UNIVERSITY OF TURKU
Turku School of Economics

Comparison of Implied Volatility and GARCH(1,1)

Evidence from the German stock market

Accounting and Finance

Bachelor's thesis

Author:
Anton Salonen

Supervisor:
Md Khaled Hossain Rafi

3.4.2024

Turku

Bachelor's thesis

Subject: Accounting and Finance
Author: Anton Salonen
Title: Comparison of Implied Volatility and GARCH(1,1): Evidence from the German stock market
Supervisor: Md Khaled Hossain Rafi
Number of pages: 33 pages
Date: 3.4.2024

Risk and thus volatility is a key concept in several different financial theories. As by definition risk is uncertain, it is necessary to forecast it, if one wants information about future volatility. This study's goal is to compare two competing volatility forecasting models: implied volatility and GARCH(1,1).

Implied volatility represents one of the two main lines of volatility forecasting, as it is calculated based on option prices. It is seen as the investors' expectations about future volatility and thus found to be more informative than models that represent the other main line of volatility forecasting. The other line is models based on financial time series data and this line is represented by the GARCH(1,1) model in this study as it is often found to outperform other similar models.

In this study the forecasts are made with data from the DAX index and the sample period of 2019–2023 is divided into two subperiods. The comparison of the models' forecasting performance is measured with three commonly used error metrics: mean squared error, root mean squared error, and mean absolute percentage error.

The results of this study suggest that the GARCH(1,1) was able to outperform the implied volatility during both of the subperiods. The forecasting performance of the models was also better during the second subperiod, which is the less volatile of the two subperiods. The results of this study are not consistent with the majority of previous studies, as implied volatility is thought to be superior in comparison to models based on financial time series data.

Key words: Risk, Volatility, Volatility forecasting, Implied volatility, GARCH model

Kanditutkielma

Oppiaine: Laskentatoimi ja rahoitus
Tekijä: Anton Salonen
Otsikko: Implisiittisen volatiliteetin ja GARCH mallin vertaaminen Saksan osakemarkkinoilla
Ohjaaja: Md Khaled Hossain Rafi
Sivumäärä: 33 sivua
Päivämäärä: 3.4.2024

Riski ja täten myös volatiliteetti ovat tärkeitä konsepteja useiden rahoituksen teorioiden taustalla. Riskillä tarkoitetaan epävarmuutta ja mahdollisuutta todennäköisestä poikkeavaan lopputulokseen, joten sen ennustaminen on tarpeellista, mikäli markkinoiden tulevasta volatiliteetista halutaan informaatiota. Tämän tutkimuksen tarkoitus on verrata kahta kilpailevaa volatiliteetin ennustemallia: implisiittinen volatiliteetti ja GARCH(1,1).

Implisiittinen volatililiteetti edustaa yhtä kahdesta volatiliteetin ennustamisen päälinjasta. Implisiittinen volatiliteetti lasketaan optiohintojen perusteella, joten sen on määritelty olevan sijoittajien odottama markkinoiden volatiliteetti. Implisiittisen volatiliteetin on todettu olevan informatiivisempi ennustemalli kuin mallit, jotka edustavat volatiliteetin ennustamisen toista päälinjaa eli aikasarjamallit. Tässä tutkimuksessa toista päälinjaa eli aikasarjamalleja edustaa GARCH(1,1)-malli, jonka on usein havaittu menestyvän muita vastaavia malleja paremmin.

Tässä tutkimuksessa ennusteet on luotu DAX-indeksille vuosilta 2019–2023. Otos on tutkimuksessa jaettu vielä kahteen pienempään osajaksoon. Mallien suoriutumista vertaillaan kolmella yleisesti käytetyllä virhemittarilla: MSE, RMSE and MAPE.

Tämän tutkimuksen tulokset viittaavat siihen, että GARCH(1,1) onnistui implisiittistä volatiliteettia paremmin ennustamaan tulevaa volatiliteettia tutkimuksen molempien ajanjaksojen aikana. Molemmat mallit pystyivät ennustamaan volatiliteettia tarkemmin tutkimuksen toisen ajanjakson aikana, jolloin markkinat olivat vähemmän volatiiliset kuin ensimmäisen ajanjakson aikana. Tämän tutkimuksen tulokset eivät ole linjassa aikaisempien tutkimusten kanssa, sillä implisiittisen volatiliteetin on havaittu olevan parempi volatiliteetin ennustemalli kuin aikasarjamallit.

Asiasanat: Riski, Volatiliteetti, Volatiliteetin ennustaminen, Implisiittinen volatiliteetti, GARCH malli

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 Introduction

One of the fundamentals of finance is the concept of risk and it is considered in numerous financial theories. Often in everyday life when talking about risk, it is understood as a chance of bad consequences or exposure to mischance. In the world of finance, risk does not always mean the chance of a loss or an unwanted outcome but also the chance of a return or an outcome that is greater or better than was statistically expected. In financial markets risk is strongly related to uncertainty and the measures calculating risk are made to calculate the size of the uncertainty (McNeil et al., 2015).

Often risk is measured with volatility, which is the spread of a random walk variable's different outcomes. In financial markets, volatility can be used as a measure to capture the aforementioned uncertainty. In other words, volatility is measured as the sample standard deviation. Sometimes volatility and thus risk can also be measured with variance. While variance simply is the square of standard deviation, the use of sample standard deviation is still preferred as it has the same measure unit as the mean, if for example, the mean is in euros, then the standard deviation would also have euros as its unit while variance would have euros squared as its measure (Poon, 2005).

Poon (2005) highlights the importance and usefulness of volatility as the measure of risk even though he also notes that volatility is not a perfect measure of risk, as it only gives information about the spread of the distribution of returns but none on the shape of the distribution. The exception for this is a situation where the returns follow a normal distribution.

While volatility should not be used as the sole measure for risk in a financial institution, it is still an important part of a great number of financial theories which makes volatility one of the most important measures in modern financial theory. Higher volatility leads to a larger variation of returns and thus higher risk (Poon, 2005).

1.1 Volatility forecasting

As volatility is a vital component of many financial theories, the topic has become vastly researched and of high interest to almost anyone operating within financial markets according to Figlewski (1997). For example, in option pricing with the basic Black & Scholes (1973) pricing formula, is the future volatility of the option the only parameter that cannot be extracted directly from the market, so it must be forecasted somehow. Volatility forecasting became of significant interest to investors in the 1990s, as the maturities of financial instruments lengthened significantly. During the 1970s, common options had maturities of a few months while now the maturities of options can even reach over ten years. This brings the market volatility's effect on the value of the option to a new level and thus makes the need for accurate volatility forecasts obvious (Figlewski, 1997).

While historical volatility can be simply calculated from the historical data as the sample standard deviation, the calculation of future values for volatility is more challenging. A common way to forecast volatility is to simply extrapolate past volatility values into the future. While this is a simple and non-time-consuming method, it hardly makes accurate forecasts, which may be problematic for investors (Figlewski, 1997).

According to Muzzioli (2010) the modern academic research regarding volatility forecasting is usually divided into two main lines: forecasting models using information from the derivatives markets and forecasting with models, that are based on financial time series data. While there are a great number of models representing both lines of research, this study compares the forecasting accuracy of implied volatility by the Black & Scholes (1973) option pricing model and the GARCH(1,1) model by Bollerslev (1986).

1.2 Objective of the study

The topic of volatility forecasting has been vastly researched in the past, but the results gained from the research are still not completely clear. The previous research has found both the GARCH model and implied volatility measures to outperform one another under different circumstances (see for example Canina & Figlewski, 1993; Christensen & Prabhala, 1998).

As most of the research is from the 1990s and the early 2000s, it is necessary to compare the models under the most recent time of highly volatile market conditions. Uddin et al. (2021) found that during the Covid-19 pandemic, the market conditions were highly volatile, especially in developed markets such as the German stock market. It is yet to be determined whether volatility forecasting in highly volatile conditions still follows the previous findings from mostly more than 20 years ago. The recent developments of the market with the time of high volatility present a great opportunity to test the volatility forecasting models and see whether the findings from previous literature still stand.

The forecasts in this study are made for a German stock market index DAX (Deutschen Aktien IndeX). Following Muzzioli (2010), the DAX index is chosen for this study for two main reasons. First, the options are of European style, which is a necessity when using implied volatility by Black & Scholes (1973) and therefore the issues of early exercise are avoided in this study. Second, the DAX index is a capital-weighted index, that was made of 30 of the major German stocks until in September 2021 the index was expanded to follow 40 of the major German stocks. The index is adjusted for stock splits, dividends, and changes in capital. Thus the assumption that dividend payments do not affect the index value is made. According to Sorokina & Booth (2022), the options in the German stock market, such as the DAX index options are known to be highly liquid and accessible. Wallmeier & Hafner (2000) state that DAX options (ODAX) are ranked among the world's most liquid index options.

Frennberg & Hansson (1995) found that implied volatility is not an accurate measure of future volatility in smaller markets, such as Sweden. This makes the Frankfurt Stock Exchange and the DAX index a qualified base for forecasts with implied volatility and the GARCH model.

In this study, the implied volatility is calculated for call options, as the option pricing model by Black & Scholes (1973) assumes that the volatility of an option is constant over the option's lifetime. Thus the implied volatility of two options with the same lifetime should be the same (Äijö 2008).

This study examines forecasts by two different volatility forecasting models during a five-year period from the beginning of 2019 until the end of 2023. To obtain information about the performance of the forecasting models during different market circumstances, the period is divided into two subperiods, which are the height of the Covid-19 pandemic and after the pandemic.

The purpose of this study is to answer the following research question: Why is the implied volatility a better forecasting method for future volatility than the GARCH model? The study also aims to identify the differences between the two forecasts and to see which one of them is the better estimator for the volatility of the German DAX index during the chosen sample period.

The forecasting ability of the two models is compared by testing two hypotheses that are made based on previous studies presented in chapter two:

Hypothesis 1:

- H0: The implied volatility is able to outperform the GARCH(1,1) during both of the subperiods.
- H1: The implied volatility is unable to outperform the GARCH(1,1) during either of the subperiods.

Hypothesis 2:

- H0: The forecasting performance of the models increases after a time of financial turbulence, such as the Covid-19 pandemic.
- H1: The forecasting performance is not affected by the market turbulence caused by the Covid-19.

1.3 Structure of the study

This study begins by presenting the theoretical framework around the subject and by introducing the topic and goals of this study. In chapter two the implied volatility and the GARCH model are introduced and then explored more in-depth to gain a further understanding about how the models work and why they were chosen for this study. At the end of chapter two, relevant studies about similar topics to this study are presented. This section focuses on the findings of these studies as they can give an insight into what possible outcomes this study may have. Thus the hypotheses of this study are made based on these previous studies about volatility forecasting and the two chosen models.

Chapter three focuses on the empirical methods used in this study. The chapter goes into the issues with estimating realized volatility, the issues with the models, and how the forecasts were made. The chapter also addresses the measures that are used to compare the two forecasting models. Chapter four then presents the findings of the study as well as discusses descriptive statistics for the DAX index and the volatility forecasts. The chapter also goes into comparing the volatility models. Finally, in chapter five the conclusions and further thoughts about the subject are presented. The chapter also addresses ideas for further research and ideas to gain more informative results to find out more about the relation of the two models.

2 Theoretical background

This study approaches volatility forecasting by choosing one model representing both of the lines found in academic research regarding volatility forecasting. (Muzzioli, 2010) The first model used in this paper is the Black-Scholes implied volatility, referred to as implied volatility. It is a byproduct of the popular option pricing model for European-style options by Black & Scholes (1973). The model gives European-style options a price based on five different variables: the price of the option, the price of the underlying asset, the strike price, the time to maturity, the risk-free interest rate, and the future value of volatility. As all before mentioned information besides the future volatility can be observed from the market, it is possible to find the volatility that the other information indicates based on the pricing formula.

Implied volatility is viewed as a market-based volatility forecast. Therefore, according to Poon (2005), it is widely considered as the superior way of forecasting volatility outperforming or matching the performance of forecasts generated with time series models. It makes use of more informative data than its financial time series counterparts.

On the other hand, it is known that the implied volatility models require several assumptions to hold for the option theory to produce useful volatility estimations. The models also suffer from market price irregularities. Even with this in mind, the forecasts with implied volatility have been shown to make use of more informative data than the financial time series versions as the implied volatility represents the knowledge and understanding that the investors possess about the market and everything that may influence the option prices (Poon, 2005).

The second model in this study is the GARCH model. The GARCH model was created by Bollerslev (1986) and in this study it represents the line of models that forecast volatility based on financial time series data. The GARCH was chosen as a competing forecast model for the implied volatility because Ederington & Guan (2005) found the GARCH model to be a better-performing model than more complex forecasting models based on financial time series data.

2.1 The Black-Scholes option pricing model

Today The Black-Scholes option pricing model holds a position as a critical part of the option pricing theory. It has become widely used both academically and within the market (Figlewski, 2008) and has become generally viewed as one of the most used and successful models in finance (Rubinstein, 1994).

While the pricing model is related closely to the prior models featured in literature, the model by Black & Scholes (1973) distinguishes itself by calculating the price of an option in a manner that ignores the investor's risk preferences and rather calculates the price of an option based on information found from the market.

It is also important to note that the option pricing model by Black & Scholes (1973) makes a number of restrictive assumptions about the market, and they see the market as having ideal conditions. The assumptions by Black & Scholes (1973), which are required for the model to stand are:

- Underlying asset's returns follow a random walk, and its volatility is constant.
- Frictionless market conditions, there are no transaction costs, trading is continuous, and there is a possibility to borrow money at a risk-free rate.
- The options are of European style, meaning that they can be exercised only at maturity.
- There are no dividend payments before the maturity of the options.

Black & Scholes (1973) proceed to demonstrate that following the assumptions, the price of the option in theory is solely reliant on the underlying asset's initial price and known constant variables.

Black & Scholes (1973) derived the option price from the no-arbitrage argument: the hedged portfolio created from a short position on the option and a long position on the underlying asset must return the risk-free rate. Under these assumptions, the pricing equation by Black & Scholes (1973) for European-style call options takes the following form:

$$C(S,t) = \phi(d_1)S - \phi(d_2)Ke^{-r(T-t)},$$

$$d_1 = \frac{\ln\left(\frac{S}{K}\right) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}},$$

$$d_2 = \frac{\ln\left(\frac{S}{K}\right) + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}.$$

where:

$\phi(d_n)$ is the cumulative standard normal distribution for variable d_n,

S is the underlying asset's spot price,

K is the option's strike price,

r is the risk-free rate,

σ is the underlying asset's volatility until the date of expiration, or implied volatility, and

$T - t$ is the time left until the option reaches maturity.

For the model to produce volatility forecasts that are accurate, it is necessary to satisfy the assumptions as well as possible. Due to these assumptions, the model is not considered to be optimal or perfect, but rather to be a useful and widely used approach to volatility forecasting in academia and in practice. The main problem with the use of this model is that it assumes volatility to be non-stochastic, meaning that it assumes volatility to be unchanged from the issuance of an option until the option reaches maturity (Canina & Figlewski, 1993).

2.2 The GARCH model

The modeling of stochastic volatility was originated by Robert Engle (1982). Bollerslev (1986) defines stochastic volatility as a volatility that changes over time as a function of past errors. Engle designed the ARCH model and then Bollerslev (1986) further developed the model into a generalized version: GARCH.

The GARCH model stands for *Generalized AutoRegressive Conditional Heteroskedasticity*. Heteroskedasticity means that variance and thus volatility vary over time, and are not constant. Conditionality means that the model makes the estimations of volatility depending on its past values (Bollerslev, 1986). Autoregressive means that the estimation is based on its previous value and a stochastic term, Pinsky & Karlin (2010) define stochasticity as randomness. The model being generalized means that the lag structure of the model is much more flexible than the one for the ARCH model. Lags mean the number of previous values that are taken into account when calculating the new volatility value (Bollerslev, 1986).

The GARCH(p, q) model can be formulated as:

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \sigma_{t-i}^2 + \sum_{j=1}^{q} \beta_j v_{t-j}^2$$

where:
σ_t is the value of volatility that the model forecasts for moment t,
ω is the white noise error term and it is assumed to be greater than zero,
α_i and β_i are coefficients determining weights for the variables σ_{t-i} and v_{t-j} $(\alpha, \beta \in [0,1])$,
v_{t-j} is the volatility of the time series at the moment $t-j$,

In this study the GARCH (1,1) is used, which implies: $p = q = 1$ (Bollerslev, 1986).

Ederington & Guan (2005) have criticized the GARCH model for giving too much weight on most recent volatility observations but nevertheless they found the model still gives a better estimate than other models. Even while keeping this in mind, Ederington & Guan (2005) note that the GARCH(1,1) model often tends to outperform more refined time series models. This is the reason for it being often used as a benchmark in comparing the volatility forecasting models. This is the main reason for choosing the GARCH(1,1) model as a comparison to the implied volatility in this paper.

2.3 Comparisons in previous literature

Studies from the 1990s comparing the forecasting performance of implied volatility and financial time series methods have come up with inconclusive and contradictive findings. Canina & Figlewski (1993) found the implied volatility to be a poor forecast for future volatility of the S&P 100 index. The study was made using data from the S&P 100 index options, which at the time were the most liquid options in the United States. Their study found the correlation between implied volatility calculated with the Black-

Scholes option pricing model and the realized volatility to be next to nonexistent. They also found that a simple historical volatility measure outperformed implied volatility in predicting future volatility, but the historical volatility measure was not able to accurately forecast future volatility either.

Day & Lewis (1992) compared the forecasting performance of estimations for future volatility made with implied volatility, GARCH and EGARCH (exponential GARCH). The forecasts were calculated for the S&P 100 index and call options on the S&P 100. They found that implied volatility forecasts may contain more information than the forecasts by GARCH and EGARCH models. As the result of their study, Day and Lewis refrained from proclaiming a ranking for the models and found that none of the models is completely able to characterize the market volatility.

Lamoureux & Lastrapes (1993) encountered ambiguous results when comparing the implied volatility to the GARCH model's forecast. The study was done by forecasting volatility for stock options of some of the world's largest companies at the time of the study. They refrained from making definitive conclusions about the predictive ability of implied volatility compared to forecasts generated with GARCH models.

Despite the mixed results from research during the 1990s, in further studies, which have sought to address methodological shortcomings of earlier works, it is indicated that forecasts made with implied volatility tend to outperform time series forecasts.

Christensen & Prabhala (1998) studied the implied volatility of call options for the S&P 100 index. They found that implied volatility outperforms forecasts made based on past volatility, one of the used models to forecast volatility based on historical data was GARCH(1,1). Their analysis found an interesting phenomenon, as the forecasting performance of implied volatility improved significantly after the stock market crash of year 1987.

Similar results to support the superiority of implied volatility's forecasting performance were yielded in a study by Harikumar et al. (2004). They examined the performance of the implied volatility and different types of GARCH models one of which was the GARCH(1,1). The analysis was made by examining currency call options data from, the Philadelphia Stock Exchange for British Pounds, Swiss Francs, and Japanese Yen. Their study found the implied volatility to outperform GARCH models.

Poon & Granger (2005) note that implied volatility is widely documented to produce biased forecasts for future volatility. The model tends to under forecast lower levels of volatility and it also tends to over forecast higher levels of volatility. On average volatility forecasts made with implied volatility tend to yield volatility forecasts with a higher level of volatility than the realized volatility. Even with this in mind, Poon & Granger (2005) still state that implied volatility is a better way to forecast volatility than models based on historical data.

3 Empirical methodology

This chapter discusses the issues related to empirical methodology. The chapter covers the estimation of the realized volatility and the evaluation criteria for the forecasts. The made decisions considering the data and the methods are significant, as they may have considerable effects on the results of this study.

3.1 Estimation of the realized volatility

As already established in this study, the volatility of a financial asset can be calculated as the sample standard deviation of the asset's returns:

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

where:
r_t is the daily return on day t and
μ is the average return over the T-day period (Poon, 2005).

The accurate estimation of realized volatility can still be challenging as the estimates are subject to noise. As the realized volatility in this study is a reference value to which the volatility forecasts are compared, it is necessary to calculate it as accurately as possible to obtain results that are informative and accurate (Poon, 2005).

The issue with approaching the realized volatility as the standard deviation of the returns is that the sample mean may not provide enough information about the true mean of the population unless the sample period is extensive enough. Intuitively one could think that taking a longer historical sample would correct this problem. While including older data leads to a larger sample size of the population, this might not make the information about the state of volatility more accurate. As volatility is calculated by comparing the value of the daily return to the mean return, the possible misinformation the mean value holds has a big impact on the estimation of the realized volatility. Enhancing the accuracy of volatility estimation often involves an alternative estimation of the theoretical mean and utilizing it in the place of the sample mean. It is a common practice to enforce a theoretical zero mean and use the square root of average squared returns as the estimate for volatility (Poon, 2005; Figlewski, 1997).

Thus, the formula to calculate realized volatility in this study takes the following form,

$$\hat{\sigma} = \sqrt{\frac{\sum_{t=1}^{T} r_t^2}{T}}$$

where:
r_t is the return at time t, and
T is the length of the sample period (Poon, 2005).

Since volatility does not remain constant over time, it is necessary to take this into account when calculating the values of weekly realized volatility. As the volatility of an asset varies over time, it is necessary to use data with intervals that are as short as possible. In this study the realized volatility of the DAX index is calculated based on the daily value of the DAX index and then the daily volatility values are converted into weekly values by multiplying with the square root of 5 as there are 5 trading days in a regular trading week (Poon, 2005).

3.2 Forecasting with Implied volatility

The volatility forecasting with implied volatility is made following Jiang & Tian (2005). In this study, this forecast is made by simulating a scenario where an investor forecasts volatility for the following week by extracting the implied volatility of an option, using the option pricing model by Black & Scholes (1973).

The source of the options data used in this study is LSEG Workspace from where monthly data for DAX index options (ODAX) with maturities from the 19th of January 2019 until the 24th of January 2024 were gathered to forecast volatility for the chosen period 2019–2023. Options for the DAX index are traded in the EUREX derivatives exchange.

To determine the options for the calculation of implied volatility, this study follows Jiang & Tian (2005), where options with seven days or less until maturity are excluded from the sample, as these options are more vulnerable to market microstructure noise than options with more time until maturity. Besides the option with less than a week until expiration, the options that are nearest to maturity are used to calculate the implied volatility. The options chosen for the calculation will be near-the-money call options, this specification is also made following Jiang & Tian (2005) as near-the-money options are more liquid, and thus the prices more informative than with options that are not near-the-money. Specifically, the nearest-to-money options have been selected for this study.

The usage of one-month Euribor rate as a proxy for the risk-free rate was made following Muzzioli (2010). As the LSEG Workspace at TSE FinanceLab does not have a license for Euribor data, the data was collected from Suomen Pankki (https://www.suomenpankki.fi/fi/Tilastot/korot/kuviot/korot_kuviot/euriborkorot_pv_chrt_fi/), as they provide daily data for the annual one-month Euribor rate on their website.

The price of the underlying asset for a DAX index option is the value of the DAX index itself. The daily values of the DAX index from January 1, 2019, to December 31, 2023, were gathered from LSEG Workspace.

As the Black & Scholes (1973) gives the implied volatility values as annual volatilities, they must be converted into weekly values, to make comparisons with the realized volatility and the volatilities forecasted with the GARCH(1,1). This is done by following the same method as earlier in this chapter dividing the annual implied volatility value by the square root of 52 as there are 52 weeks in a year.

3.3 Forecasting with GARCH(1,1)

Out-of-sample testing is widely considered to be the “gold standard” of volatility forecasting with financial time series models. Conducting out-of-sample forecasts gives much more information about the model’s predictive ability than estimating the model’s parameters with historical data to see if a model could fit the realized volatility data. As in practice the forecasts are made before there is any realized data from the period, in this study the forecasts will be made as they would have been done when forecasting the future (Sahiner, 2022).

Out-of-sample forecasts for volatility with the GARCH model can be made by using two different methods which are a recursive forecast and a rolling window forecast. In this study, the out-of-sample forecasts with GARCH(1,1) are generated with a recursive method which means that every GARCH(1,1) forecast is based on a sample period where the model is fitted to the sample data and then the model makes an out-of-sample forecast for the decided forecasting period (Sahiner, 2022).

In this study the out-of-sample forecasts are generated for every financial quarter based on the realized volatility data from the previous five years. This simulates a situation where an investor wants to forecast the weekly volatility of the DAX index for next the financial quarter one day at a time. The sample used in this study for fitting the model to make a forecast is the logarithmic returns of the DAX index from the previous five years. The method was chosen following Huang (2011).

The forecasts with GARCH(1,1) in this study are made with EViews following the instructions and settings by Aljandali & Tatahi (2018).

3.4 Evaluation of the forecasts

To obtain more informative results about the performance of the two forecasting models, the examined period of 2019–2023 is divided into two subperiods. The first period is the time during the height of the Covid-19 pandemic 2019–2021 and the second is after the height of the pandemic 2022–2023. The division into subperiods is made, to obtain information about the forecasting performance of the models during different market conditions: a market that is highly volatile and a market after a time of high volatility.

The forecasting performance of the two models is evaluated with three different statistical measures, chosen following Ching & Siok (2013): mean squared error, root mean squared error and mean absolute percentage error. Ching & Siok (2013) note that when evaluating the performance of forecasting models in practice, it is rare that one model dominates the other with respect to all evaluation measures. This is commonly solved by comparing the average performance across all chosen evaluation measures. While none of the models alone are a perfect indicator for the performance of the forecasting model as a whole, with all of the three used at the same time, the average result of the comparison can give a good indication on the models’ ability to forecast volatility during the chosen period in comparison to one another (Ching & Siok, 2013).

The mean squared error or MSE is often applied in studies comparing forecasting performance. It has a tendency to more harshly penalize larger forecasting errors than other commonly used measures of error, thus it is considered to be the most appropriate measure to determine which models avoid making large errors. The MSE is written as,

$$MSE = \sum_{t}^{n} \frac{e_t^2}{n}$$

$$e_t = y_t - \hat{y}_t,$$

where:
y_t is the actual value at moment t, and
$\hat{y}_t$ is the value of the estimation at moment of time t (Ching & Siok, 2013).

The root mean squared error or RMSE is the measure that is favored the most among practitioners and academics. The RMSE is written as,

$$RMSE = \sqrt{\sum_{t}^{n} \frac{e_t^2}{n}}$$

where the variables are the same as for MSE. Just like the MSE the RMSE gives equal weight to all the errors, thus the errors that happen at different times during the forecast are weighted equally (Ching & Siok, 2013).

Mean absolute percentage error or MAPE is written as,

$$MAPE = \sum_{t=1}^{n} \frac{\left|\frac{e_t}{y_t}\right| \cdot 100}{n}$$

where, n is the number of data points in the sample and $\left|\frac{e_t}{y_t}\right| \cdot 100$ is the absolute percentage error of the forecasted values. The variables are the same as for MSE (Ching & Siok, 2013).

4 Empirical results

In this chapter, the empirical findings of the study are presented. At the beginning of this chapter, the performance of the DAX index between 2019 and 2023 is discussed as well as a graphical illustration of the DAX index volatility forecasts and realized volatility is presented. The second part of this chapter is reserved for the analysis of the forecasting performance of the two competing models.

4.1 Descriptive statistics

Figure 1 shows the weekly closing values of the DAX index between 2019 and 2023. The graph illustrates that the DAX index's value is considerably affected by the highly volatile market conditions during the height of the Covid-19 pandemic in the spring of 2020. During the chosen period the DAX index reached its highest value towards the end of the period in December of 2023 and the lowest values are found during the height of the pandemic in March and April of 2020. To conclude, the chosen sample period consists of times with high volatility, which could possibly influence the performance of the two forecasting models.

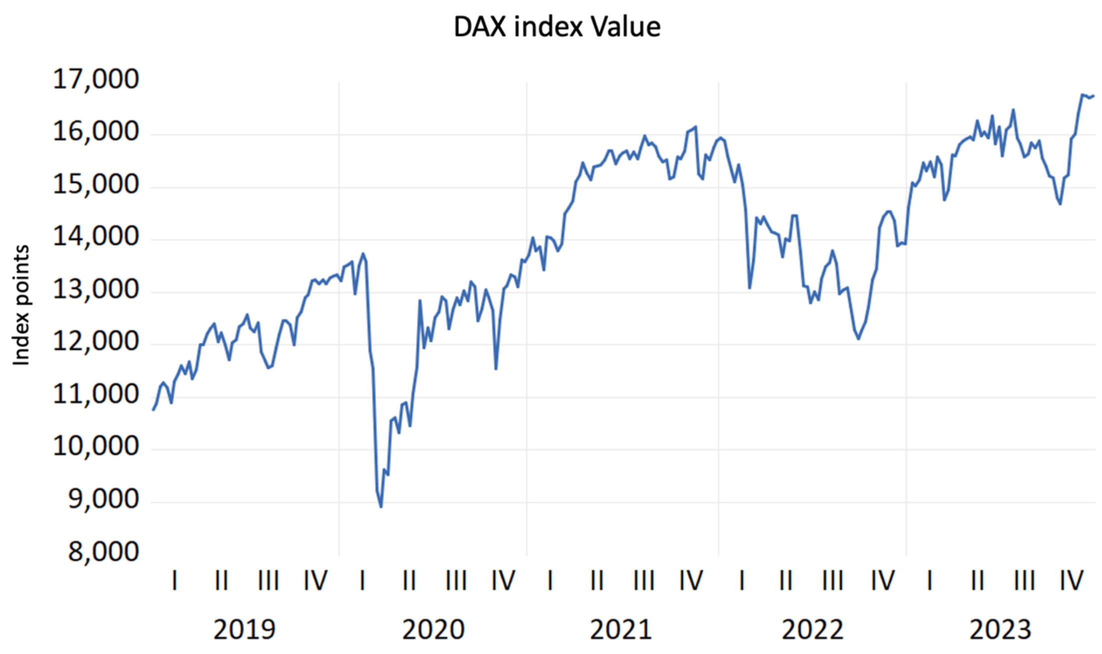

Figure 1 The value of the DAX index

Table 1 presents descriptive statistics for the DAX index ($n = 261$) and the weekly logarithmic returns during the sample period. The mean of the weekly returns is near zero but slightly positive as well as the median weekly return. This can also be seen in figure 1, where the DAX index's value is clearly higher at the end of the sample period than at the beginning of the period. The lowest weekly return of $-0{,}22330$ was observed on March 14, 2019, during the earlier stages of the Covid-19 pandemic, while the highest return value was observed four weeks later March 11, 2019.

Table 1 Descriptive statistics of the DAX index

	Value of the index	Weekly return
Mean	13843,22	0,002
Median	13786,29	0,003
Max	16759,22	0,104
Min	8928,5	–0,223
Standard deviation	1693,69	0,031

Table 2 provides descriptive statistics of the realized volatility, implied volatility, and the volatility forecasted with GARCH(1,1) for the whole sample period from 2019–2023. According to the means and medians, both forecasts are positively biased and seem to have forecasted volatility values that are larger than the realized volatility. While both forecasts seem to be positively biased, according to the mean and median in this study the implied volatility gives the investor even more positively biased results than the GARCH(1,1). The maximum value given by the GARCH(1,1) is very similar to the realized volatility, while both the maximum and minimum values of the implied volatility are significantly different from realized volatility values.

Table 2 Descriptive statistics of the volatility forecasts

	Realized Volatility	Implied volatility	GARCH(1,1)
Mean	0,019	0,029	0,027
Median	0,016	0,025	0,023
Min	0,002	0,0000001	0,013
Max	0,107	0,113	0,107
Standard deviation	0,013	0,017	0,013

4.2 Results of the forecast performance analysis

Figure 2 illustrates the results given by the two different volatility forecasting models and the realized volatility. In the graph, it seems that the implied volatility especially seems to over forecast higher levels of volatility. This is supported by findings in table 2 as the average value of the implied volatility is higher than that of realized volatility. The spike in volatility during 2020 is also easily identified from the graph and it seems that implied volatility forecasts high volatility slightly earlier than it actually realized.

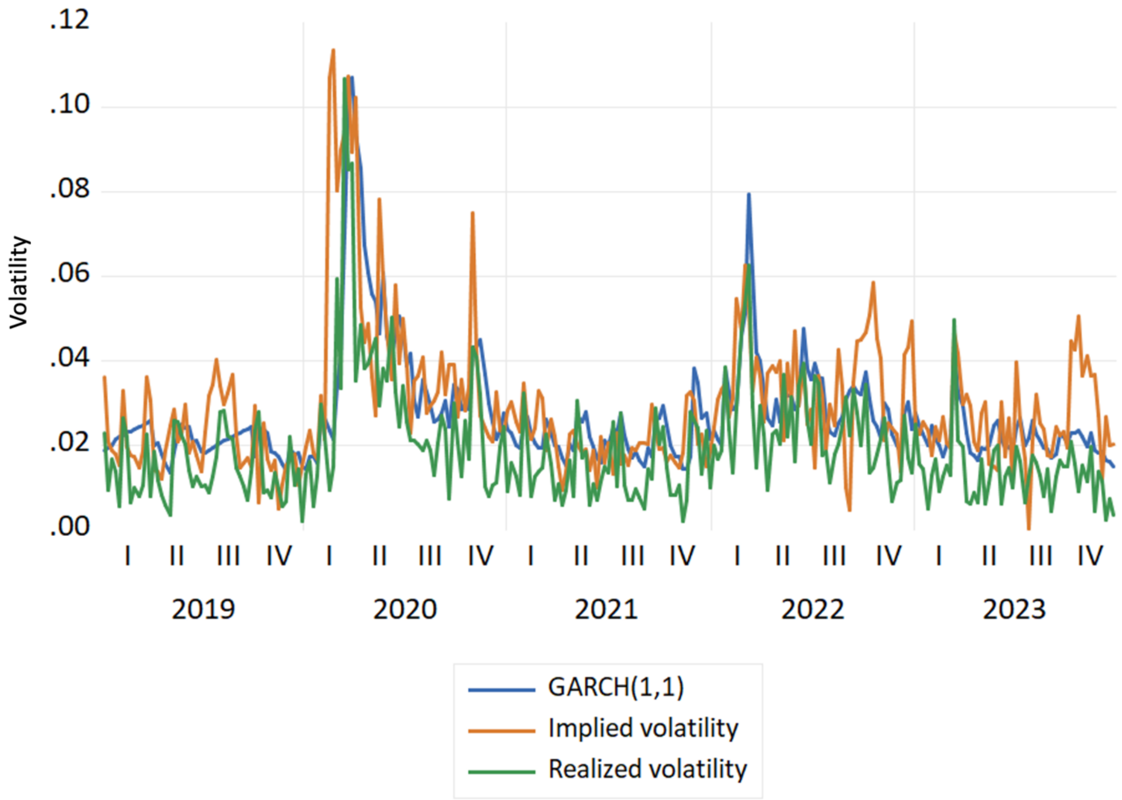

Figure 2 Forecasts and realized volatility

Table 3 provides the error statistics to compare the forecasts made with implied volatility and GARCH(1,1) for the first subperiod ($n = 157$). The forecast made with GARCH(1,1) outperforms the forecast made with implied volatility over all of the three error statistics. According to these three error statistics, during the first subperiod the GARCH(1,1) was able to perform superiorly in comparison to the implied volatility.

Table 3 Forecast Error Statistics for the first subperiod

Model:	MSE		RMSE		MAPE	
	Level	Rank	Level	Rank	Level	Rank
Implied Volatility	0,0003	2nd	0,018	2nd	101,162	2nd
GARCH(1,1)	0,0002	1st	0,013	1st	86,292	1st

In table 4 the error statistics for the second subperiod are provided ($n = 104$). The forecast made with GARCH(1,1) is able to outperform the implied volatility also during the second subperiod. According to the error measures chosen in this study, the used GARCH(1,1) method is a forecast in comparison to the used implied volatility method.

Table 4 Forecast Error Statistics for the second subperiod

Model:	MSE		RMSE		MAPE	
	Level	Rank	Level	Rank	Level	Rank
Implied Volatility	0,0003	2nd	0,017	2nd	123,767	2nd
GARCH(1,1)	0,0001	1st	0,010	1st	75,286	1st

In table 5 are the results for the comparison of the models to themselves during the two different subperiods. The values are represented as the second subperiod value subtracted by the second subperiod value. This means that a positive value implicates better performance during the first subperiod and vice versa. The GARCH(1,1) is able to outperform the first subperiod forecast in every error category during the second subperiod. Meanwhile the implied volatility forecast is able to outperform the first subperiod forecast in two of the three categories: MSE and RMSE but in the third category MAPE, the first subperiod forecast is able to outperform the forecast made for the second subperiod.

Table 5 Comparison of the model performance during both subperiods

Model:	MSE	RMSE	MAPE
Implied Volatility	–0,00002	–0,0005	9,007
GARCH(1,1)	–0,00002	–0,001	–3,548

5 Conclusions and Summary

The objective of this study was to provide insight into the comparison of volatility forecasts made with the GARCH(1,1) model and implied volatility. The first hypothesis presented in this study was that the implied volatility would be able to outperform the GARCH(1,1) during both of the subperiods. This hypothesis was made based on previous studies examining the differences between volatility forecasting models based on information from the derivatives markets and models based on financial time series data (see, for example Christensen & Prabhala, 1998; Harikumar et al., 2004; Poon & Granger 2005).

The results yielded in this study do not support the first hypotheses and thus supports rejecting the null hypothesis "The implied volatility is able to outperform the GARCH(1,1) during both of the subperiods" and accepting the alternative hypothesis "The implied volatility is unable to outperform the GARCH(1,1) during either of the subperiods". Thus resulting in the conclusion that in this study, measuring forecasting performance with the MSE, RMSE and MAPE the GARCH(1,1) by Bollerslev (1986) is actually the superior forecasting method over both of the two subperiods in comparison to the implied volatility by Black & Scholes (1973).

The second hypothesis made in this study is that the forecasting performance of the two models improves after a time of financial turbulence and a time of higher volatility such as the market conditions were between the two subperiods. This hypothesis was made based on the findings of Christensen & Prabhala (1998) that are presented in chapter two of this study.

For the GARCH(1,1) the results of this study support accepting the null hypothesis "The forecasting performance of the models increases after a time of financial turbulence, such as the Covid-19 pandemic". This is supported because all of the three chosen error metrics show that the forecast during the second subperiod has less error than the forecast during the first subperiod. For the forecast made with implied volatility, the results are not quite as clear as with the GARCH(1,1). The forecast with implied volatility for the second subperiod is able to outperform the first subperiod when measuring error with MSE and RMSE, but when measuring with MAPE, the results are the opposite as the first subperiod forecast has a lower MAPE value than the second subperiod forecast. In this, the choice comes down to the investor's preferences. If the investor prefers a model where the ratio of the forecasts that are made are close to each other, then they would be inclined to prefer the model that has the lower MAPE value. If the investor prefers the model to just have less error, then they would be inclined to prefer the model where the MSE and RMSE values are lower.

As previously mentioned in chapter three, this study determines the performance of the model as an average across all of the error metrics. As two of the three error metrics suggest so, the implied volatility makes a more accurate forecast during the second subperiod than during the first subperiod and the null hypothesis "The forecasting performance of the models increases after a time of financial turbulence, such as the Covid-19 pandemic".

In this study the implied volatility could not outperform the GARCH(1,1) even while that was expected in this study based on previous studies. This implies the fact that as the implied volatility is based on investors' expectations, the investors were unable to make accurate predictions and the turbulent market conditions caused the expectations to not be accurate. In this situation a financial time series model such as the GARCH(1,1) that makes predictions based on historical data shows that it can still be a useful tool to an investor for volatility forecasting.

The main problem behind making volatility forecasts with implied volatility was highlighted by Canina & Figlewski (1993) to be the fact that it assumes the volatility over an option's lifetime to be non-stochastic. This assumption is faulty as can be seen by calculating weekly volatility for the DAX index during the sample period. A faulty assumption for the model to hold implies that the results produced by the model may also be faulty. This can result in a situation such as this study found, where the volatility forecast made with implied volatility is not as accurate as a forecast made with another model such as GARCH(1,1).

Another possible reason for why the GARCH model was able to outperform implied volatility is due to the fact noted by Ederington & Guan (2005), that the GARCH(1,1) gives so much weight on recent volatility observations. This could possibly create a scenario where the unexpectedly functioning market can be accurately forecasting by closely following the previous few volatility values. Though this is a theory and it should be studied by comparing forecasts made by GARCH models with different lag structures to see how the weight put on different lag periods affects the forecasting ability of the models.

Even though this study found evidence that supports the GARCH(1,1) being a superior volatility forecasting model in comparison to implied volatility, it still is not appropriate to declare it as a superior model in comparison to the implied volatility in all situations and samples. The implied volatility has been shown extensively to outperform the GARCH(1,1) and thus it still remains a sufficient tool for volatility forecasting. (see for example Christensen & Prabhala, 1998; Harikumar et al., 2004; Poon & Granger, 2005)

Further research on this topic still seems necessary as studies have shown contradicting results. To gain more understanding about the forecasting ability of the two models in comparison to each other should the future studies include a longer sample period than this study. This is to ensure that the results of future studies can have even more reliable results that this study. It is also necessary to study different markets and not only the German market. A study to gain further insight into the topic should research most of the worlds largest markets. If a study with longer sample period and with more markets included yields similar results to this one, would it then be relevant to discuss whether the GARCH(1,1) has for some reason overtaken implied volatility in forecasting ability.

Instead of declaring clear superiority of one single model, it would be more informative for an investor to make volatility forecasts using multiple different models and comparing the forecasts to be able to understand why such forecasts were made and what assumptions the models make of the financial markets. This way an investor may gain more insight into the possible future volatility values. It is still important to keep in mind the fact that this study noted in chapter one: future volatility is uncertain and based on chance, meaning that perfectly predicting it is difficult if not impossible. An investor can only make forecasts and compare the forecasts to one another to see their compared forecasting performance.

To summarize, this study focuses on volatility forecasting with two competing models: the GARCH(1,1) by Bollerslev (1986) and implied volatility by Black & Scholes (1973). The two models represent the two main lines of volatility forecasting that are recognized in literature. The expectation based on previous literature was that the implied volatility should produce more accurate volatility forecasts than the GARCH(1,1) model and the research question was made based on that assumption as

“Why is the implied volatility a better forecasting method for future volatility than the GARCH model”. Surprisingly the study found results that oppose the made expectations as the GARCH(1,1) model was able to outperform implied volatility’s forecasting performance. Based on previous literature, the reasons for this could be that the investors’ expectations on the future volatility were inaccurate, the faulty assumption that volatility is non-stochastic affected the forecast enough for the GARCH(1,1) to outperform it or that the GARCH(1,1) was able to be reactive enough because of its lag structure to make more accurate forecasts in a quickly changing market to outperform implied volatility.

References

Aljandali, Abdulkader – Tatahi, Motasam (2018) Modelling Volatility in Finance and Economics: ARCH, GARCH and EGARCH Models. *Economic modelling with EViews,* 143–195.

Black, Fischer – Scholes, Myron (1973) The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, Vol. 81 (3), 637–654.

Bollerslev, Tim (1986) Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, Vol. 31 (3), 307–327.

Canina, Linda – Figlewski, Stephen (1993) The Informational Content of Implied Volatility. *The Review of Financial Studies*, Vol. 6 (3), 659–681.

Ching, Mun L. – Siok, Kun S. (2013) Comparing the performances of GARCH-type models in capturing the stock market volatility in Malaysia. *Procedia Economics* and Finance, Vol. 5, 478–487.

Christensen, Bent J. – Prabhala, Nagpurnanand (1998) The relation between implied and realized volatility. *Journal of Financial Economics*, Vol.50 (2), 125–150.

Day, Theodore E. – Lewis, Craig M. (1992) Stock Market Volatility and the Information Content of Stock Index Options. *Journal of Econometrics*, Vol. 52, 267–287.

Ederington, Louis H. – Guan, Wei (2005) Forecasting Volatility. *Journal of Futures Markets*, Vol. 25 (5), 465–490.

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

Figlewski, Stephen (1997) Forecasting Volatility. *Financial Markets, Institutions & Instruments*, Vol 6 (1), 1–88.

Figlewski, Stephen (2008) Estimating the Implied Risk Neutral Density for the U.S. Market Portfolio. *Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle, eds. T. Bollerslev – J.R. Russell – M. Watson*. Oxford University Press, Oxford, England.

Frennberg, Per – Hansson, Björn (1995) An Evaluation of Alternative Models for Predicting Stock Volatility: Evidence from a Small Stock Market. *Journal of International Financial Markets, Institutions and Money*, Vol. 5, 117–134.

Harikumar, Sankaran – De Boyrie, Maria E. – Pak, Simon J. (2004) Evaluation of Black-Scholes and GARCH models using currency call options data. *Review of quantitative finance and accounting*, Vol. 23 (4), 299–312.

Huang, Kun (2011) Modelling Volatility of S&P 500 Index Daily Returns: A comparison between model based forecasts and implied volatility. Hanken School of Economics.

Jiang, J.J. – Tian, Y.S. (2005) The Model-Free Implied Volatility and Its Information Content. *The Review of Financial Studies*, Vol. 18 (4), 1305–1342.

Lamoureux, Christopher G. – Lastrapes, William D. (1993) Forecasting Stock Return Variance: Toward an Understanding on Stochastic Implied Volatilities. *The Review of Financial Studies*, Vol. 6 (2), 293–326.

McNeil, Alexander J. – Frey, Rüdiger – Embrechts, Paul (2015) Quantitative risk management: concepts, techniques and tools. Revised edition. Princeton: Princeton University Press.

Mohammadi, Mohammad (2016) Prediction of α-stable GARCH and ARMA-GARCH-M models. *Journal of forecasting*, Vol. 36 (7), 859–866.

Muzzioli, Silvia (2010) Option-based forecasts of volatility: an empirical study in the DAX-index options market. *The European Journal of Finance*, Vol. 16 (6), 561–586.

Pinsky, Mark A. – Karlin Samuel (2010) An introduction to stochastic modeling. Academic Press.

Poon, Ser-Huang (2005) A practical guide to forecasting financial market volatility. John Wiley & Sons, Chichester, England.

Poon, Ser-Huang. & Granger, Clive (2005). Practical issues in forecasting volatility. *Financial Analysts Journal*, Vol. 61 (1), 4–556.

Rubinstein, Mark (1994) Implied Binomial Trees. *Journal of Finance*, Vol. 69 (3), 771–818.

Sahiner, Mehmet (2022) Forecasting Volatility in Asian Financial Markets: Evidence from Recursive and Rolling Window Methods. *SN Business & Economics*, Vol. 2, (10) 157–157.

Sorokina, Nonna Y. – Booth, David E. (2022) Does liquidity matter when crisis is brewing? Equity returns and intraday liquidity in DAX futures market. *Journal of Accounting and Finance*, Vol. 22 (1) 116–132.

Uddin, Moshfique – Chowdhury, Anup – Anderson, Keith – Chaudhuri, Kausik (2021) The effect of COVID – 19 pandemic on global stock market volatility: Can economic strength help to manage the uncertainty? *Journal of Business Research*, Vol. 128, 31–44.

Wallmeier, Martin – Hafner, Reinhold (2000) The Dynamics of DAX Implied Volatilities. *University of Augsburg Working Paper.*

Äijö, Janne (2008) Implied volatility term structure linkages between VDAX, VSMI and VSTOXX volatility indices. *Global Finance Journal*, Vol. 18 (3), 290–302.