by Newey-West (1987) and used by Charoenwong et. al. (2008) in their evaluation of volatility. Poon and Granger (2003) argue that it is very important to examine if the forecasted volatility is biased in some way. The forecast is unbiased when α=0 and β=1 and downwardly biased if $\alpha > 0$ and $\beta = 1$ or $\alpha = 0$ and $\beta > 1$.[59] A downward bias means that the volatility is underestimated by the model. Downward bias can be corrected and taken into account, and is thus not a major problem. The most common scenario is that $\alpha > 0$ and $\beta < 1$, which is a more serious problem than a downward bias. If the forecasting model suffers from this type of behaviour the model will under-forecast low volatility and over-forecast high volatility. This is problematic since it is impossible to determine if the subsequent period is a low or a high volatility period.

[59] Granger and Poon (2003) p.503

There is an important difference between bias and predictive power. A biased forecast could yield a good prediction of the future volatility if the bias can be corrected and taken into account. The uncertainty of the predicting power of an unbiased model with big forecast errors could on the other hand be completely useless in practice.[60]

4.3.3 Information Content

Instead of evaluating the forecasting accuracy and the predictive power of the forecasting models, Jorion (1995) suggests that it would be useful to look at the information content of the forecasting models. The basic idea is to test the information content of the volatility forecast with respect to the realized volatility one day ahead.

$$RV_t = \alpha + \beta\hat{\sigma}_t + \varepsilon_t \qquad (23)$$

Where RV_t denotes the realized volatility one day ahead, $\hat{\sigma}_t$ denotes the forecasted volatility at time t. The information content test is not designed to test if volatility forecast models give the best prediction about the future volatility for the entire forecast horizon. Instead the information content test is designed to test if the volatility forecast can say anything at all about the volatility one day ahead. Hence we test if the volatility forecast contains any information about the future and not only information about the historical volatility. The test is conducted by estimating equation 23 with OLS. As above we correct the standard errors with the Newey-West estimator to account for the assumed autocorrelation and heteroskedasticity. To determine the information content we check if the slope coefficient $\beta > 0$. The information content test is performed for all forecasting models and we rank the models based on their performance.

4.4 Test of Additional Information in Implied Volatility

We follow the methodology of Becker et. al. (2007) when testing for additional information in implied volatility beyond that contained in the model based forecasts (MBFs). In order to investigate whether or not there is any additional explanatory power in VSTOXX we first have to decompose VSTOXX into two components. The first component is constructed to contain the same information as the combination of the MBFs, we call this $\text{VSTOXX}^{\text{MBF}}$.

[60] Granger and Poon (2003) p.491

The second component contains any additional information that may or may not be useful for forecasting. The second component is called $VSTOXX^*$.

4.4.1 Decomposition of Implied Volatility

We start by calculating the model based forecast (MBF) for each time series model. In order to match the forecast horizon with the horizon of implied volatility the average volatility forecast during the following calendar month is calculated. All MBFs at time t are stored in a vector ω_t. The objective is to decompose VSTOXX into the components:

$$VSTOXX_t = VSTOXX_t^{MBF} + VSTOXX_t^* \tag{24}$$

Since we have the implied volatility and the MBFs we can write the above relation as

$$VSTOXX_t = \gamma_0 + \gamma_1 \omega_t + \varepsilon_t \tag{25}$$

Any additional information beyond that captured by the MBFs is now in the residual of the above regression model. What we have done is to construct a new time series $VSTOXX^*$ which is made orthogonal to the vector of MBFs hence:

$$VSTOXX_t^* = \hat{\varepsilon}_t \tag{26}$$

4.4.2 Construction of the Realized Volatility Vector

Since the intention is to test if there is any additional explanatory power of realized volatility in $VSTOXX_t^*$ we, following Becker, construct a vector of realized volatility. Since VSTOXX is defined to capture the volatility of the underlying during the following calendar month we therefore construct the realized volatility vector to include the average realized volatility during the 21 trading days following the date of the forecast. The realized volatility is calculated as:

$$\overline{RV}_{t+T1} = \sqrt{\frac{1}{T_1}\sum_{i=1}^{T_1} \ln\left(high/low\right)^2} \tag{27}$$

Where T_1 is the number of days over which the average realized volatility is calculated. In order to test if there is any additional information in implied volatility for shorter time horizons we include the following forecast horizons in the realized volatility vector **RV**.

$$\mathbf{RV} = \left\{ \overline{RV}_{t+1,}\ \overline{RV}_{t+5,}\ \overline{RV}_{t+10,}\ \overline{RV}_{t+15,}\ \overline{RV}_{t+21} \right\} \qquad (28)$$

To test if there is any information regarding the changes in realized volatility incorporated in the implied volatility we create the following vector.

$$\mathbf{dRV} = \left\{ d\overline{RV}_{t+1,}\ d\overline{RV}_{t+5,}\ d\overline{RV}_{t+10,}\ d\overline{RV}_{t+15,}\ d\overline{RV}_{t+21} \right\} \qquad (29)$$

Where $d\overline{RV}_{t+j} = \overline{RV}_{t+j} - RV_t$

4.4.3 Test of Additional Information in VSTOXX Using the F-test

When it comes to testing the potential additional information content of the implied volatility Becker et. al. (2007) propose two different testing strategies. An intuitive test of additional information in VSTOXX would be to run the following linear regression

$$VSTOXX_t^* = \beta \mathbf{RV} + u_t \qquad (30)$$

Where β is the vector of slope coefficients from the linear regression. If all slopes of β are equal to zero this would mean that there is no additional information regarding future realized volatility contained in $VSTOXX_t^*$. The reader should be aware that the formulation of the test in equation 30 does not imply that realized volatility in any way causes $VSTOXX_t^*$ it is just a way to investigate the existence of a linear relationship between $VSTOXX_t^*$ and future realized volatility. There are however some issues when it comes to estimating the above regression. One of the underlying assumptions of Ordinary Least Squares (OLS) is that the error term u_t is not serially correlated. If this assumption is violated we can not safely use the standard errors to perform inference around the parameters.[61] Since the forecast is made multiple-steps-ahead, the forecasted values are overlapping and therefore most likely

[61] See for example Brooks (2008) p. 149

correlated. The joint null hypothesis of zero slope coefficients is tested by the F-test statistic. To account for the possible autocorrelation in the residuals, Becker et. al. (2007) make use of the adjusted test F-statistic suggested by Harvey and Newbold (2000). The adjusted F-statistic was developed to evaluate encompassing forecasts and constructed to account for overlapping data. In this thesis we face exactly the problem of overlapping data and we therefore find it motivated to implement this non standard test procedure even though it is not thoroughly tested. In practice Harvey and Newbold make adjustments to the variance covariance matrix to allow for serial correlation.[62] The adjusted F-statistic is calculated as

$$F = \frac{1}{K-1} n\hat{\beta}' \hat{D}^{-1} \hat{\beta} \tag{31}$$

Where K is the number of estimated parameters, n is the number of observations, $\hat{\beta}$ is the vector of estimated parameters and $\hat{D} = \hat{M}^{-1}\hat{Q}\hat{M}^{-1}$ where $\hat{M} = n^{-1}X'X$. X is a $n \times K$ matrix of $K-1$ independent variables plus a column of ones for the intercept and $\hat{Q}$ is a $K \times K$ matrix in which each element can be calculated as:

$$\hat{q}_{ij} = \frac{1}{n}\left[\sum_{t=1}^{n} x_{it}x_{jt}u_t^2 + \sum_{m=1}^{h-1}\sum_{t=m+1}^{n} x_{it}x_{j,t-m}u_t u_{t-m} + \sum_{m=1}^{h-1}\sum_{t=m+1}^{n} x_{i,t-m}x_j u_t u_{t-m}\right] \tag{32}$$

Where u_t is the residual at time t from equation 30 and h is the number of steps-ahead of the forecast.[63] In this thesis we intend to forecast up to 21 days ahead and h is therefore set to 21. As usual the F-statistic is used to test the joint null hypothesis that all slopes in the vector β of equation 30 are equal to zero.

4.4.4 Test of Additional Information in VSTOXX Using GMM

Becker et. al. (2007) point out that the robustness of the test statistic proposed by Harvey and Newbolt is only tested under short forecast horizons and that the results therefore should be interpreted with caution, especially if the forecasting horizon is increased.[64] In order to increase the reliability of the results we follow Becker et. al. and use Generalized Method of

[62] Harvey and Newbold (2000) p. 473*f*
[63] See appendix for the Eviews code used to calculate the F-statistic
[64] Becker et. al. (2007) p. 2542

Moments (GMM) to investigate the existence of a linear relationship between realized volatility and *VSTOXX**.

We start by giving a brief overview of the GMM-estimation technique. The basic idea is to estimate a linear or non linear equation by fulfilling a number of moment conditions by varying the parameters to be estimated. To define the moment conditions we choose a set of instrument variables. The instrument variables could be identical to the independent variables of the regression to be estimated. Imagine if there is one independent variable (regressor) and we want to estimate a linear regression with an intercept. If we define the moment conditions that the regressor should be orthogonal to the residual and that the sum of squared residuals should be zero, we would have a system of 1+1=2 equations. Since there are two parameters to be estimated and two moment conditions, the system of equations is just identified. For example if we want to estimate the equation $y_t = \alpha + \beta_1 x_t + \varepsilon_t$ we get the following set of moment conditions.

$$\begin{aligned} \sum (y_t - \alpha - \beta_1 x_t)^2 &= 0 \\ \sum (y_t - \alpha - \beta_1 x_t) x_t &= 0 \end{aligned} \tag{33}$$

Since the system of equations is just identified there exists an analytical solution to the above optimization problem. In fact this is exactly the same thing as estimating the equation using OLS. Recall that the OLS-estimator minimizes a function of squared residuals under the assumption that the independent variables are uncorrelated with the residual. If more instruments are added to the above system of equations then the system is overidentified which would make it impossible to find an exact analytical solution. The estimator will now try to minimize the objective function through a numerical procedure by varying the parameters in order to satisfy the conditions as closely as possible. Imagine if we expanded the list of instruments to include not only x_t but also the new variable z_t. The restriction would be the same as before, z_t should be uncorrelated with the residual. This additional restriction would be referred to as an overidentifying restriction since there are now three equations but only two unknowns. We would then re-estimate the model. If z_t was entirely uncorrelated with the residual then the GMM estimator would find the exact same solution as before. Furthermore we can test the validity of the additional restriction using the so called J-

statistic which is the value of the objective function times the number of observation. The null hypothesis to be tested is that the overidentifying restrictions are valid. In the above example we only have one overidentifying restriction. In other words if the J-statistic is small enough we cannot reject the null that z_t is orthogonal to the residual. To express this differently, the smaller the value of the objective function to minimize (which leads to a smaller J-statistic), the more likely it is that the assumed moment condition actually fits the dataset at hand. GMM has the advantage that estimation does not require any assumption about distribution of the residuals. Furthermore GMM also can handle heteroskedasticity of unknown form.[65]

Recall equation 25 above where VSTOXX is regressed on the model based forecasts in order to construct $VSTOXX^*$. $VSTOXX^*$ is linearly independent or orthogonal to the MBFs and simply defined as the residual of equation 25. Above we described how to use OLS to estimate equation 25. We could however have used another estimator such as GMM since the OLS-estimator can be seen as a special case of GMM. If we were to estimate equation 25 using GMM we would have the same model specification as before.

$$VSTOXX_t = \gamma_0 + \gamma_1 \omega_t + \varepsilon_t \qquad (34)$$

Recall that the elements of ω_t are the model based forecasts. Estimating equation 34 using GMM also requires us to supply a set of instrument variables. As mentioned above the way GMM works is that it minimizes a function by varying the parameters to be estimated in order to satisfy a pre-defined set of conditions. More formally if we were to estimate the parameters $\gamma(\gamma_0, \gamma_1')$ of equation 34 we would minimize the function $V = \mathbf{M}'\mathbf{H}\mathbf{M}$ where $\mathbf{M} = T^{-1}(\varepsilon_t(\gamma)'\mathbf{Z}_t)$ is a $K \times 1$ vector of moment conditions, $\mathbf{H}$ is a $K \times K$ weighting matrix and $\mathbf{Z}_t$ is a vector of instrument variables.

If we estimate the above regression using GMM with the moment conditions that the residuals should be orthogonal to the elements of ω_t we obtain the same parameter estimates as we would if we used OLS. Becker et. al. suggest that the vector of instruments should be expanded to include, not only the regressors of equation 34, but also the vector of realized volatility $\mathbf{RV}$. If GMM is able to find parameters so that the element of ω_t and $\mathbf{RV}$ are

[65] For more information on GMM estimation see for example Verbeek (2004) p. 159*ff*

uncorrelated with the residuals of equation 34, this would imply that (i) $VSTOXX^*$ is orthogonal to all MBFs and (ii) $VSTOXX^*$ is orthogonal to **RV**, which means that VSTOXX contains no additional information beyond that contained in the model based forecasts.

We test if the instrument variables are uncorrelated with the residual by calculating the J-statistic $J = T\mathbf{M}'\mathbf{H}\mathbf{M}$, which is χ^2-distributed with $K - \dim(\gamma)$ degrees of freedom under the null hypothesis of zero correlation between the residual and the overidentifying instrument variables. In this case the overidentifying instruments are equal to the elements of the realized volatility vector and the J-test subsequently tests if $VSTOXX^*$ (the residual from equation 34) is uncorrelated with the realized volatility series.

This could also be expressed as GMM being a way to test the existence of additional information in $VSTOXX^*$ by simultaneously estimating equation 25 and 30. More precisely we estimate equation 25 with the imposed restriction that all elements of the slope vector in equation 30 are equal to zero. The authors of this thesis argue that although somewhat complex, the GMM-framework is justified due to the uncertain properties of the F-test described above.

5 Result

In this section we describe the result of the above methodology. We start by presenting the data material and its statistical properties. We continue by describing the forecast result from the used forecast methods. In the last section we present the test result of additional information in implied volatility.

5.1 Data and Descriptive Statistics in the Estimation Period

The dataset consists of the stock index Dow Jones EURO STOXX 50 (DJE50) between 1998-02-26 and 2009-04-01 (2820 daily observations). The realized volatility is constructed from high/low data on DJE50 obtained from Datastream. Estimation of the GARCH-model is done on continuously compounded returns of DJE50. The log returns are calculated from the DJE50 price series obtained from Datastream. The implied volatility series (VSTOXX) is obtained from stoxx.com.[66] We use a rolling window of 1000 observations to calculate the parameters for the chosen forecast models. Since the rolling window is rolled one day in each step we obtain 1820 in-sample periods which are based on 1000 historical observations of the daily realized volatility.

Table 2 Descriptive statistics of realized volatility

	Mean	Std.Dev	Skewness	Kurtosis	Jarque-Bera (JB)	P-Value JB	Min	Max
Mean	1,62	0,998	2,205	11,006	4617,085	0	0,333	8,286
Min	0,978	0,453	1,576	6,563	946,617	0	0,24	4,292
Max	2,148	1,304	4,366	32,224	38762,4	0	0,498	10,884
Std.Dev.	0,409	0,267	0,559	5,103	5878,791	0	0,099	1,636

This table summarizes the descriptive statistics of the realized volatility. We calculate the descriptive statistics from a window of 1000 observations, the window is rolled 1820 times for each statistical property. The descriptive statistics are aggregated and the mean value, min value, max value and standard deviation are shown for each statistical property.

Table 2 describes the descriptive statistics of the 1820 in-sample observations between 2002-02-05 and 2009-04-01. Both the mean and the standard deviation (Std.Dev) fluctuate through out the sample period which indicates that the realized volatility is not constant over time. An interesting observation is that the Jarque-Bera test shows that normally distributed realized volatility can be rejected in all sample periods. The distribution of realized volatility can be categorized as leptokurtic since it has fatter tails and is more peaked than the normal

[66] http://www.stoxx.com/data/historical/historical_strategy.html 2009-04-15

distribution.[67] This property is very common in financial time series. Non-normally distributed realized volatility is also found by Becker et. al. (2007) in their study of the VIX index. The result is thus not unexpected.

Table 3 Autocorrelation of realized volatility

Lag	1	2	3	4	5	10	20	100	252
Mean	0,597	0,581	0,569	0,545	0,533	0,471	0,374	0,147	0,113
Min	0,313	0,298	0,256	0,265	0,236	0,176	0,103	-0,122	-0,074
Max	0,757	0,759	0,771	0,748	0,756	0,72	0,707	0,814	0,833
Std.Dev	0,119	0,125	0,137	0,134	0,141	0,14	0,152	0,216	0,195
Percent Significant	100%	100%	100%	100%	100%	100%	100%	74%	54%
Total Insignificant	0	0	0	0	0	0	0	465	835

This table summarizes the autocorrelation structure of the realized volatility. We calculate the autocorrelation for each lag length from a window of 1000 observations, the window is rolled 1820 times.

As we can see in table 3, the realized volatility is autocorrelated at various lag lengths. The autocorrelation indicates that the present volatility depends on previous values of volatility. The autocorrelation is positive and significant in all cases up until the 20th lag. The positive autocorrelation indicates that the high (low) realized volatility is followed by high (low) volatility. Volatility is very persistent since autocorrelation in some periods is significant up until the 252nd lag. Hence shocks in the realized volatility die out very slowly and impact the volatility for a long time into the future. As mentioned in section 3.4.3 Granger et. al. investigate the statistical properties of volatility and conclude that the realized volatility has a long memory. The result of autocorrelated volatility is thus in line with previous research.

Table 4 Augmented Dickey-Fuller Unit Root Test H0 = Realized Volatility has a Unit Root

	Mean	Min	Max	Unit Roots	Number of Obs	Unit Root Not Rejected at 5%
Number of Lags	5,309	1	14		1820	
T-Stats	-4,854	-11,77	0,984		1820	
P-Value		0	0,997	164	1820	9,01%

This table summarizes the Augmented Dickey-Fuller unit root test of the realized volatility. We conduct the test from a window of 1000 observations, the window is rolled 1820 times. The test indicates that the null hypothesis cannot be rejected in 9 percent of the cases i.e. 164 of 1820.

Table 4 indicates that the null hypothesis of unit root in the realized volatility series is rejected in most cases. The result is however somewhat inconclusive since the null is not rejected in 164 of 1820 cases (9 percent). On the other hand we expect the null not to be rejected in 5 percent of the cases. A unit root indicates that the time series is non-stationary which has

[67] See Brooks (2008) p.162 for discussion

implications for e.g. the OLS estimator. If the assumption of stationarity is violated we can not draw any reliable conclusions from the result of the regression.[68]

5.2 Model Estimation

The parametric models used for forecasting the realized volatility over 21 days is estimated on an in-sample of 1000 observations. The models are re-estimated every day with a rolling window of 1000 observations which, means that the forecast is repeated 1820 times. Since the in-sample estimation window consists of 1000 observations and a forecast window of 21 observations, the authors of the thesis makes the implicit assumption that the estimated parameters from the in-sample period is sufficient to explain the variation of the realized volatility over the next 21 days. Furthermore a deterministic structure is assumed since historical data is used to forecast the future. We do however believe that the structure changes over time and it is therefore motivated to re-estimate the parameters. Stable and significant coefficients from the parametric models over different forecast periods indicate that the realized volatility follow some kind of deterministic structure. We examine the structure and the coefficients from the three different parametric forecast models. The first parametric model estimated is the ARMA(2,1), the stability of the parameters are displayed in table 5 and figure 3.

Table 5 ARMA(2,1) model estimation

ARMA(2,1)	C	T-Stat C	P-Value C	AR1	T-Stat AR1	P-Value AR1
Mean	1,61	8,44	0,00	0,97	20,90	0,00
Min	0,95	0,11	0,00	0,81	7,59	0,00
Max	8,44	20,05	0,91	1,23	31,02	0,00
Std.Dev.	0,47	4,30	0,04	0,07	1,98	0,00
Percent Significant			99,6%			100%
	AR2	**T-Stat AR2**	**P-Value AR2**	**MA1**	**T-Stat MA1**	**P-Value MA1**
Mean	0,00	0,05	0,37	-0,75	-22,70	0,00
Min	-0,25	-5,66	0,00	-0,91	-62,54	0,00
Max	0,07	1,64	1,00	-0,43	-4,17	0,00
Std.Dev.	0,08	1,66	0,27	0,04	5,04	0,00
Percent Significant			15%			100%

This table summarizes the coefficient estimates of the ARMA(2,1) model. We calculate each coefficient from a window of 1000 observations, the window is rolled 1820 times. The number of times each coefficient is significant is expressed in percent.

[68] See for example Brooks (2008) p. 319

The estimated coefficients of the ARMA(2,1) model are relatively stable over time. The constant coefficient C fluctuates over time and there is a wide span between its min and max value. The purpose of estimating the constant coefficient is to ensure that the OLS estimator gives unbiased estimates of the coefficients.[69] However, if we were to include the constant coefficient in the forecast it would lead to increasing volatility every day over the forecasting period. Accordingly the constant coefficient is excluded in the forecast period. The AR(1) and MA(1) coefficients are stable over time and are always significant. The AR(2) coefficient is insignificant in many estimations but its inclusion is motivated by its common use in the volatility forecasting literature.[70]

Figure 3 ARMA(2,1) Parameter stability and R^2 in the in-sample period

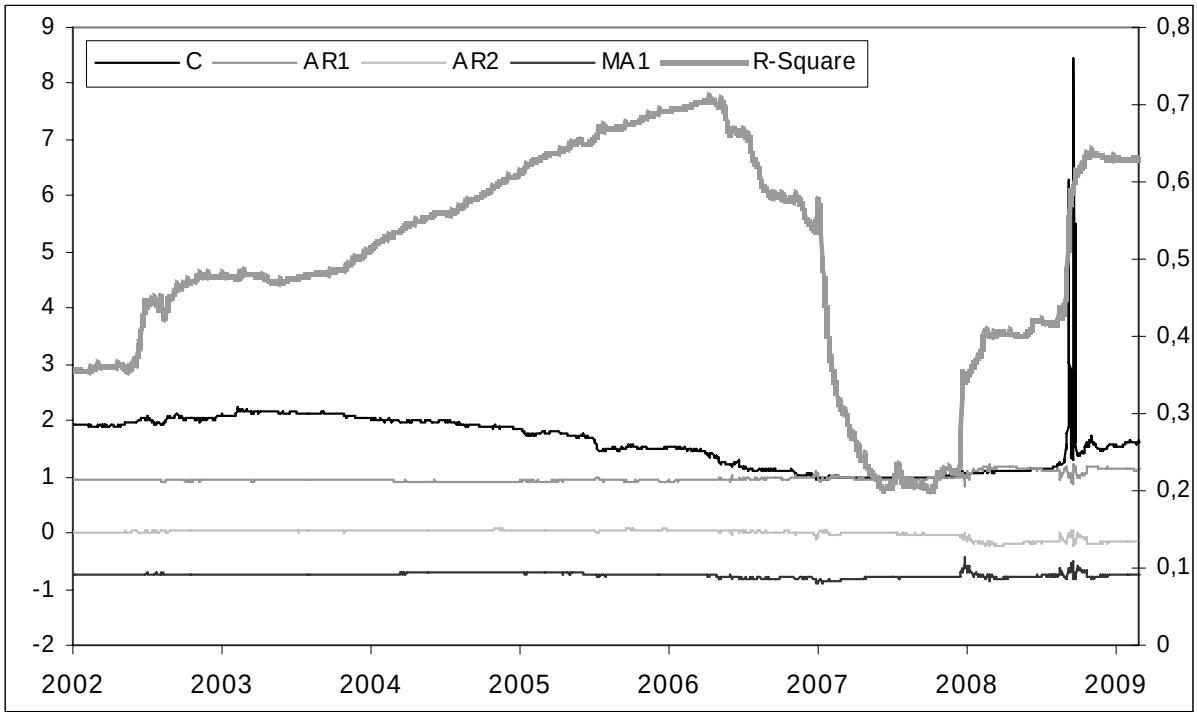


Figure 3 displays how the estimated parameters from the ARMA(2,1) model fluctuate over time. The parameters seem to be fairly stable but the constant parameter experiences drastic fluctuations over the last quarter of 2008 during the turbulent months of the financial crisis. The R^2, which describes the explanatory power in the in-sample period is also plotted in figure 3. The ARMA(2,1) model seems to be better at explaining the variation in the realized volatility than the ARIMA(1,1,1) model that is presented in figure 4. It is important to note

[69] Brooks(2008) p.131
[70] See section on methodology in section 4

that high explanatory power of past volatility does not necessarily mean that the model can forecast the future with high accuracy. We also check the stationarity condition of the coefficient estimates and find that all roots lie within the unit circle which implies stationarity.

Table 6 ARIMA(1,1,1) model estimation

ARIMA(1,1,1)	C	T-Stat C	P-Value C	AR1	T-Stat AR1	P-Value AR1
Mean	0,00	-0,06	0,78	0,03	0,81	0,39
Min	0,00	-1,70	0,09	-0,06	-1,46	0,00
Max	0,01	1,08	1,00	0,34	9,66	1,00
Std.Dev	0,00	0,38	0,18	0,11	2,79	0,29
Percent Significant			0%			17,6%
	MA1	**T-Stat MA1**	**P-Value MA1**			
Mean	-0,80	-34,42	0,00			
Min	-0,93	-71,75	0,00			
Max	-0,65	-15,48	0,00			
Std.Dev.	0,05	9,25	0,00			
Percent Significant			100%			

This table summarizes the coefficient estimates of the ARIMA(1,1,1) model. We calculate each coefficient from a window of 1000 observations, the window is rolled 1820 times. The number of times each coefficient is significant is expressed in percent.

Figure 4 ARIMA(1,1,1) Parameter stability and R-Square

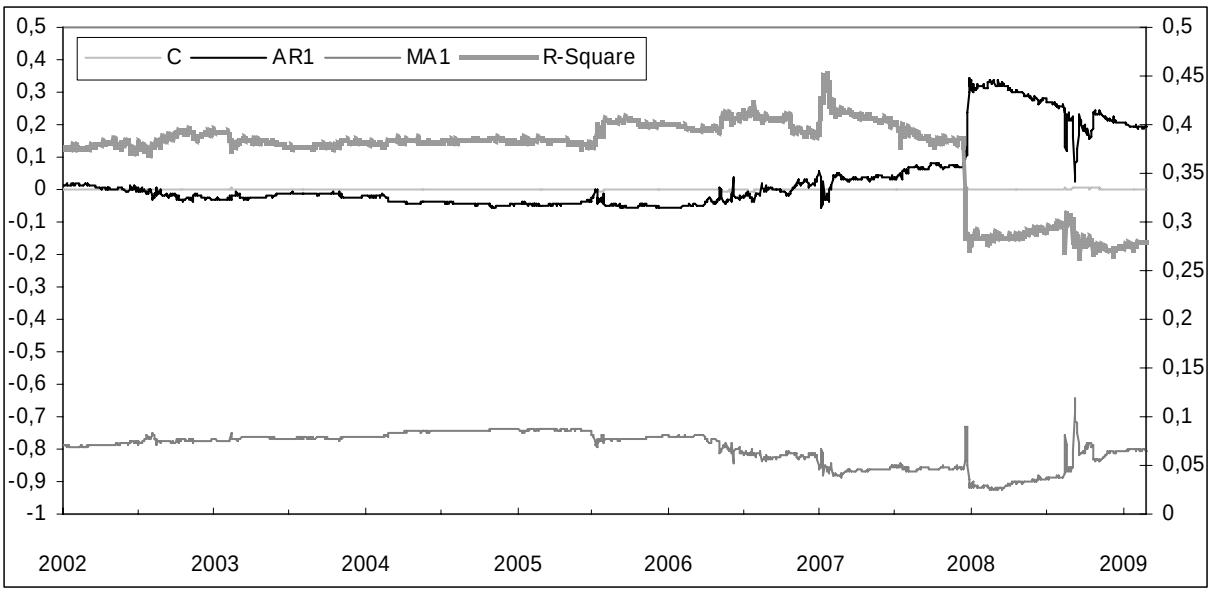


Table 6 and figure 4 indicate that the coefficients of the ARIMA(1,1,1) model are rather stable over time. There is some fluctuation in the beginning of 2008 and in the turbulent months in the end of 2008. The constant coefficient C is once more included to obtain unbiased coefficients in the OLS estimation. Since it would be a-theoretical if the volatility would follow a deterministic increasing trend over time, we expect the constant to be zero.

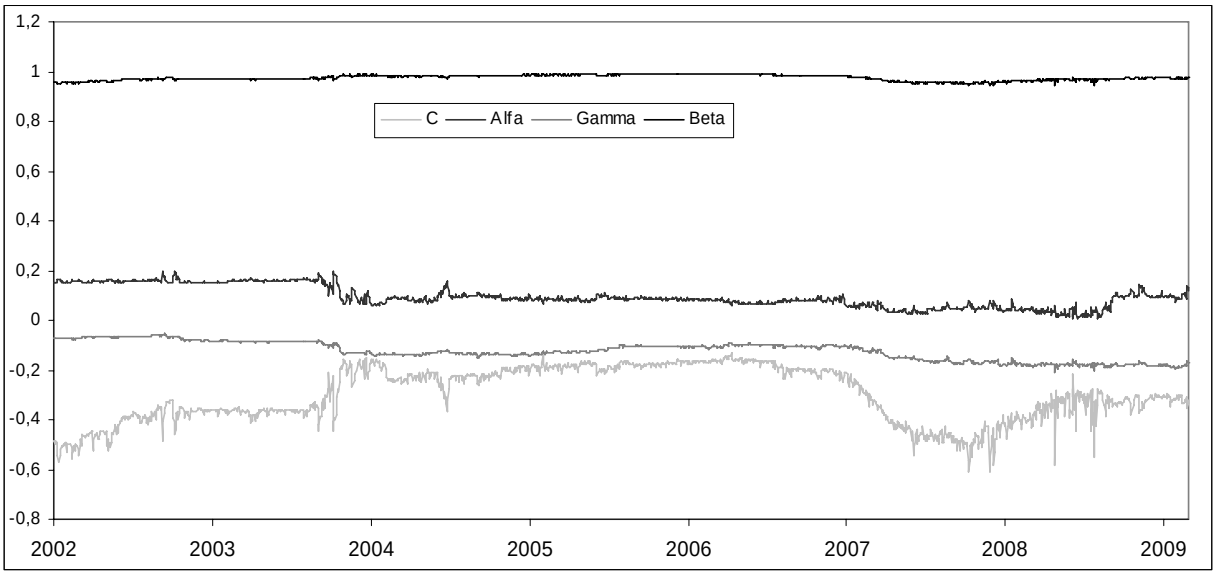
The result that the constant coefficient C is insignificant is therefore an expected result. The MA(1) coefficient is significant over the whole sample period and negative. The AR(1) coefficient varies between being negative and positive and is also insignificant in many cases. The model is also found to be stationary by an investigation of the roots of the characteristic equation.

Table 7 EGARCH model estimation

EGARCH	C	T-Stat C	P-Value C	Alfa	T-Stat Alfa	P-Value Alfa
Mean	-0,300	-4,981	0,000	0,094	3,629	0,025
Min	-0,610	-8,658	0,000	0,006	0,429	0,000
Max	-0,134	-3,430	0,001	0,198	7,112	0,668
Std.Dev	0,109	0,886	0,000	0,043	1,211	0,074
Percent Significant			100%			87%
	Gamma	**T-Stat Gamma**	**P-Value Gamma**	**Beta**	**T-Stat Beta**	**P-Value Beta**
Mean	-0,123	-6,765	0,000	0,975	206,967	0,000
Min	-0,209	-13,183	0,000	0,943	61,521	0,000
Max	-0,056	-3,046	0,002	0,991	413,546	0,000
Std.Dev	0,037	2,041	0,000	0,011	97,183	0,000
Percent Significant		100%				100%

This table summarizes the coefficient estimates of the EGARCH model. We calculate each coefficient from a window of 1000 observations, the window is rolled 1820 times. The number of times each coefficient is significant is expressed in percent.

Figure 5 EGARCH Parameter stability in the in-sample period



We estimate the four coefficients of the EGARCH-model using maximum likelihood. The coefficients are significant in most of the estimations. The Alpha coefficient, which captures the symmetric reaction to past shocks, is positive but not always significant. Gamma, which

captures the asymmetry of shocks, is negative and significant in all estimations. This indicates that there a is negative relationship between return and volatility, which is in line with previous empirical findings. The constant coefficient C is the only parameter that is really unstable over time. We also test if there are any ARCH-effects left in the error term after the EGARCH model has been estimated. We regress five lags of the squared standardized residuals on itself. The joint null hypothesis of zero slope coefficients is tested both with the χ^2 and the F-distributed test-statistic. We can reject the null in only 3 percent of the cases which indicates that the EGARCH-model captures the variation in variance quite well.

5.3 Data and Descriptive Statistic in the Forecast Period

We forecast over 1820 days between 2002-02-05 and 2009-04-01. Before we evaluate the forecasting ability of each model the statistical properties of the forecasts are examined. The forecast models we use are the EGARCH model, the ARIMA(1,1,1) model, the ARMA(2,1) model, the implied volatility from the VSTOXX index and the benchmark model based on historical volatility. In table 8, the statistical properties of the forecasts are displayed together with the realized volatility (RV) for different time horizons, where the column RVt21 exactly matches the horizon of the model based forecasts and VSTOXX. All numbers are expressed on a daily basis.

Table 8 Statistical properties of the forecast models

	EGARCH	ARIMA(1,1,1)	ARMA(2,1)	VSTOXX	Benchmark	RVt1	RVt5	RVt10	RVt15	RVt21
Mean	0,713	1,617	1,184	1,579	1,621	1,618	1,620	1,623	1,625	1,627
Median	0,558	1,239	0,903	1,312	1,737	1,199	1,223	1,223	1,259	1,266
Std.Dev	0,396	1,086	0,805	0,793	0,411	1,271	1,120	1,082	1,060	1,041
Skew	1,312	2,070	3,084	1,504	-0,299	2,478	2,138	1,931	1,827	1,741
Kurt	1,088	5,247	14,574	1,841	-1,473	8,622	5,700	4,106	3,414	2,880
Corr(RV21)	0,755	0,833	0,756	0,838	0,009	0,778	0,908	0,959	0,987	1,000
Corr(VIX)	0,870	0,931	0,890	1,000	0,152	0,792	0,864	0,865	0,854	0,838

This table summarizes the statistical properties of the average 21 day ahead forecast and the properties of the average t=1,5,10,15,21 day ahead realized volatility. The forecast window of 21 days is rolled 1800 times.

The realized volatility is positively skewed and experiences excess kurtosis, i.e. the distribution of the realized volatility is leptokurtic. The distribution of the model based forecasts experience excess kurtosis and positive skewness. The level of skewness and kurtosis differ between the models but all models except the benchmark model are able to capture the leptokurtic behaviour. As expected, the distribution of the benchmark model is notably different from the other models and the realized volatility.

The ARIMA(1,1,1) model and VSTOXX experience approximately the same mean and median as the realized volatility over 21 days. This is an indication that the models provide accurate forecasts of the realized volatility in levels. The ARMA(2,1) model and the EGARCH model differ in mean and median compared to realized volatility, which indicates a bias in the forecast. The benchmark model performs well in terms of mean and median. However, when the correlation with realized volatility is examined, it can be concluded that the benchmark model is uncorrelated with the realized volatility. This is a strong indication of poor ability to forecast the realized volatility. The model based forecasts are overall highly correlated with the realized volatility, which is an indication that the models have some forecast ability of future volatility. ARIMA(1,1,1) and VSTOXX seem to be the best models to predict realized volatility, since they experience the highest correlation with realized volatility. Even though EGARCH and the ARMA(2,1) differ in mean and median, the high correlation with the realized volatility indicates that the model produces fairly good estimates of the change in volatility from one period to the next.

Figure 6 Average 21 days ahead realized volatility in the forecasts period

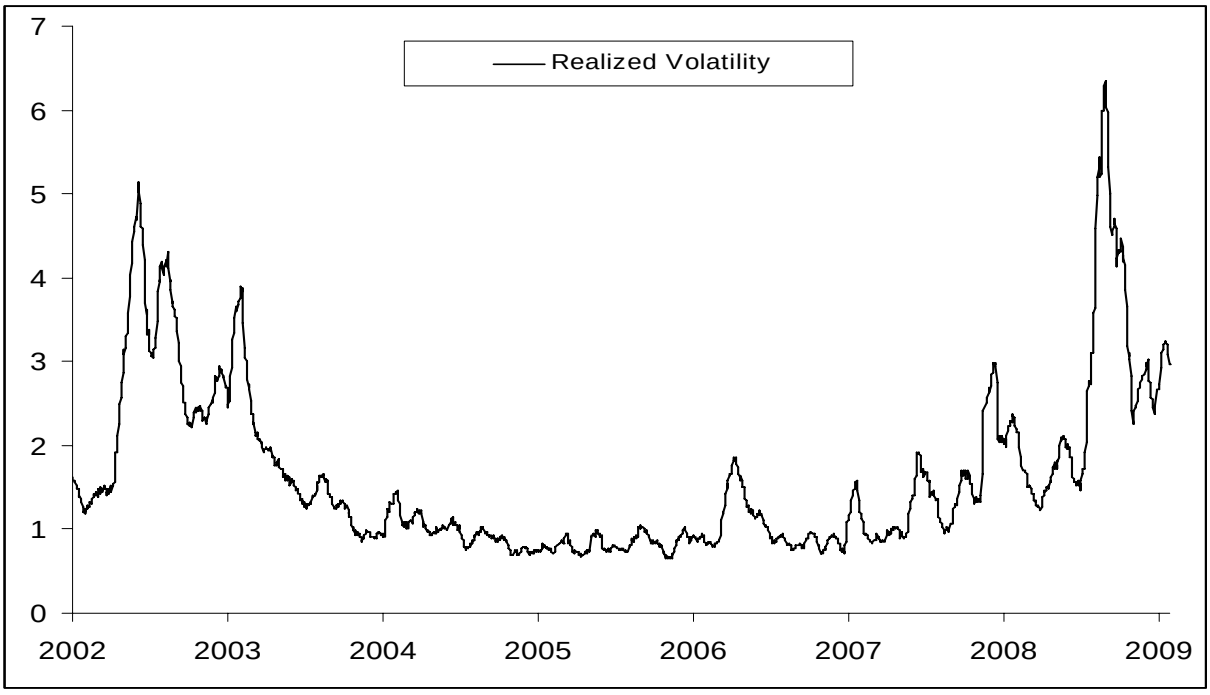


Figure 6 displays the average 21 day ahead realized volatility for the entire sample period. Worth noticing is that the realized volatility is high in the beginning of the sample period and in the end of the sample period in the turbulent month of the financial crisis. As we can see volatility is time varying.

5.4 Performance Measure Analysis

In this section the forecast models are evaluated using different performance measures. The performance measures evaluate the accuracy of a forecast by comparing the out-of-sample forecast with the realized (actual) value in each time step. The result of the different performance measures are summarized in table 9.

Table 9 Performance Measures Result

	EGARCH	ARIMA (1,1,1)	ARMA (2,1)	VSTOXX	Benchmark
MSE	1,45	0,38	0,66	0,33	1,24
MAE	0,91	0,38	0,52	0,33	0,87
Theil-U	0,73	0,35	0,44	0,32	1,00

This table summarizes the result of different performance measures for the forecast models. Definition of each measure can be found in section 4.

The Mean Squared Error (MSE) is a quadratic loss function and measures the average squared unexplained volatility between the model estimate and the realized volatility. We can conclude that VSTOXX produces the least average error. The ARIMA(1,1,1) model is ranked second after VSTOXX. The EGARCH model is even worse off than the naïve benchmark model.

The Mean Absolute Error (MAE) is a measure of the average absolute error between the forecast model and the realized volatility. The ARIMA(1,1,1) model and VSTOXX produce the most accurate forecasts according to the MAE measure.

The Theil-U statistic measures if the more complex forecasting models yield different forecasting results than the naïve benchmark model. VSTOXX and ARIMA(1,1,1) yield almost the same result and produce fairly low Theil-U statistics. EGARCH produces a worse Theil-U statistic than the other parametric models although it still outperforms the benchmark model. The difference between MAE/MSE and Theil-U is that Theil-U describes relative errors, i.e. the Theil-U statistic decreases when the level of realized volatility increases. In table 8, where the descriptive statistics of the forecasts are presented, we can see that EGARCH has a lower mean and median than realized volatility. This indicates that EGARCH is wrong in level which leads to poor performance according to MAE and MSE.

Table 10 Performance Measures Ranking

	EGARCH	ARIMA(1,1,1)	ARMA(2,1)	VSTOXX	Benchmark
MSE	5	2	3	1	4
MAE	5	2	3	1	4
Theil-U	4	2	3	1	5

This table summarizes the individual ranking for the forecast models based on the different performance measures.

One conclusion that can be drawn from table 10 is that the EGARCH model produces the worst forecast of realized volatility compared to the other sophisticated models. The ARIMA(1,1,1) model and VSTOXX perform almost equally well with a minor advantage for VSTOXX. The ARMA(2,1) model could be described as an average achiever. The identical ranking by MAE and MSE indicates that the model that deviates most frequently also has the greatest deviations in terms of magnitude. The fact that VSTOXX produces smallest forecast errors is in line with our expectations formed by previous research.

5.5 Predictive Power Analysis

The predictive power of the forecasting models are examined in this section. The predictive power is determined by the coefficient of determination, R^2. The realized volatility series is the average volatility during the 21 trading days following the forecast. Before the predictive power analysis is performed, the statistical properties of the forecast period are examined. The augmented Dickey-Fuller test cannot reject the null hypothesis that the realized volatility is non-stationary, i.e. has a unit root. In order to get a stationary series of realized volatility we divide the sample into three parts, the first part stretches from 2002-05-05 to 2003-05-07 and contains 316 observations. This part of the sample is also found to be non-stationary using the augmented Dickey-Fuller test. The middle part of the sample stretches from 2003-05-07 to 2007-06-18 and contains 1051 observations. This part is found to be stationary by the augmented Dickey-Fuller test. The third part of the sample stretches from 2007-06-18 to 2009-03-03 and is found to be non-stationary by the augmented Dickey-Fuller test. There is no support in economic theory for volatility being non-stationary. Non-stationarity of volatility would imply that the volatility would increase in every time step and finally explode. The result of non-stationary realized volatility series can therefore be disregarded from an economic point of view. From an econometric point of view the non-stationarity has to be considered since some of our methods require stationary data. The use of the OLS-

estimator on non-stationary data can lead to spurious regressions with abnormally high R^2 s.[71] In order to ensure stationarity and reliable results from the predictive power test, we run it on two different periods. The first period consists of the entire sample and the second period of the middle (stationary) part of the entire sample i.e. 2003-05-07 to 2007-06-18.

Table 11 Augmented Dickey-Fuller test

Sample Period 2002-02-05 to 2009-03-03				**Sample Period 2003-05-07 to 2007-06-18**			
Augmented Dickey-Fuller		T-Stat	P-Value	Augmented Dickey-Fuller		T-Stat	P-Value
		-2,314	0,168			-3,934	0,002
Test critical values:	1% level	-3,434		Test critical values:	1% level	-3,436	
	5% level	-2,863			5% level	-2,864	
	10% level	-2,568			10% level	-2,568	
H0 Realized Volatility has a unit root is not rejected				H0 Realized Volatility has a unit root is rejected			

This table summarizes the Augmented Dickey-Fuller unit root test of the realized volatility. The test is conducted over two different sample periods. The number of lags to include in the auxiliary is determined by Schwartz Information Criterion. Maximum number of lags is set to 24. For the entire sample period 23 lags were included while 22 lags were included in the shorter sample.

The result of the predictive power analysis for the two sample periods are presented in table 12 and table 13. All regressions are estimated with OLS with the average realized volatility over the 21 days following the forecast as the dependent variable. The standard errors are adjusted to account for heteroskedasticity and autocorrelation using Newey-West.

Turning the attention to table 12 and 13 we can see that all slope coefficients are significantly different from zero, indicating that all forecast models have explanatory power of future realized volatility. The R^2 coefficient, which describes the explanatory power of each model, is reasonably high. The R^2 coefficient is especially high in the sample period where we cannot rule out non-stationarity using the augmented Dickey-Fuller test. The R^2 coefficient decreases for all models in the stationary period (table 13). This is an indication that the non-stationarity is leading to a spurious relationship. The explanatory power of the EGARCH changes the least between the two periods. The best predictor of realized future volatility seems to be VSTOXX since it produces the highest R^2 over the whole sample period and is very close to the explanatory power of EGARCH in the shorter sample period.

[71] Brooks (2008) p.319

Table 12 Predictive power

Sample Period 2002-02-05- 2009-03-03						
	C	**VSTOXX**	**ARIMA(1,1,1)**	**ARMA(2,1)**	**EGARCH**	**R-Square**
Coefficient	-0,110	1,099**				0,702
Std Error	0,064	0,052				
P-Value	0,085	0,000				
Coefficient	0,335*		0,798*			0,694
Std Error	0,062		0,048			
P-Value	0,000		0,000			
Coefficient	0,468*			0,977**		0,572
Std Error	0,093			0,087		
P-Value	0,000			0,000		
Coefficient	0,209*				1,986*	0,570
Std Error	0,067				0,113	
P-Value	0,002				0,000	
Coefficient	0,041	0,613*	0,380*			0,723
Std Error	0,064	0,123	0,109			
P-Value	0,521	0,000	0,001			
Coefficient	-0,094	1,040**		0,065		0,703
Std Error	0,058	0,078		0,094		
P-Value	0,107	0,000		0,490		
Coefficient	-0,118	0,977**			0,282	0,705
Std Error	0,062	0,115			0,235	
P-Value	0,058	0,000			0,231	

* Indicates that the coefficient is significantly different from zero. ** indicates that the coefficient is significantly different from zero and that the null of a coefficient equal to one cannot be rejected. All standard errors are corrected to account for heteroskedasticity and autocorrelation.

Furthermore EGARCH is dominated by VSTOXX in the whole sample period. In the shorter sample period both EGARCH and VSTOXX are significant, which indicates that EGARCH potentially holds some additional information beyond that contained in VSTOXX.

When we look at the entire sample period, presented in table 12, VSTOXX seems to be an unbiased estimator of future realized volatility, since the constant coefficient C is insignificant and the null hypothesis of a slope coefficient equal to one cannot be rejected at the 5 percent level. However, since we cannot ensure that the time series is stationary, the results should be interpreted with caution.

Table 13 Predictive power

Sample Period 2003-05-07- 2007-06-18						
	C	**VSTOXX**	**ARIMA(1,1,1)**	**ARMA(2,1)**	**EGARCH**	**R-Square**
Coefficient	0,239*	0,683*				0,537
Std Error	0,058	0,050				
P-Value	0,000	0,000				
Coefficient	0,402*		0,599*			0,496
Std Error	0,048		0,045			
P-Value	0,000		0,000			
Coefficient	0,355*			0,779*		0,432
Std Error	0,061			0,073		
P-Value	0,000			0,000		
Coefficient	0,489*				1,010**	0,549
Std Error	0,037				0,064	
P-Value	0,000				0,000	
Coefficient	0,253*	0,461*	0,233*			0,555
Std Error	0,056	0,102	0,099			
P-Value	0,000	0,000	0,018			
Coefficient	0,205*	0,532*		0,243		0,553
Std Error	0,059	0,089		0,125		
P-Value	0,001	0,000		0,052		
Coefficient	0,343*	0,323*			0,585*	0,572
Std Error	0,063	0,110			0,155	
P-Value	0,000	0,003			0,000	

* Indicates that the coefficient is significantly different from zero. ** indicates that the coefficient is significantly different from zero and that the null of a coefficient equal to one cannot be rejected. All standard errors are corrected to account for heteroskedasticity and autocorrelation.

All other estimations have a slope coefficient significantly different from zero, thus indicating biased forecasts. If the constant is greater than zero and the slope coefficient smaller than one, the forecast overestimates the volatility in high volatility periods and underestimates it in periods of low volatility. Finally we can conclude that the explanatory power of the forecast models is fairly high even when the stationary condition of the OLS-estimator is satisfied. The model based forecasts perform fairly similar to VSTOXX in terms of predictive power.

5.6 Information Content Analysis

The information content test is designed to test if the forecast models contain any information about the volatility one day ahead. The information content test allows us to determine if the

forecast models contain any useful information about the future. The results are presented in table 14 below.

Table 14 Information content test

Sample Period 2002-02-05- 2009-03-03						
	C	**VSTOXX**	**ARIMA(1,1,1)**	**ARMA(2,1)**	**EGARCH**	**R-square**
Coefficient	-0,383*	1,266*				0,629
Std Error	0,077	0,061				
P-Value	0,000	0,000				
Coefficient	0,108*		0,933*			0,639
Std Error	0,044		0,034			
P-Value	0,014		0,000			
Coefficient	0,197*			1,199*		0,582
Std Error	0,069			0,063		
P-Value	0,005			0,000		
Coefficient	0,018				2,259*	0,488
Std Error	0,075				0,141	
P-Value	0,813				0,000	

This table summarizes the result of the information content test. The test is conducted on 4 forecasts models and the dependent variable is the realized volatility the following day. * indicates significantly different from zero at 5 percent level.

As we can see all slope coefficients are significantly different from zero. In other words we can conclude that there is some information about the one-step-ahead realized volatility contained in the forecasts.

5.7 Analysis of Additional Information in VSTOXX

In this section we present the result when testing if implied volatility contains any extra information of the realized volatility beyond what is already incorporated in the model based forecasts. Two different methods are used to examine the additional information content of VSTOXX presented in separate sections below. In this section we only analyze the period where the augmented Dickey-Fuller test indicates stationarity, i.e. 2003-05-07 to 2007-06-18.

5.7.1 Test of Additional Information in VSTOXX Using the F-test

In order to examine if VSTOXX contains any additional information compared to the model based forecasts, $VSTOXX^*$ is tested using a modified F-statistic which takes overlapping samples into account. $VSTOXX^*$,as defined in equation 26, is the part of VSTOXX which is

not captured by the model based forecasts. The vector of independent variables ω_t has the following elements $\omega_t = \{ARMA(2,1), ARIMA(1,1,1), EGARCH\}$. The slope coefficients for this regression are presented in table 15. The regression coefficients and standard errors presented in table 15 should be interpreted with some caution. Since the different forecasts are highly correlated the regression suffers from multicolinearity. A symptom of this is the negative coefficient in front of the ARMA forecast, which becomes positive if the ARIMA-model is excluded from the regression. Recall that the purpose of running equation 25 is to produce a residual that is orthogonal to all model based forecasts and therefore the coefficients are of minor interest. The residuals are also heavily autocorrelated which makes the standard errors of the coefficient biased downward. Please note that the coefficients are identical to those of the GMM-estimation with zero overidentifying restrictions presented in panel A of table 18 below. Recall that that a GMM-estimation with zero overidentifying restrictions is equivalent to an OLS-estimation.

Table 15 Slope coefficients of the orthogonalizing regression

	C	ARMA(2,1)	ARIMA(1,1,1)	EGARCH
Coefficient	0,478	-0,860	0,958	0,778
Std Error	0,014	0,063	0,054	0,040
P-Value	0,000	0,000	0,000	0,000
R-Square	0,86			

The table summarizes the coefficients obtained when estimating equation 25. The sample period for the regression is set to 2003-05-07- 2007-06-18.

We now test if there is any additional information about future realized volatility in $VSTOXX^*$. We construct the realized volatility vector **RV** which includes the elements RV_{t+j} which is the j-day ahead average realized volatility. The slope coefficient for each forecast horizon j is presented in table 16. The joint null hypothesis of zero slope coefficients is tested using the adjusted F-statistic proposed by Harvey and Newbold (2000).

Table 16 Test of additional information using F-statistics

RV							Harvey-Newbold Statistics	
	C	j=1	j=5	j=10	J=15	j=21	F-Stat	P-Value
Coefficient	-0,048	0,012	0,030	-0,016	0,027	-0,006	1,082	0,368
T-Stat	-3,625	1,211	1,030	-0,293	0,352	-0,104		
Coefficient	-0,029	0,028					2,210	0,137
T-Stat	-3,554	3,944						
Coefficient	-0,046		0,045				1,366	0,243
T-Stat	-4,161		4,395					
Coefficient	-0,047			0,046			1,039	0,393
T-Stat	-3,922			4,107				
Coefficient	-0,048				0,047		0,914	0,471
T-Stat	-3,824				3,985			
Coefficient	-0,048					0,047	0,822	0,534
T-Stat	-3,640					3,779		

The table summarize the result of the testing if there is any additional information in VSTOXX* about the level of future realized volatility. The F-statistics is used to test the joint null hypothesis of zero slope coefficients for the average j=1,5,10,15,21day ahead realized volatility. The test is conducted on the sample period 2003-05-07 to 2007-06-18.

Inspecting table 16 we can conclude that the null hypothesis of all slope coefficients being equal to zero cannot be rejected for any composition of the **RV**-vector. This is a clear indication that VSTOXX does not contain any extra information content about the level of volatility other than what is already captured by the model based forecasts.

Table 17 Test of additional information using F-statistics

dRV							Harvey-Newbold Statistics	
	C	j=1	j=5	j=10	J=15	j=21	F-Stat	P-Value
Coefficient	0,000	0,015	0,038	0,004	0,041	-0,052	35,836	0,000
T-Stat	-0,019	1,498	1,328	0,070	0,550	-1,020		
Coefficient	0,000	0,035					64,806	0,000
T-Stat	0,011	5,397						
Coefficient	0,000		0,047				24,432	0,000
T-Stat	0,050		5,812					
Coefficient	0,000			0,045			17,926	0,000
T-Stat	0,081			5,387				
Coefficient	0,000				0,110		15,304	0,000
T-Stat	0,042				5,066			
Coefficient	0,000					0,134	10,989	0,000
T-Stat	0,039					4,725		

The table summarize the result of the testing if there is any additional information in VSTOXX* about the changes in future realized volatility. The F-statistics is used to test the joint null hypothesis of zero slope coefficients for the change in average j=1,5,10,15,21 day ahead realized volatility. The test is conducted on the sample period 2003-05-07 to 2007-06-18.

Table 17 displays the result of regressing $VSTOXX^*$ on changes in realized volatility. The joint null hypothesis of zero slope coefficients is tested with the adjusted F-statistic. The elements of **dRV** are the change in volatility between t and the average volatility during the

period t to $t+j$. The result of the joint null hypothesis, that the slope coefficients are equal to zero, can be rejected for all compositions of **dRV**. In other words the test of the coefficients indicates that VSTOXX contains useful information beyond that contained in the model based forecasts when it comes to predicting changes in volatility.

5.7.1 Test of Additional Information in VSTOXX Using GMM-Estimation

As we concluded earlier the properties of the adjusted F-statistic are uncertain during longer forecast horizons. We therefore follow Becker et. al. (2007) and increase the reliability of the results by also testing for additional information using GMM-estimation. As described above GMM-estimation is a simultaneous estimation of equation 34 (identical to equation 25) and test of the orthogonality of $VSTOXX^*$ to future realized volatility.

Table 18 Test of additional information explaining **RV** using GMM

Panel A	**C**	**ARMA(21)**	**ARIMA(1,1,1)**	**EGARCH**	**J-Test**
Coefficient	0,478	-0,860	0,958	0,778	
T-Stat	8,109	-3,758	4,472	4,332	
P-Value	0,000	0,000	0,000	0,000	NA
Instruments	C ARMA(2,1) ARIMA(1,1,1) EGARCH				
Panel B					
Coefficient	0,484	-0,933	1,079	0,658	
T-Stat	8,893	-4,080	5,156	3,897	
P-Value	0,000	0,000	0,000	0,000	0,573
Instruments	C ARMA(2,1) ARIMA(1,1,1) EGARCH RVT1 RVT5 RVT10 RVT15 RVT21				
Panel C					
Coefficient	0,507	-0,993	1,086	0,704	
T-Stat	8,862	-4,180	4,962	4,110	
P-Value	0,000	0,000	0,000	0,000	0,165
Instruments	C ARMA(2,1) ARIMA(1,1,1) EGARCH RVT1				
Panel D					
Coefficient	0,459	-0,827	0,962	0,747	
T-Stat	8,069	-3,672	4,574	4,147	
P-Value	0,000	0,000	0,000	0,000	0,179
Instruments	C ARMA(2,1) ARIMA(1,1,1) EGARCH RVT5				
Panel E					
Coefficient	0,468	-0,843	0,979	0,727	
T-Stat	7,962	-3,651	4,592	4,075	
P-Value	0,000	0,000	0,000	0,000	0,105
Instruments	C ARMA(2,1) ARIMA(1,1,1) EGARCH RVT21				

The table summarize the result of the testing if there is any additional information in VSTOXX* about the level of future realized volatility. The P-value of the J-Statistic is used to test the joint null hypothesis that all overidentifying restrictions are valid. The test is conducted on the sample period stretching from 2003-05-07-2007-06-18.

Panel A presents the results of the GMM-estimation with zero overidentifying restrictions. As we stated above the results are identical to those obtained from the OLS-estimation of equation 25.

In panels B through E we then add different compositions of the **RV**-vector as instruments. Each instrument is assumed to be uncorrelated with the residual under the null hypothesis. In panel B through E we have more instruments than parameters to estimate. These superfluous instruments (the elements of **RV**) form what we refer to as the overidentifying restrictions. We assume under the null that all elements of **RV** are orthogonal to the residual, i.e. the correlation between future realized volatility and the residual is zero. This is equivalent to obtaining slope coefficients equal to zero when regressing **RV** on $VSTOXX^*$. The p-value of the J-test indicates that we cannot reject the null hypothesis of the overidentifying restrictions being valid. In other words, we cannot reject that the correlation between $VSTOXX^*$ and the elements of **RV** is equal to zero which verifies the results obtained when using the adjusted F-statistic.

Table 19 Test of additional information explaining **dRV** using GMM

Panel A	**C**	**ARMA(21)**	**ARIMA(1,1,1)**	**EGARCH**	**J-Test**
Coefficient	0,614	-1,073	0,740	1,268	
T-Stats	11,484	-4,930	3,629	7,614	
P-Value	0,000	0,000	0,000	0,000	0,042
Instruments	C ARMA(2,1) ARIMA(1,1,1) EGARCH dRVT1 dRVT5 dRVT10 dRVT15 dRVT21				
Panel B					
Coefficient	0,617	-1,077	0,745	1,261	
T-Stats	11,196	-4,783	3,549	7,519	
P-Value	0,000	0,000	0,000	0,000	0,001
Instruments	C ARMA(2,1) ARIMA(1,1,1) EGARCH dRVT1				
Panel C					
Coefficient	0,594	-1,049	0,868	1,033	
T-Stats	10,335	-4,471	3,841	5,658	
P-Value	0,000	0,000	0,000	0,000	0,001
Instruments	C ARMA(2,1) ARIMA(1,1,1) EGARCH dRVT5				
Panel D					
Coefficient	0,609	-1,162	1,047	0,861	
T-Stats	10,795	-4,271	3,930	4,335	
P-Value	0,000	0,000	0,000	0,000	0,009
Instruments	C ARMA(2,1) ARIMA(1,1,1) EGARCH dRVT21				

The table summarize the result of the testing if there is any additional information in VSTOXX* about the changes in the future realized volatility. The P-value of the J-Statistic is used to test the joint null hypothesis that all overidentifying restrictions are valid. The test is conducted on the sample period stretching from 2003-05-07- 2007-06-18.

We also test the elements of the **dRV** using GMM. Note that we have excluded the case with zero overidentifying restrictions since this is identical to the results presented in panel A of table 18. The results are interpreted the same way as before. If the p-value of the J-test is significant we can reject the null hypothesis of zero correlation between the element of **dRV** and the residual. Once again we can verify the result of the adjusted F-statistic.

6 Conclusion

In this section we discuss and draw conclusions from the empirical results. We first give some general observations about the empirical findings. We then discuss the results from the perspective of the research questions asked in the problem specification.

The empirical investigation of the data material suggests that volatility experiences most of the properties that has been suggested by previous research in the field. Volatility is leptokurtic and very persistent. Furthermore it seems like volatility has an asymmetric reaction to shocks where negative news lead to greater volatility than positive news.

6.1 Predictability of Volatility

According to our empirical finding volatility can be predicted to some extent both by time series models and implied volatility (VSTOXX). The forecasts share the leptokurtic property of realized volatility and are highly correlated with future realized volatility. In the predictive power test the models produce R^2 around 0.5 in the part of the sample that is shown to be stationary. For the entire sample period the predictive power of the models are slightly higher, here R^2s lie in the range 0.6 to 0.7. Even though the results from the entire sample period should be interpreted with some caution there are still strong indications of fairly good predictability. Although we do not fully succeed in modelling the underlying structure of volatility it seems like our deterministic models at least are able to capture and forecast some of the structure. The information content test also indicates that all models hold fairly much information about the one-day-ahead volatility.

6.2 Relative Performance of Volatility Forecasts

The evaluation of the different forecasts methods indicates that implied volatility (VSTOXX) generally outperforms the model based forecast in a bilateral comparison. In accordance with our expectations, VSTOXX is performing quite well in terms of the performance measures. Furthermore VSTOXX dominates some of the model based forecasts in the pairwise predictive power test. VSTOXX fails to dominate the ARIMA(1,1,1) both in the partial sample and in the entire sample period and EGARCH in partial sample period. This indicates that these time series models capture some property of future realized volatility not contained

in implied volatility. It is however worth mentioning that the time series models fail to dominate VSTOXX. EGARCH has a marginally higher explanatory power in the partial sample period. Even though there exists some contradictive results, we find indications that VSTOXX is outperforming the model based forecast in pairwise comparisons.

6.3 The Bias of Volatility Forecasts

The bias of the volatility forecast is investigated by inference around the intercept and slope parameter in the predictive power test. Recall that an unbiased forecast will result in an intercept of zero and unity in the slope coefficient. We do find VSTOXX to be an unbiased estimator although only in the entire potentially non-stationary sample period. The observed bias of volatility forecasts is in line with much of the previous research in the field. Granger and Poon (2003) argue that a bias of implied volatility as a volatility forecast could indicate a structural mispricing of the options used to derive it. A significant intercept in the predictive power test of VSTOXX could also indicate the existence of a volatility risk premium as suggested by e.g. Chernov (2001, cited in Granger and Poon, 2003) who analyzed the VIX.

6.4 Additional Information

The last leg of the analysis involved testing if implied volatility (VSTOXX) holds any additional information beyond that captured by the model based forecasts. The null hypothesis of no additional information was tested both by the adjusted F-statistic and in a GMM framework to ensure reliability of the results. We found no indication of additional information in VSTOXX once VSTOXX had been orthogonalized to the vector of model based forecasts. This result is in line with the findings of Becker et. al. (2007) in their study of S&P 500 and VIX. We do however find that VSTOXX holds some information about changes in future realized volatility not captured by the time series models. This result contradicts the findings of Becker et. al. One intuitive explanation for the results could be that VSTOXX is derived from the market participant's expectations about the future and that it therefore is able to predict changes in future volatility.

6.5 Concluding Remarks

During the work on this thesis we have been forced to make some delimitations which might have affected the results. First of all, choosing other time series models could have yielded different results. On the other hand, the models actually chosen have been proven to produce high-quality forecasts in the past and were therefore highly motivated. Another issue regards

the length of the rolling estimation window for parameter estimation. Here we once more relied on previous empirical findings when deciding the length of the estimation window. Relatively good model fit and parameter stability indicate that the length of the estimation window is fairly good. Furthermore it could always be argued that the sample could have been expanded and chosen to include other indices or asset classes to ensure robustness of the results. Since our primary interest was to see if the results from VIX translated to a European setting we made the choice to only investigate VSTOXX. The test of additional information in implied volatility could favourably be investigated in e.g. the currency market.[72] This and other improvements of our work are left to future studies.

[72] Currency options are often traded over-the-counter and are therefore always traded at-the-money which reduces the problem induced by volatility smiles

7 References

7.1 Published References

Alizadeh, S.Brandt. M.W, Diebold F.X. 2002, "Range-Based Estimation of Stochastic Volatility" *The Journal of Finance*, Vol. 57, No. 3 (Jun., 2002), pp. 1047-1091

Andersen, T, G, Bollerslev, T, 1998 "Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecasts" *International Economic Review,* Vol. 39, No. 4 , pp. 885-905

Andersen, T, G, Bollerslev, T, Diebold, F,X, Labys, P 2003, "Modeling and Forecasting Realized Volatility" *Econometrica* Vol. 71, No. 2 , pp. 579-625

Bekaert G., Wu G., 2000, "Asymmetric Volatility and Risk in Equity Markets", *The Review of Financial Studies*, Vol. 13, No. 1, p. 1-42

Becker, R. Clements .A.E, White, S.I (2007), "Does implied volatility provide any information beyond that captured in model-based volatility forecasts?" *Journal of Banking & Finance*, Vol: 31 No: 8 p. 2535-2549

Black F., Scholes M., 1973, "The Pricng of Options and Corporate Liabilities", *The Journal of Political Economy*, Vol. 81, No. 3, p. 637-654

Blair B, J, Poon S, Taylor S, J. 2001, "Forecasting S&P 100 volatility: The incremental information content of implied volatilities and high-frequency index returns", *Journal of Econometrics,* Vol. 105, No. 1, p. 5-26

Bollerslev T, 1986, "Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics,* Vol. 31, No. 3 , pp. 307-328

Brooks C., 2008, Introductory Econometrics for Finance 2nd edition, Cambridge University Press

Canina, L, Figlewski, S, 1993, "The Informational Content of Implied Volatility", *Review of Financial Studies,* Vol. 6, No. 3, p. 659-681.

Champbell J.Y., Lo A.W., MacKinlay A.C., 1997, The Econometrics of Financial Markets, Princeton University Press

Demeterfi K., Derman E., Kamal M., Zou, J., 1999 "More than You Ever Wanted to Know about Volatility Swaps (But Less than Can Be Said)", Goldman Sachs Quantitative Strategies Research Notes

Dunis C.L, Huang X., 2002, “Forecasting and trading currency volatility: An application of recurrent neural regression and model combination” *Journal of Forecasting* 2002, Vol 21, No 5 (2002), pp 317-354.

Engle R.F., Lilien. D.M., Robins P. 1987, “Estimating Time Varying Risk Premia in the Term Structure: The Arch-M Model” *Econometrica*, Vol. 55, No. 2 pp. 391-407

Figlweski S., 1997, ”Forecasting volatility”, *Financial Markets, Institutions & Instruments*, Vol. 6, No. 1, p. 1-88

Glosten R., Jagannathan R., Runkle D.E, 1993, ”On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks” Vol. 48, No. 5, p. 1779-1801

Granger C.W.J., Ding Z., Spear S., 2000, “Stylized Facts on the Temporal and Distributional Properties of Absolute Returns: An Update” Working paper, University of California San Diego

Granger, C.W.J., Newbold, P., 1976, Forecasting Economic Time Series, Academic Press

Granger C.W.J, Poon S-H, 2003, “Forecasting Volatility in Financial Markets: A Review”, *Journal of Economic Literature*, Vol. 41, No. 2, p. 478-539

Jorion, P, 1995, “Predicting Volatility in the Foreign Exchange Market”, *The Journal of Finance*, Vol. 50, No. 2 , pp. 507-558

Harvey, D, Newbold, P, 2000 “Tests for multiple forecast encompassing” Journal of Applied Econometrics, Vol. 15, No. 5 , pp. 471-482

Hull, J.C, 2006, Options, Futures and Other Derivatives 6th edition, Prentice Hall

Hull J.C., White A., 1998, “Value at risk when daily changes in market variables are not normally distributed”, *Journal of Derivatives*, Vol. 5, No. 3, p. 9-19

Hull.J.C., White,A., 1987, “The Pricing of Options on Assets with Stochastic Volatilities” *The Journal of Finance*, Vol. 42, No. 2 (Jun., 1987), pp. 281-300

Lamoureux, C,G. Lastrapes W,D. 1993, “Forecasting Stock-Return Variance: Toward an Understanding of Stochastic Implied Volatilities”, *Review of Financial Studies,* Vol. 6, No. 2, p. 293-326

Li Kai, 2002, “Long-memory versus option-implied volatility predictions”, *Journal of Derivatives,* Vol. 9, No. 3 , pp. 9-25

Mandelbrot B., 1963, ”The Variation of Certain Speculative Prices”, *The Journal of Business*, Vol. 36, No. 4, p. 394-419

Martens M., Zein J., 2004, “Predicting financial volatility: High-frequency time-series forecasts vis-a-vis implied volatility”, *Journal of Futures Markets*, Vol. 24, No. 11, p. 1005-1028

Nelson, D, B. 1991, “Conditional Heteroskedasticity in Asset Returns: A New Approach” *Econometrica,* Vol. 59, No. 2 , pp. 347-371

Newey W, K, West, K, D, 1987, “Hypothesis Testing With Efficient Method Of Moments Estimat” International Economic Review, Vol. 28, No. 3 , pp. 777-788

Pong S., Shackleton M.B., Taylor S., Xu Xinzhong, 2004, “Forecasting currency volatility: A comparison of implied volatilities and AR(FI)MA models”, *Journal of Banking & Finance*, Vol. 28, No. 10, p. 2541-2563

Taylor, J, W, 2004 “Volatility forecasting with smooth transition exponential smoothing”, *International Journal of Forecasting,* Vol. 20, No. 2 , pp. 273-286

Veerbeek, M, 2008, A guide to modern econometrics, John Willy and Son

7.2 Electronic References

Dow Jones STOXX Index Guide (2009), Downloaded 2009-04-15, Web address: http://www.stoxx.com/download/indices/rulebooks/djstoxx_indexguide.pdf

Dow Jones STOXX Historical Data Downloaded 2009-04-15, Web address: http://www.stoxx.com/data/historical/historical_strategy.html

Eviews 6 User Guide II, (2007-03-09), Downloaded 2009-05-10, Web address: http://www.eviews.com/download/download.html

8 Appendix

A1 Eveiws Code for Forecasting with EGARCH

```
scalar noobs=2820
scalar nw=1000
matrix(2820,21) garchforecast
vector(21) forecastresult

for !i=nw to noobs
        smpl !i-nw+1 !i
        equation garch_n
        garch_n.arch(egarch) stoxxr c
        garch_n.makegarch garchcondvar
        for !j=1 to 21
                'For the forecast one-day-ahead, the conditional variance used as input
                'is the last observation of the conditional variance series garchcondvar
                if !j=1 then
                        forecastresult(1)=@EXP(garch_n.@coefs(2)+
                        garch_n.@coefs(3)*@ABS(resid(!i)/@SQRT(garchcondvar(!i)))+
                        garch_n.@coefs(4)*(resid(!i)/@SQRT(garchcondvar(!i)))+
                        garch_n.@coefs(5)*@LOG(garchcondvar(!i)))
                 endif
                'For forecasts more than one-day-ahead, the conditional variance used as input
                'is equal to the last forecasted conditional variance
                if !j>1 then
                        forecastresult(!j)=@EXP(garch_n.@coefs(2)+
                        garch_n.@coefs(3)*@ABS(@SQRT(forecastresult(!j1))/@SQRT(garchcondvar(!i)))+
                        garch_n.@coefs(4)*(@SQRT(forecastresult(!j-1))/@SQRT(forecastresult(!j-1)))+
                        garch_n.@coefs(5)*@LOG(forecastresult(!j-1)))
                endif
        next !j
        rowplace(garchforecast,@transpose(forecastresult),!i)
next !i
```

A2 Eviews Code for Forecasting with ARIMA(1,1,1)

```
scalar noobs=2820
scalar nw=1000
matrix(2820,21) armaforecast
vector(21) forecastresult

for !i=nw to noobs
        smpl !i-nw+1 !i
        equation arma_n
        arma_n.ls d(rvdaily) c ar(1) ma(1)
        for !j=1 to 21
                if !j=1 then
                        forecastresult(1)=rvdaily(!i)+
                        arma_n.@coefs(2)*(rvdaily(!i)-rvdaily(!i-1))+
                        arma_n.@coefs(3)*resid(!i)
                endif
                if !j=2 then
                        forecastresult(!j)=forecastresult(!j-1)+
                        arma_n.@coefs(2)*(forecastresult(!j-1)-rvdaily(!i))
                endif
                   if !j>2 then
                        forecastresult(!j)=forecastresult(!j-1)+
                        arma_n.@coefs(2)*(forecastresult(!j-1)-forecastresult(!j-2))
                endif
        next !j
        rowplace(armaforecast,@transpose(forecastresult),!i)
next !i
```

A3 Eviews Code for Forecasting with ARMA(2,1)

```
scalar noobs=2820
scalar nw=1000
matrix(2820,21) armaforecast
vector(21) forecastresult

for !i=nw to noobs
	smpl !i-nw+1 !i
	equation arma_n
	arma_n.ls rvdaily c ar(1) ar(2) ma(1)
	for !j=1 to 21
		if !j=1 then
			forecastresult(1)=arma_n.@coefs(2)*rvdaily(!i)+
			arma_n.@coefs(3)*rvdaily(!i-1)+
			arma_n.@coefs(4)*resid(!i)
		endif
		if !j=2 then
			forecastresult(2)=arma_n.@coefs(2)*forecastresult(1)+arma_n.@coefs(3)*rvdaily(!i)
		endif
		if !j=3 then
			forecastresult(3)=arma_n.@coefs(2)*forecastresult(2)+arma_n.@coefs(3)*forecastresult(1)
		endif
		if !j>3 then
			forecastresult(!j)=arma_n.@coefs(2)*forecastresult(!j-1)+
			arma_n.@coefs(3)*forecastresult(!j-2)
		endif
	next !j
	rowplace(armaforecast,@transpose(forecastresult),!i)
next !i
```

A4 Eviews Code for Calculating the Adjusted F-statistic

```
'Declare the number of observations, number of independent variables, number of overlapping
‘observations in the forecast and number of the first observation
!n=1051
!noindependent=2
!h=21
!startobservation=1314

'Declare the vector of independent variables and the residual vector
matrix(1051,!noindependent) independent
vector(1051) residualvector

'Declare and estimate the OLS-equation
‘In this case with the independent variables rvt1 rvt5 rvt10 rvt15 rvt21
equation testorto
testorto.ls vstoxxstar c rvt1 rvt5 rvt10 rvt15 rvt21

'Declare the beta vector
vector(!noindependent) betavector

'Store regression coefficients in the beta vector
betavector(1)=testorto.@coefs(1)
betavector(2)=testorto.@coefs(2)
betavector(3)=testorto.@coefs(3)
betavector(4)=testorto.@coefs(4)
betavector(5)=testorto.@coefs(5)
betavector(6)=testorto.@coefs(6)
```

```
for !i=1 to 1051
        'Store the realized volatility series in the independent  variable vector
        ' the first column is filled with ones for the intercept
        independent(!i,1)=1
        independent(!i,2)=rvt1(!i+!startobservation)
        independent(!i,2)=rvt5(!i+!startobservation)
        independent(!i,2)=rvt10(!i+!startobservation)
        independent(!i,2)=rvt15(!i+!startobservation)
        independent(!i,2)=rvt21(!i+!startobservation)
        'Store the residual from the regression in the residual vector
        residualvector(!i)=resid(!i+!startobservation)
next !i

'Declare the Q matrix
matrix(!noindependent,!noindependent) Q

for !i=1 to !noindependent
        for !j=!i to !noindependent
                'a,b and c refer to the three different summations in equation 32
                !a=0
                !b=0
                !c=0
                for !t=1 to !n
                        !a=!a+independent(!t,!i)*independent(!t,!j)*residualvector(!t)^2
                next !t
                for !m=1 to !h-1
                        for !t=!m+1 to !n
                                !b=!b+independent(!t,!i)*independent(!t-!m,!j)*residualvector(!t)*residualvector(!t-!m)
                        next !t
                next !m
                for !m=1 to !h-1
                        for !t=!m+1 to !n
                                !c=!c+independent(!t-!m,!i)*independent(!t,!j)*residualvector(!t)*residualvector(!t-!m)
                        next !t
                next !m
                'Since the Q matrix is symmetrical, two elements are calculated in each iteration
                Q(!i,!j)=(1/!n)*(!a+!b+!c)
                Q(!j,!i)=(1/!n)*(!a+!b+!c)
        next !j
next !i


'Calcualte the m matrix
matrix(!noindependent,!noindependent) m
m=(1/!n)*@transpose(independent)*independent

'Calculate the d matrix
matrix(!noindependent,!noindependent) dmatrix
dmatrix=@inverse(m)*Q*@inverse(m)

'Calculate the f-statistic and store in the vector fresult
vector(1) fresult
fresult=(1/(!noindependent-1))*!n*@transpose(betavector)*@inverse(dmatrix)*betavector

'Calculate the p-value for the f-statistic and store in the vector fresultpvalue
vector(1) fresultpvalue
fresultpvalue=1-@cfdist(fresult(1),!noindependent-1,!n-!noindependent+2)
```