Forecasting Volatility: A Comparative Study of GARCH and Implied Volatility Using the NOK/SEK Exchange Rate

Oskar Laestander

Kandidatuppsats i matematisk statistik
Bachelor Thesis in Mathematical Statistics

Kandidatuppsats 2023:17
Matematisk statistik
September 2023

www.math.su.se

Matematisk statistik
Matematiska institutionen
Stockholms universitet
106 91 Stockholm

Mathematical Statistics
Stockholm University
Bachelor Thesis **2023:17**
http://www.math.su.se

Forecasting Volatility: A Comparative Study of GARCH and Implied Volatility Using the NOK/SEK Exchange Rate

Oskar Laestander*

September 2023

Abstract

Volatility forecasting is an integral task in finance, with strong implications for option pricing, risk management, and investment strategies. Two commonly used methods for volatility forecasting include the Generalised Autoregressive Conditional Heteroscedasticity (GARCH) model, a traditional statistical approach, and Implied Volatility derived from option prices, which very much reflect the market's expectations of future volatility.

This thesis presents a comparison of the performance of the standard GARCH(1,1) model and Implied Volatility when forecasting the volatility of the NOK/SEK exchange rate, using Naive forecasts as a reference point. The study examines the Mean Squared Error (MSE) of the forecasts produced by both methods, and evaluates their accuracy when compared to the true realised volatility. The study relies on data from mid-2013 to mid-2023 obtained from Yahoo Finance and Eikon.

The results indicate that overall, the GARCH(1,1) model generates more accurate forecasts than Implied Volatility, as measured by the MSE. However, the superiority of the GARCH model is not universal, with Implied Volatility outperforming in certain years. Noteworthy, there is a strong correlation between the performance of GARCH and Implied Volatility, with both methods experiencing difficulties in years of high volatility. Additionally, the Naive forecasts offer relatively accurate forecasts, even performing the best of all methods in 2014. The findings support previous research suggesting the complex nature of volatility prediction, with no one-size-fits-all solution.

The study identifies research gaps, particularly in comparing GARCH and Implied Volatility in the context of exchange rates, and suggests promising avenues for future research. The findings of this thesis contribute to the theoretical understanding of volatility forecasting, and offer practical insights for market participants.

*Postal address: Mathematical Statistics, Stockholm University, SE-106 91, Sweden. E-mail: oskar.laestander@gmail.com. Supervisor: Ola Hössjer, Mohamed El Khalifi.

Acknowledgments

I would like to express my gratitude to my supervisors, Ola Gerton Hössjer and Mohamed El Khalifi, for their support throughout this project. Their expertise, constructive feedback, and patience have been instrumental in shaping this thesis. Their encouraging words and guidance have provided me with the necessary confidence to overcome the challenges that arose during this journey.

In addition, I would like to acknowledge the invaluable assistance of ChatGPT, developed by OpenAI. The around-the-clock availability and resourcefulness of this AI system have greatly influenced the structure of this thesis and provided several insightful suggestions for improvement. The assistance I received from ChatGPT in developing the code and ideas used in this thesis has significantly improved my efficiency and deepened my understanding of the concepts.

I wish to clarify that despite the instrumental assistance provided by ChatGPT, all the text in this thesis is entirely written by me. This contribution served as an aid for my understanding and my expression of the ideas and not as a source for the text.

Lastly, I would like to extend my appreciation to my family, friends, and colleagues who have provided encouragement and inspiration throughout this process. Your support has been invaluable and has made this journey all the more rewarding.

This journey of knowledge would not have been possible without your contributions. For that, I am grateful.

Contents

1 Introduction

1.1 Background and Rationale

Financial markets are regarded as unpredictable in many ways, characterised by their seemingly random behaviour. All while the prices of stocks, bonds, commodities, currencies, real estate and many other types of assets elude predicability, the volatility of the returns show potential for forecasting. This study will focus on just that; Comparing the accuracy when forecasting the volatility in returns, using two common methods: GARCH (Generalised Auto-Regressive Conditional Heteroskedasticity), often utilised by professionals in the field, and Implied Volatility, an interesting alternative as it represents the consensus of the market rather than a specific model.

1.2 Statement of the Problem

In this work we aim to investigate the following thesis statement: "This study aims to compare the accuracy of the Standard GARCH(1,1) model when compared to Implied Volatility when forecasting the volatility in the log returns of the NOK/SEK exchange rate, by estimating the Mean Squared Error of the forecasted volatilities when compared to the realised volatilities."

1.3 Purpose of the Study

The objective of this study is to evaluate the usefulness of the GARCH model, as well as the Implied Volatility. Both GARCH and Implied Volatility serve as good benchmarks when forecasting volatilities, as GARCH is so commonly used, and the failure of Implied Volatility would imply a large arbitrage in the market. By examining these two methods, the hope is to provide insights into the usefulness of their underlying model assumptions, as well as the rigidity of the market.

1.4 Significance of the Study

Understanding volatility forecasting is instrumental in finance. Accurate forecasts could not only provide more profitable opportunities for investors, but the ability to anticipate volatility can also contribute to economic stability and better risk management.

1.5 Scope of the Study

This research is based on almost 20 years of daily data for the NOK/SEK exchange rate, from two independent sources, dated from 2003-12-01 to 2023-06-05. Additionally, the historical data on Implied volatility ranges from 2013-01-01 to 2023-06-02. For the GARCH model a rolling window approach is utilised, with a constant window size of 1000 days. The abundance of data for the exchange rate allows GARCH volatility forecasts to be made to compare with the already available IV (Implied Volatility) data.

2 Literature Review

2.1 Overview of Volatility Forecasting Models

Volatility is often an indicator of risk for an asset, as it corresponds to uncertainty in financial markets [Capinski and Zastawniak, 2011]. Volatility is the measure of the fluctuations in price over a period of time, and is very often used to give the asset an indication of risk. The ability to forecast volatility is essential in many aspects: portfolio management, risk management, option pricing, hedging strategies and many other reasons. Commonly, the forecasting models for volatility can be categorised as follows:

2.1.1 Statistical (Parametric) Models

These are often time-series models, usually making assumptions on the generation of sample data. The model parameters are estimated based on historical data and often using maximum likelihood methods. Examples of statistical models include ARMA and GARCH models.

2.1.2 Stochastic Models

Stochastic models do not make any parametric assumptions about the volatility process. Instead the volatility is modeled as an unobservable process, stochastic and of unknown structure. Models like these allow for great flexibility, but they are often computationally intensive.

2.1.3 High-Frequency-Based Models

These are models based on high-frequency data. This could be intraday data, sometimes including the data of every trade occuring. Such data can capture the dynamics of the asset more precisely, but they are resource heavy as they might be limited by storage capacity or computational restrictions.

2.1.4 Implied Volatility Models

These models estimate future volatility from already observed prices of assets. The most commonly used method is using the Black-Scholes option pricing formula and inferring the future volatility from it as the volatility that gives the pricing identical to the observed price.

2.2 Call and Put Options

Options are securities tied to an underlying asset, with their value derived from the price of said asset. More specifically, the option gives the holder the right to buy or sell the underlying asset for a set price on or before a set date. The price is called the strike price and the last date for exercising the option is called the expiration date. The difference between the call and the put option is: the call option gives the holder the right to buy the underlying asset for the strike price, while the put option gives the holder the right to sell the underlying asset for the strike price. It is also important to note that while the option gives the holder the right to exercise the option, it does not come with any obligations [Capinski and Zastawniak, 2011].

There are some variations to these options. The most common are European and American options. The difference between these two are when the holder has the right to exercise the option. For the American option the holder has the right to exercise the option (buy or sell) at any date up until the expiration date. In the case of a European option, the holder only has the right to exercise the option on the expiration date itself.

Typically when buying a call option, the holder is betting on the price of the underlying asset to go up more than expected. Similarly a holder of a put options is betting on the price of the underlying asset to go down more than expected. The strike price would typically be close to the expected price of the underlying asset at the expiration date, but there are different ways to estimate the expected price, leading to some differences in the price of the options.

A common method to set a price of an option is using the Black-Scholes option pricing formula, which will be discussed in the next section. The formula leverages factors such as price of the underlying asset, time to expiration and volatility to estimate the fair price of the option.

2.3 The Black-Scholes Option Pricing Formula

This formula was derived to find fair prices for European call and put options. Black and Scholes published the formula in 1973 and later received the Nobel prize in Economic Sciences for their work [Black and Scholes, 1973]. The formula makes certain assumptions on the market: the financial market is efficient meaning there is no arbitrage, there are no costs for transactions, the risk-free interest rate is constant, the volatility σ of the underlying asset is constant and the log returns of the underlying asset are normally distributed.

The formula for pricing a call option can be written as:

$$C(S_t) = N(d_1)S_t - N(d_2)Ke^{-r(T-t)}$$

where

- $d_1 = (\log \frac{S_t}{K} + (r + \frac{\sigma^2}{2})(T-t))\frac{1}{\sigma\sqrt{T-t}}$
- $d_2 = d_1 - \sigma\sqrt{T-t}$
- $C(S_t)$ is the price of the call option at time t
- S_t is the price of the underlying asset at time t
- K is the strike price of the call option
- $N()$ is the cumulative standard normal distribution function
- r is the risk free return
- T is the time of expiration

The formula for the put option is similar, but with small differences taking into account the payoff from selling the underlying asset, rather than buying it [Capinski and Zastawniak, 2011].

Note that volatility σ is used in the formula, while not being known. The volatility is often estimated with historic volatility values and as the formula assumes a constant volatility this is straightforward. However, the constant volatility assumption is often too simplistic and therefore an obvious flaw of the model.

2.4 The ARMA Model

Auto Regressive Moving Average (ARMA) models are used to understand and explain values in a time series. It is possible to make predictions, but with varying results. The ARMA model uses the dependencies between the last observation and a certain number of lagged observations and residual errors to estimate and forecast the upcoming values in the time series.

The model includes the parameters p and q, deciding the number of lagged observations to take into account for the Auto Regressive and the Moving Average components respectively. The model can be written as

$$X_t = \varepsilon_t + \sum_{i=1}^{p} \varphi_i X_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i},$$

where

- X_t is the time series value at time t
- φ_i are the parameters of the autoregressive part of the model
- θ_i are the parameters of the moving average part of the model
- ε_t is the error term at time t, assumed to be a white noise process.

Models like these are useful and flexible, but they do come with assumptions. For example the ARMA model assumes the time series to be stationary, meaning time does not change the properties of the model. Often an ARIMA (Auto Regressive Integrated Moving Average) is used to account for changes over time [Box et al., 2016, Sundberg, 2022].

2.5 GARCH Models in Volatility Forecasting

Generalised Autoregressive Conditional Heteroscedasticity (GARCH) models were introduced in 1986 by Bollerslev as a way to forecast volatility in time series. Over time GARCH models have only increased in popularity and today they are widespread and adopted by many, not only in the financial sector, but they are also used when predicting weather and in other fields of research. The GARCH model [Bollerslev, 1986] is an intuitive extension of the ARCH model [Engle, 1982] as it incorporates more parameters by adding an Auto Regressive component to the model forecasting of the conditional variance of the noise of the time series. In the ARCH model only a Moving Average component is used to model the conditional variance of the noise [Box et al., 2016].

The conditional variance of the GARCH model can be defined as

$$\sigma_t^2 = \omega + \sum_{i=1}^{q'} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{p'} \beta_i \sigma_{t-i}^2$$

where

- $\sigma_t^2 = \mathrm{Var}(\epsilon_t | \epsilon_{t-1}, \epsilon_{t-2}, \ldots)$ is the conditional variance of the log return residual at time t
- ω is a constant
- p' is the number of lagged conditional variances included in the model
- q' is the number of lagged residuals included in the model
- α_i are ARCH parameters, representing the effect of past residuals
- β_i are GARCH parameters, representing the effect of past variances
- ϵ_{t-i} are past log return residuals with respect to a mean value.

The largest strength of GARCH models comes from their ability to capture volatility clustering, a characteristic present in most financial time series data. Periods of high volatility σ_t are often followed by periods of high volatility and vice versa. This kind of behaviour is advantageously modelled using GARCH models and it and provides valuable insights into market fluctuations. Bursts of volatility are quite common and by using GARCH models they become slightly more predictable.

In general GARCH models are flexible and they can account for both positive and negative shocks to the market. Even long term effects can be modelled, allowing for effects from volatility shocks to persist for long periods of time. There are different variations of GARCH such as Exponential GARCH (EGARCH) and Fractionally Integrated GARCH (FIGARCH) which allow for different types of long term memory within the model.

The GARCH model does not come without limitations. As the underlying data for the model is always historic, the GARCH model might often fail to capture one time events, or sudden changes in volatility in general. Additionally, the GARCH model assumes a stationary mean return and might fail for data with distinct trends or structural breaks.

In spite of these drawbacks, due to the effectiveness in modelling volatility, GARCH models are widely used for forecasting volatility in financial markets. This makes the standard GARCH model a prime candidate as a model for this study.

2.6 Implied Volatility in Volatility Forecasting

Implied Volatility is not a model, but a unique way of forecasting volatility, as it represents the entire market's expectations of the future, based on the prices of today. The Implied Volatility data in this study is derived from the prices of call and put options, and the Black-Scholes option pricing formula. Using this formula, the implied volatility can be seen as the consensus of the market's view of future volatility.

Using the Black-Scholes pricing formula and the actual prices of the options, the only unknown variable left in the formula is the volatility σ. The resulting nonlinear equation can easily be solved and it gives us the Implied Volatility we are looking for. The data available for this study already contains the implied volatility, calculated by the Black-Scholes formula.

Implied volatility is often used in finance, especially in risk management, option pricing and general financial strategies. In terms of forecasting volatility, Implied Volatility has repeatedly been able to provide useful and trustworthy information about the future [Siriopoulos and Fassas, 2019].

The largest advantage of looking at Implied Volatility is that it elegantly incorporates all information available, even agreed upon prices for future trades, such as the call and put options. Unlike models based on historic returns, Implied Volatility is solely based on the expectations of the market.

It is worth noting that Implied Volatility has its limitations too. It can be influenced by human behaviour, supply and demand imbalances in the options market and changes of the market's risk aversion. Previous studies have found that Implied Volatility performs especially well in periods of financial stress, where the market expects volatility to rise.

As Implied Volatility is such a contrast to traditional methods of forecasting volatility based on historical data, it is of interest to compare the two approaches.

2.7 Comparison of GARCH and Implied Volatility Models

Comparisons between GARCH models and Implied Volatility has been done before, as the different approaches to volatility forecasting come with different assumptions and advantages.

GARCH models are rooted in historical time series data to create forecasts. The flexibility of GARCH models allows them to capture delicate phenomena such as volatility clustering and leverage effects. However, due to the nature of the assumptions, GARCH models will always to some extent forecast the future in such a way that it resembles the past. Due to this fact, dramatic effects such as a financial crisis or structural changes in the market or economy might be ignored or missed. Events like these might cause volatility patterns to change dramatically, paralysing GARCH models until they are refit with new data.

Implied Volatility on the other hand does not really take any time series data of the asset of interest into account, but only the prices of call and put options. It forms around the consensus of the market, and in that way, indirectly takes all information available to the market into account. However Implied Volatility is not without drawbacks, as it simply relies on our methods of pricing options. Other factors affecting the accuracy of Implied Volatility can be risk aversiveness in the market, supply and demand of options and transaction costs. This might cause the Implied Volatilities to differ from the true values of expected future volatilities.

Previous studies comparing GARCH and IV methods in terms of volatility forecasting have yielded mixed results. Some say IV outperforms GARCH, while some say it is the other way around. The relative performance of the two methods often depends on the underlying dataset of the study, and the forecasting horizon. The split opinions on the quality of the GARCH and IV volatility forecast methods enourage us to look more closely into previous research.

2.8 Research Gaps

While plenty of studies comparing GARCH and IV exist [Bunjaku and Näsholm, 2010, Schmidt, 2021], there are huge insights to be found in exploring the forecasting abilities of volatility forecasting models. Many studies have applied the models to well known, high liquid markets such as the United States or other major economies, often using the most traded market indices on the planet such as the S&P500 [Dai et al., 2020]. Fewer studies have examined how well these models perform on smaller markets such as the Nordic markets.

Although additional research can often be found on stock markets or major currencies [Bollerslev, 1986, Engle, 1982], less attention has been given to other asset classes such as commodities or less traded pairs of currencies. In this study we hope to fill this gap by using the NOK/SEK currency exchange rate as the underlying data set. This data is not particularly interesting on an international level, but the knowledge of how GARCH and IV perform on the NOK/SEK exchange rate can give clues as to how GARCH and IV would perform on other smaller, regional markets.

While much research has been done on forecasting volatility and measuring the accuracy of the values of these forecasts [Poon and Granger, 2003], less research has been done on how to predict directions of volatility change. A better understanding of the abilities of forecasting methods to predict direction of changes could be valuable to practitioners and researchers in the field.

Another topic that has not been the subject of much research is the impact of the difference in window size, forecast horizon and other time aspects of the forecasting. In this study we employ a static forecast horizon for the GARCH model but with a rolling window approach when fitting models parameters to data.

This thesis aims to contribute to the literature by gently filling in some of the abovementioned gaps, hopefully contributing to a more nuanced understanding of the strengths and weaknesses of volatility forecasting models.

3 Methodology

3.1 Data Collection

The data for this study was collected from two independent sources: Yahoo Finance and Eikon. The data from Yahoo Finance was primarily used for the NOK/SEK exchange rates, even though the Eikon dataset included both the exchange rates and the Implied Volatilities [Yahoo-Finance and Eikon, 2023].

The Yahoo Finance dataset includes daily observations of open, high, low, close and adjusted close prices for the exchange rate. This data spans from December 1st of 2003 to June 5th of 2023, and it is based on a high liquid market, meaning that many trades are observed.

The Eikon dataset includes only the daily exchange rates and the Implied Volatilities. The data is from January 1st of 2013 to June 2nd of 2023. This dataset is based on a lower liquid market which can be seen in terms of some dates with missing data.

As a result of the quality of data, the Yahoo Finance dataset is primarily used as the source for exchange rate data. In the rate instances of missing data, the dataset is complemented by the Eikon dataset. Furthermore the datasets have been visually compared and verified against each other, ruling out faulty observations, lending credibility to the data used in this study.

No pre-processing was needed for the data, except for the normalisation of the volatility data. As the Implied Volatility was already annualised, the same was done to the true observed volatilities and the forecasted volatilities from the GARCH forecasts. The data was analysed using RStudio and coded using the R programming language [RPr, 2023].

The true volatility values were calculated using the historical data available. For each existing day in the dataset, the true volatility was determined as the standard deviation of the log returns over the subsequent 21 days.

In short, the data collection process was straightforward and faced minor obstacles. Still a lot of work has been done ensuring the appropriate forecasting methods and normalisations of volatilities have been used.

While the data from Eikon is not publically available, all data from Yahoo Finance is easily accessible to anyone. On the other hand, historical Implied Volatility data is only available in the Eikon database. The exchange rate data obtained can be seen in Figure 1, comparing the data obtained from Yahoo Finance and Eikon.

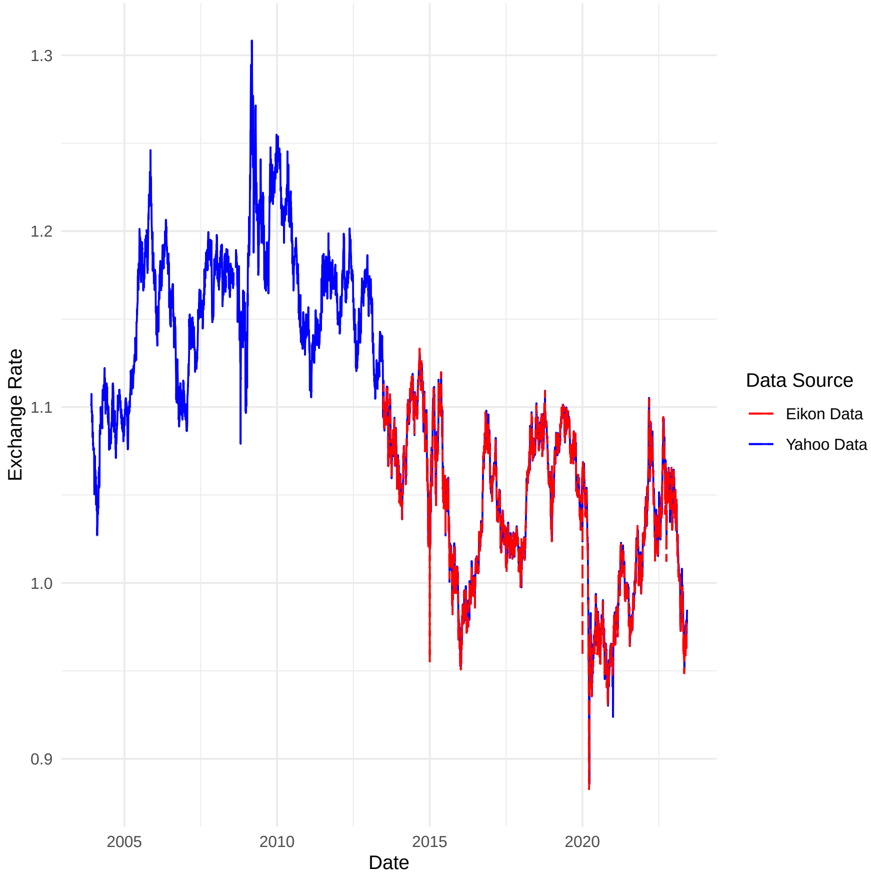

Figure 1: Exchange Rate Data: Eikon vs. Yahoo.

3.2 Description of GARCH Models

The GARCH model is characterised by its parameters p' and q'. The most commonly used version of the GARCH(p', q') model is the GARCH(1,1), which we will use in this study as well. This choice was guided by common practice within the financial sector, making this an easy decision. In the formula for the conditional variance of the noise, the GARCH(1,1) model includes one lag of the conditional variance and one lag of the squared residual. It also includes a mean model for the log exchange rate, in this case specified as an ARMA(1,1) model, allowing us to capture the Auto Regressive and the Moving Average aspects of the log returns.

The model was fit using the 'rugarch' package in R, more specifically the 'ugarchfit' function [Galanos, 2022]. The specification of the function were set to a standard GARCH model (sGARCH) with (1,1) as the order of the model. The parameter estimation method was set to 'hybrid' employing a combination of a line search algorithm and a trust-region method, allowing for some efficiency in the parameter fitting while maximum likelihood estimation remains the core method used.

The data used for the modelling was the log return data extracted from the NOK/SEK daily exchange rates, with a 1000 day window size rolling window approach. This means the model was fitted every day, using the last 1000 days of data. Using the model a 21 day volatility forecast was made, for every observation in the dataset.

The goodness of fit was verified through diagnostic checks of the residuals. In particular, the residuals were evaluated using QQ-plots, in this case verifying that the residuals indeed come from a student's t-distribution, as specified in the model. A smaller sample of the fitted models were analysed in a QQ-plot where only some of the models showed indication of having a misspecified distribution of the residuals, as can be seen in Figure 2. This step is crucial as it ensures the GARCH model to be fitted optimally. Initially the standard normal distribution was used, but it had to be replaced by the student's t-distribution in order to ensure a good fit of the model.

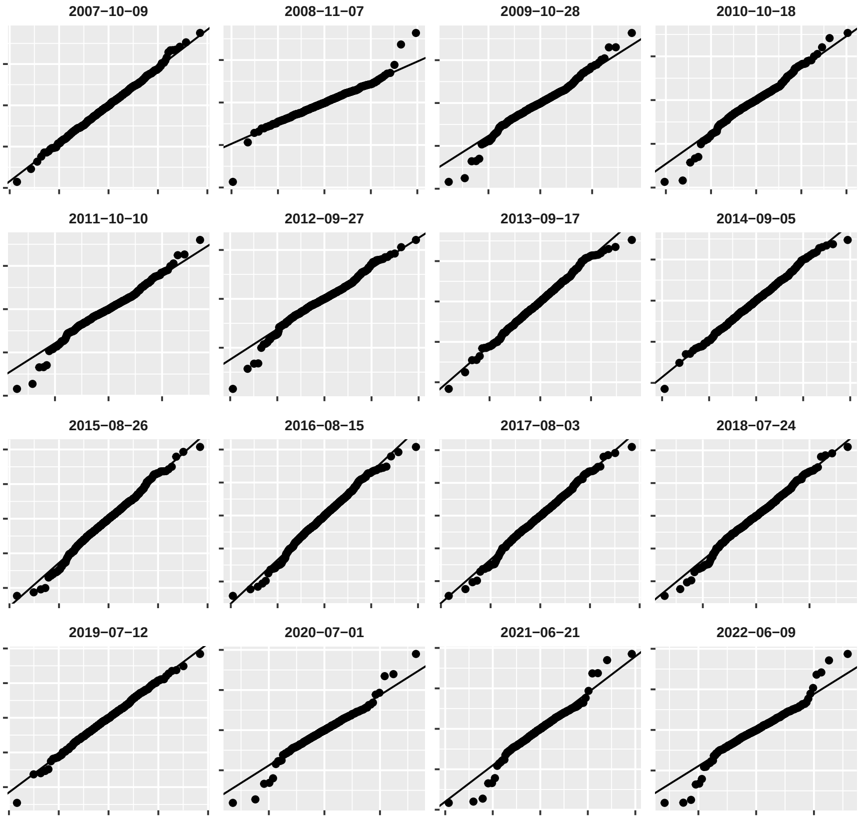

Figure 2: Periodically selected QQ-Plots compared to the theoretical Student's t distribution.

3.3 Description of Implied Volatility

Implied Volatility is a measure derived only from the pricing of options, as well as a pricing formula, provided in the form of Black-Scholes option pricing formula. The Implied Volatility was directly provided from the Eikon dataset, meaning no calculations were needed. The data was calculated based on the prices of the call and put options available and the transactions made on the option market for the NOK/SEK currency exchange.

The options expire once every month, consisting of typically 21 days of trading. As a result, the forecast horizon is changing day by day, shrinking up until the last day of the trading month. The first volatility forecast based on the Implied Volatility is consequently a 21 day forecast, and day by day the forecast horizon shrinks until a 1 day forecast is obtained at the last day of the trading month.

This is in contrast to the constant forecast horizon of 21 days from the GARCH model. A more complicated approach could have been used, but it is unlikely that his would have given very different results at the end, as the GARCH model provided stable forecasts for the entire 21 day horizon.

3.4 Data Normalisation and Mean Squared Error Calculation with Confidence Intervals

The Implied Volatility data from Eikon was already annualised when recieved. Since this is also common practice, such a normalisation was used for all volatilities in this study. This means that the true observed volatilities, and the volatilities σ_t forecasted by using the GARCH model were converted to annual data. This is usually done by multiplying the daily volatility with $\sqrt{252}$ as this is the typical number of trading days in a year. The volatilites are also referred to in percentages, meaning they have been multiplied with 100.

The measure of accuracy in this study is simply the Mean Squared Error (MSE). The MSE emphasises larger errors over smaller ones due to the square, and it is a commonly used metric for assessing the average size of errors. The Standard Error (SE) was calculated for each forecasting method using the 'dplyr' package, part of the 'tidyverse' package for R. The MSE was then calculated for the entire dataset, and also for the data grouped by year.

To address the uncertainty within the samples, bootstrapped confidence intervals were created for the MSE values. Bootstrapping is a computer-intensive, brute force method to assess the accuracy of the estimates [Chernick, 2008]. In this study, 1000 bootstrap samples are generated randomly from the entire, underlying dataset. The quantiles of the 1000 samples are then used to estimate confidence intervals, using the 2.5th and 97.5th percentiles of the resulting distribution of bootstrapped MSE values. The bootstrap method allows for the capture of uncertainty of the MSE estimates, giving further insights into the results of the thesis.

3.5 Naive Volatility Forecast

To have another, simple comparison and benchmark, naive volatility forecasts were also calculated. The naive forecasts simply consist of the average volatility of the last 21 days, forecasting that same volatility for the next 21 day period. This makes them simple to analyse and calculate. Because of the nature of naive estimates, they are simply the true volatilities delayed, as seen in Figure 3.

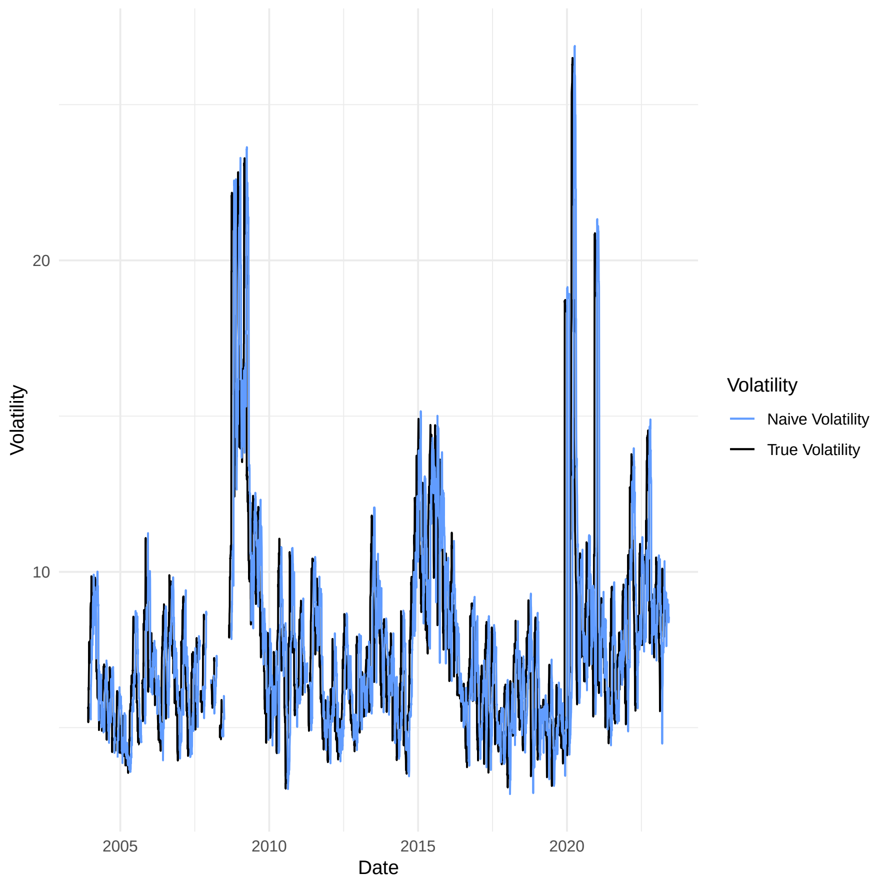

Figure 3: True vs. naive volatility during 2003-2023.

3.6 Methodological Justification

The GARCH model was chosen as a representative of classic, parametric forecasting models and also because of its popularity within the financial field. It is a flexible model, well suited to the task.

Implied Volatility was chosen as a metric of what prices the market is actually willing to pay for options, which prices should be directly derived from the volatility. The Implied Volatility should reflect the consensus of the market of what volatility is expected. Despite its simplicity, the Implied Volatility method has shown to provide accurate forecasts, making it a worthwhile benchmark against more prestigious methods such as GARCH.

The Mean Square Error was chosen as the evaluation metric for this study. MSE quantifies the average, squared difference between forecasts and observed values, pronouncing larger errors over smaller ones. Other metrics such as Mean Absolute Error or Mean Absolute Percentage Error could have been used. But MSE was recommended due to its wide acceptance in financial mathematics, as well as its property of penalising large errors harschly.

The choice of the NOK/SEK exchange was partly made due to a limited amount of available data. Another reason for choosing NOK/SEK was because of the regional relevance. In addition, since the Norwegian and Swedish economies share many similarities, the exchange rate between the currencies is an interesting subject of study.

For the forecasting horizons some decisions had to be made. As described in section 3.3. the forecast horizons are changing in size. For every trading month they start at 21 days, and then they shrink one day at a time. This is nothing we can change as this is how the Implied Volatility estimates reflect the time to expiration of the options. To maintain consistency with the longest forecasts of the Implied Volatility, a forecast horizon of 21 days was chosen for the GARCH forecasts. Obviously there is a difference here, but using a constant horizon for the GARCH model allows for more consistent forecasts, and it also allows for easier comparisons with future studies on the subject.

When evaluating the performance of GARCH, Implied Volatility and the Naive forecasts, the MSE was examined both for the entire period and for each year. This allows for a more comprehensive analysis of the difference in performance, highlighting variations in performance at different periods in time.

Moreover, the bootstrapped confidence intervals were added to the methodology, to address the randomness of sampling, as well as possibly provide further insights into the differences of method accuracies. As no underlying distribution could be found for the residuals for any of the model forecasts, analytical confidence intervals were impossible. The bootstrap technique however, was chosen due to the lack of an underlying distribution, but also because it is a robust method for computing confidence intervals while making few assumptions about the distribution of the underlying data.

In general, the choices made for this study were guided by the will to provide a meaningful comparison between the GARCH and IV models, while allowing for future comparisons.

4 Results

4.1 Presentation of Findings

The main findings from the data analysis will be discussed in this section. The focus is to compare the performance in forecasting accuracy of the GARCH and IV methods, while also comparing with Naive forecasts as a point of reference.

From the entire dataset for which IV data exists, the MSE was calculated annually and for the whole time period. In total, the GARCH model had the lowest MSE of 9.23 (CI: 8.04 - 10.57), followed by the MSE of the IV forecasts with a MSE of 10.77 (CI: 9.38 - 12.20). Lastly the benchmark metric of the Naive forecasts scored an MSE of 13.60 (CI: 12.00 - 15.25). This would suggest that, in general, the GARCH model provided the most accurate forecasts of volatility over the analysed period.

In addition to the total MSE, annual examinations were made. This was done to identify any potential trends and limitations to the different models. Apart from the last year 2023 the IV method had the single best

MSE of 2.16 in the year 2016. However there is no apparent trend as to which method of forecasting is the best. Every method, even the Naive forecasts, performed the best for some year. However, it is interesting to note that the GARCH method never performed the worst of the methods, something that cannot be said about the IV and Naive forecasts. The worst MSE for one single year was obtained for the Naive forecasts in 2020 with an MSE of 67.27. The year 2020 was a bad year for all methods, with GARCH performing the best among the three methods, but this year still gave the worst MSE value of 45.56, for GARCH, over the period.

The year 2020 stands out as clearly the most challenging year for all methods, and this is likely due to the outbreak of Covid-19 during that particular year. However, these methods are expected to handle even such events, but as can be seen from Figure 4, there were some issues for all methods when forecasting the volatility in 2020. Comparing the MSE we can see some correlations between a higher MSE and a higher volatility.

Overall, GARCH seems best suited for predicting volatility, providing both the best overall MSE, but also never performing the worst and never being far off the best MSE. All methods faced challenges in different period, most notable in 2020, really showcasing the difficulties in volatility forecasting.

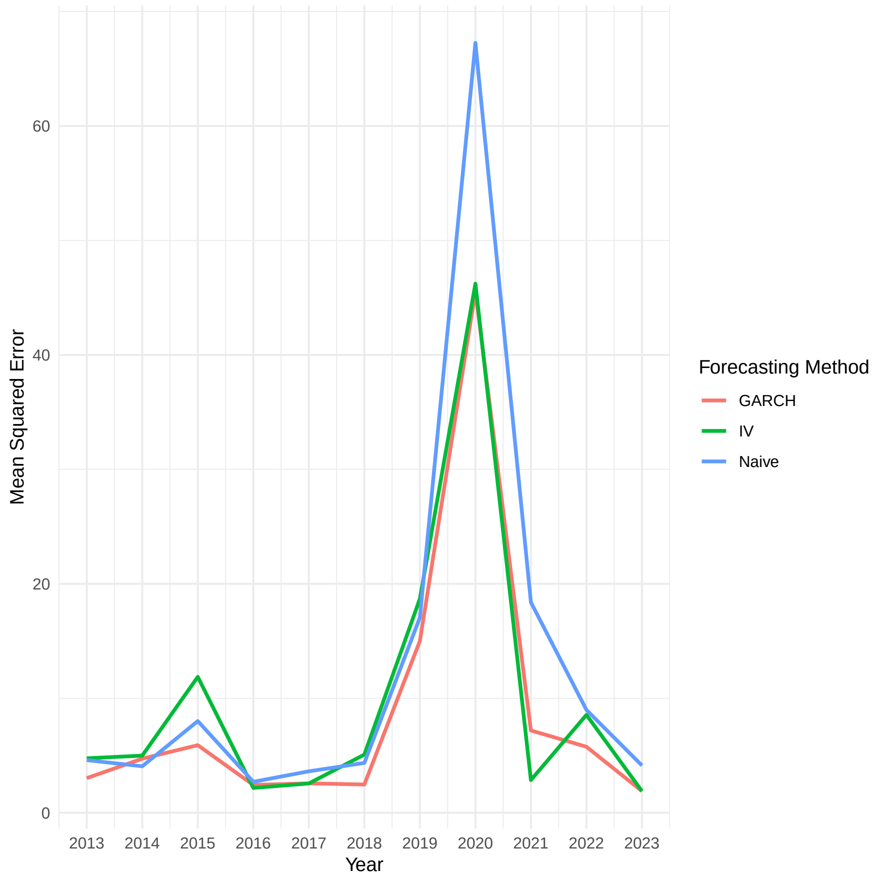

Figure 4: Annual Mean Squared Error during 2013-2023, for each forecasting method.

In the next section will discuss the implications of the results of section 4.1.

4.2 Analysis and Interpretation of the Results

Overall, the MSE values, accompanied by the bootstrapped confidence intervals, offer a fair, but shallow understanding of the performance of each model. The confidence intervals quantify the uncertainty of our MSE estimates, reinforcing our understanding of the reliability of each method. However a more intricate, year by year analysis reveals that the relative accuracy of the different models change over time. It is particularly interesting to note that each individual model had at least one year where it outperformed the other two methods.

The year 2020 presents a real challenge for all the tested models, due to the disturbances to the economy caused by the Covid-19 pandemic. All methods performed poorly in forecasting the volatility during this year. Most likely this is due to a very short period of extremely high volatility, separating this year from other years with similar average volatility levels. During this period GARCH and IV performed very similar, both outperforming the Naive forecasts, reaffirming that the GARCH and IV forecasting methods are somewhat resilient to market stress.

The comparatively bad forecasting performance of the Naive model suggests that this method faces particular challenges during drastic market fluctuations.

The parallell performances of the GARCH and IV methods during both their best and worst years of performance might indicate some underlying similarities in their reactions to market events. It is likely that many investors already incorporate GARCH predictions in their analysis, which would embed the volatility forecasted by GARCH into the prices of the options used for the IV. However, the lack of a consistent pattern of relative performance between the GARCH and IV methods vouches for more complexity in forecasting volatility, highlighting the importance of considering different methods when forecasting future volatility.

In summary, the GARCH method showed superior accuracy in reference to the lowest overall MSE and it is well suited in times of great market uncertainty. Meanwhile, the other methods show glimmers of prominence, as can also be observed in the overlapping confidence intervals. This once again underlines the importance of using not only one method for volatility forecasts and that there is no one-size-fits-all solution to the problem.

5 Discussion

5.1 Comparison with Previous Research

The findings of this study offer further knowledge in the field and provides insights into the body of research associated with GARCH and IV models.

Our results resonate with the findings of Christensen and Prabhala [1998], that the information provided by option prices and historical volatility are somewhat identical. This was seen in the comparisons between the forecasting performance of the GARCH and IV methods in this study.

According to the meta analysis of Poon and Granger [2003] there is no single model consistently outperforming the others. In this study we found GARCH to be slightly more accurate than IV in general, but when comparing the methods on an annual basis no clear winner could be found. The findings in this study seem to match the conclusions of Poon and Granger.

The observation in this thesis, that whereas there is no clear winner among the GARCH and IV methods, the naive forecasting method seems inferior in general, echoes the findings of Engle [1982], who introduced the ARCH model.

However more recent research made by Basri and Sumitra [2019] suggests that naive models are still competetive due to their simplicity and resilience to pure randomness. This can be observed in the annual comparisons in this study, declaring the naive forecasts as superior for the year 2014.

In conclusion, this study offers contributions to the existing literature in the form of a case study on the NOK/SEK exchange rate. The findings reinforce some previous research and offer new information on the relative performance between volatility forecasting models under different market conditions.

5.2 Significance of Findings

The finding of this work are valuable both in the aspect of financial research, but also in terms of the practical applications in the finance industry. In the context of the academic sphere, this study contributes to the ongoing debate around different methods of volatility forecasting and their efficiency. The academic sphere benefits in particular from the relative performance comparison between the GARCH and IV methods, that was conducted in this thesis. From this specific case study of the GARCH and IV methods we reinforce the current consensus that there is no one-size-fits-all model for forecasting volatility.

In practice, this study provides insights to traders, risk managers and financial institutions in general dealing with the NOK/SEK exchange rate, but it also highlights the importance of volatility awareness in general. The idea that one method can be completely trusted for forecasting volatility is dismissed in this study, while confirming that the existing methods can give valuable forecasts with varying accuracy. This work advises caution in using one sole method.

The fact that GARCH overall performed the best emphasises the need to account for volatility clustering and varying volatility over time. The periods where the naive forecasts performed the best can help to remind the financial industry of the randomness of financial markets, proving that even the more advanced, prominent methods have limitations.

Lastly, the difficulties in forecasting volatility during 2020 underlines the notion that unknown situations can always arise, which are very challenging to model, without previous knowledge on similar economic conditions. It signals the need for rigid risk management during market turmoil.

5.3 Limitations and Challenges

This study has been subjected to some limitations and challenges. First of all, the limited scope of the data in terms of the time span, and also the rather specific choice of exchange rate, makes the study not as general and useful to everyone. While 20 years corresponds to a substantial amount of data, it might leave this research irrelevant to other time periods. Since both the NOK and SEK currencies are not that large in market cap, other pairs of currencies might behave differently.

In addition, the Implied Volatility data was even more limited. The data available was for 10 years (2013-2023) leaving us with an even smaller time period for analysis. This means our conclusions are not as reliable as they could have been, and they are therefore likely to be more impacted by randomness.

There are some flaws to the Implied Volatility data itself. Since the data is derived from option prices using the Black-Scholes option pricing formula, the assumptions that the formula rely on is a limitation to our comparison. The most important assumption is that of constant volatility, which may very well not hold true in reality. Even though the IV method performs relatively well, the constant volatility assumption should be considered a flaw of the method. The Black-Scholes formula also assumes no trading costs when buying or selling options. This might not be the case in reality, even though transaction costs may be insignificantly low. Supply and demand on the option market might affect the prices, possibly skewing the results.

The rolling window approach for parameter estimation might limit the quality of the GARCH model forecasts. A rather large window of 1000 days is used in this study. Still the rolling window approach relies on the assumption that the latest data is the only data relevant, and this might not always be true.

Lastly, it is important to acknowledge that while MSE was used in this study, other metrics exist which might have yielded different results. Such metrics include Mean Absolute Error, Mean Absolute Percentage Error, Mean Absolute Scaled Error and others. Noting the alternatives, it is clear that MSE is not the only possible option, and it might not give the most suitable results. It was chosen in this study because of its properties; punishing larger errors more over smaller errors. But the MSE performance criterion might give skewed results in the presence of outliers, or even more during periods of higher volatility, as in the specific case of this study.

Despite these limitations, this study offers some valuable insights and new knowledge when forecasting volatility. As always, the results should be interpreted with these limits in mind.

6 Conclusion

6.1 Summary of Key Findings

In this study, our aim was to compare the accuracy of the standard GARCH model when compared to Implied Volatility derived from option prices. These methods were compared when forecasting volatility of the logarithm of the NOK/SEK exchange rate, using naive volatility forecasts as a benchmark. A comprehensive approach has been used, including modeling, forecasting and calculating errors to measure the forecasting performance of each method. The key findings are listed as follows:

GARCH Dominance

The analysis revealed that GARCH, while outperformed on occasions, consistently performed on par or close to the best performing forecasting method each year, often performing the best of all. Over the entire period, and as evidenced by the lowest MSE, GARCH was the best method when forecasting future volatility.

Yearly Variation

While the GARCH method often performed well, significant yearly variation in volatility was observed, as well as varying forecasting performance of all the models evaluated. Notably, GARCH and IV often had its best and worst years strongly correlated.

Consistency of Naive Forecasts

Despite the simplicity of the method, the Naive forecasts yielded consistent results, often on par with the GARCH and IV methods. Although yielding the highest MSE, the simplicity of the naive volatility forecasts speaks in favor of this method in comparison to the more sophisticated methods presented in this study.

GARCH and IV Correlation

There was a visibly noticeable correlation between the MSE of GARCH and IV, offering some insights into the overall difficulty of forecasting during periods of different circumstances, often affected by the general volatility at the time.

6.2 Implications for Theory and Practice

The findings of this study have some implications, not only for the theoretical understanding of financial volatility, but also to the practical application of financial forecasting and risk management.

Theoretical Implications

Affirmation of GARCH Models The results of this study provide empirical support to the general consensus in the field on the efficacy of GARCH models when forecasting volatility.

Performance of Implied Volatility This study is consistent with previous research on Implied Volatility, providing insights into the performance and limitations of Implied Volatility forecasting. This study provides further knowledge on the application of Implied Volatility in financial forecasting.

Practical Implications

Tool for Risk Management In practice, particularly in risk management and derivative pricing, the insights of this study could be utilised, providing more knowledge on the confidence of the forecasts available. The evidence of the superiority of GARCH models for volatility forecasting could increase the adoption rate of GARCH models in risk modelling practices.

Basis for Strategy Development The comparisons of this study, evaluating Naive, GARCH and IV forecasts provide a rigid foundation for developing investing and hedging strategies.

Signal for Model Refinement This study could ignite the curiosity of researchers and practitioners to explore new ways to refine and improve current models, especially regarding Implied Volatility due to its low cost of calculation.

Indicator for Sensitivity Analysis The observed correlation between GARCH and IV performance suggests further studies on their similar sensitivities to economic events and circumstances, which could improve the reliability of forecasting. As mentioned in section 4.2. it is likely prices of options are set using GARCH methods, which could partly explain the correlation between GARCH and IV performance.

6.3 Suggestions for Future Research

This study offers various suggestions for future research. At first, different GARCH models such as GARCH(p', q') with $p' > 1$ and/or $q' > 1$, EGARCH and FIGARCH could be compared and evaluated to assess whether their increased sophistication provides higher performance when forecasting volatility, compared to the GARCH(1,1) model used in this study. Other methods such as Machine Learning models, and more advanced Neural Network methods would be interesting to evaluate in this context.

As the NOK/SEK exchange rate was used as the underlying data for this study, it begs the question of what the results would have been for other exchange rates, as well as for various market indices, stock prices or commodities. Studies with different kinds of underlying data could provide more comprehensive insights into the performance of each of the forecasting methods.

When quantifying the results, this study used the MSE as the performance measure. Studies using alternative metrics such as MAPE or MAE could offer valuable insights or even lead to different conclusions.

More research could be done on the interpretation of Implied Volatility as it is often seen as the reflection of the market as a whole. As such IV could reveal elements not available through other, more analytical models, as IV could reflect the results of human behaviour and other phenomena in our civilisation. Still, the option prices are likely often grounded in some volatility forecasts, as the future volatility is vital for setting fair market prices.

In general, this research opens some doors and provides a solid starting point for further research in the field, hopefully inspiring more studies on volatility forecasting.

References

R Programming Language, 2023. URL https://www.r-project.org/.

K. Basri and I. D. Sumitra. Comparison of forecasting the number of outpatients visitors based on naïve method and exponential smoothing. In *IOP*, 2019.

F. Black and M. Scholes. The pricing of options and corporate liabilities. *The University of Chicago Press*, 1973. URL https://www.jstor.org/stable/1831029?origin=JSTOR-pdf.

Tim Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 1986.

E. P. Box, Gwilym M. Jenkins, Gregory C. Reinsel, and Greta M. Ljung. *Time Series Analysis: Forecasting and Control*. Wiley, 2016.

Bujar Bunjaku and Armin Näsholm. Forecasting volatility: A comparison study of model based forecasts and implied volatility. Lunds Universitet, Sweden. http://lup.lub.lu.se/student-papers/record/1627421, 2010.

Marek Capinski and Tomasz Zastawniak. *Mathematics for Finance: An Introduction to Financial Engineering*. Springer, 2011.

Michael R. Chernick. *Bootstrap Methods: A Guide for Practitioners and Researchers*. Wiley, 2008.

B.J. Christensen and N.R. Prabhala. The relation between implied and realized volatility. *Journal of Financial Economics*, 1998. doi: 10.1016/S0304-405X(98)00034-8.

Zhifeng Dai, Huiting Zhou, Xiaodi Dong, and Jie Kang. Forecasting stock market volatility: A combination approach. 2020. doi: 10.1155/2020/1428628.

F. Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 1982. doi: 10.2307/1912773.

Alexios Galanos. *Rugarch*, 2022. URL `https://cran.r-project.org/web/packages/rugarch/rugarch.pdf`.

Ser-Huang Poon and Clive W.J. Granger. Forecasting volatility in financial markets. *Journal of Economic Literature*, 2003. doi: 10.1257/002205103765762743.

Ludwig Schmidt. Volatility forecasting: Performance of garch models: A study on nordic indices during covid-19. Umeå Universitet, Sweden, 2021.

Costas Siriopoulos and Athanasios P. Fassas. Implied volatility indices - a review. *Elsevier*, 2019. doi: 10.1016/j.qref.2020.07.004.

Rolf Sundberg. *Lineära Statistiska Modeller*. 2022.

Yahoo-Finance and Eikon. Nok/sek exchange rate and implied volatility data, 2023. URL `https://github.com/oskarlaestander/Forecasting-volatility-data.git`.

Appendix

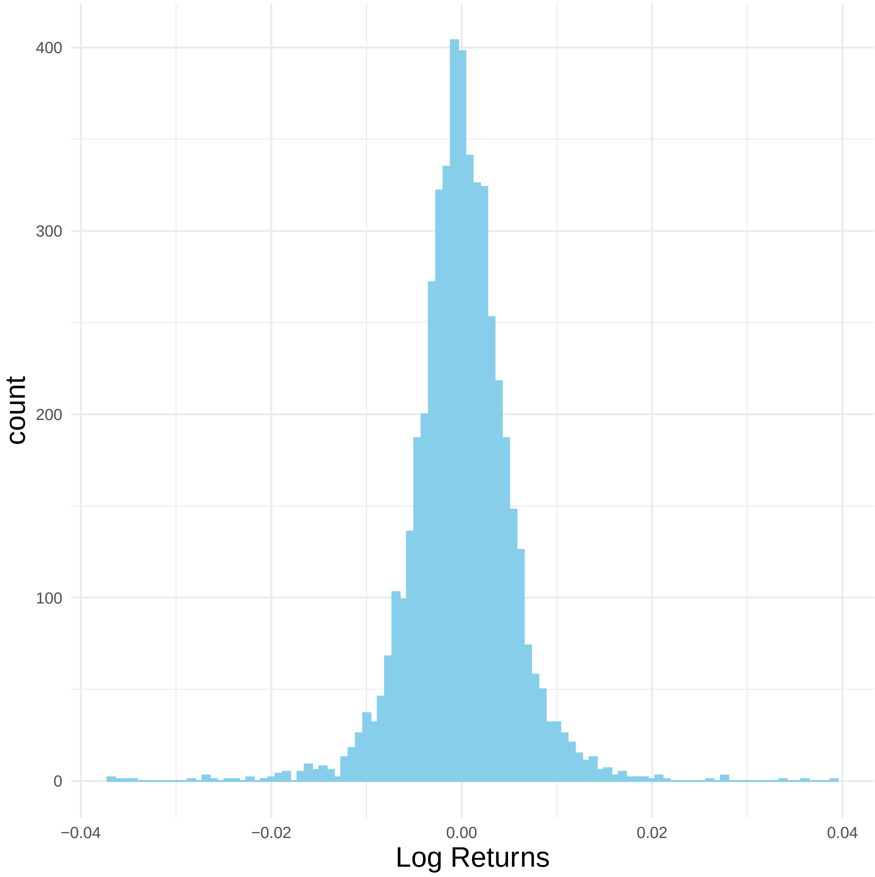

Figure 5: Histogram of log returns.

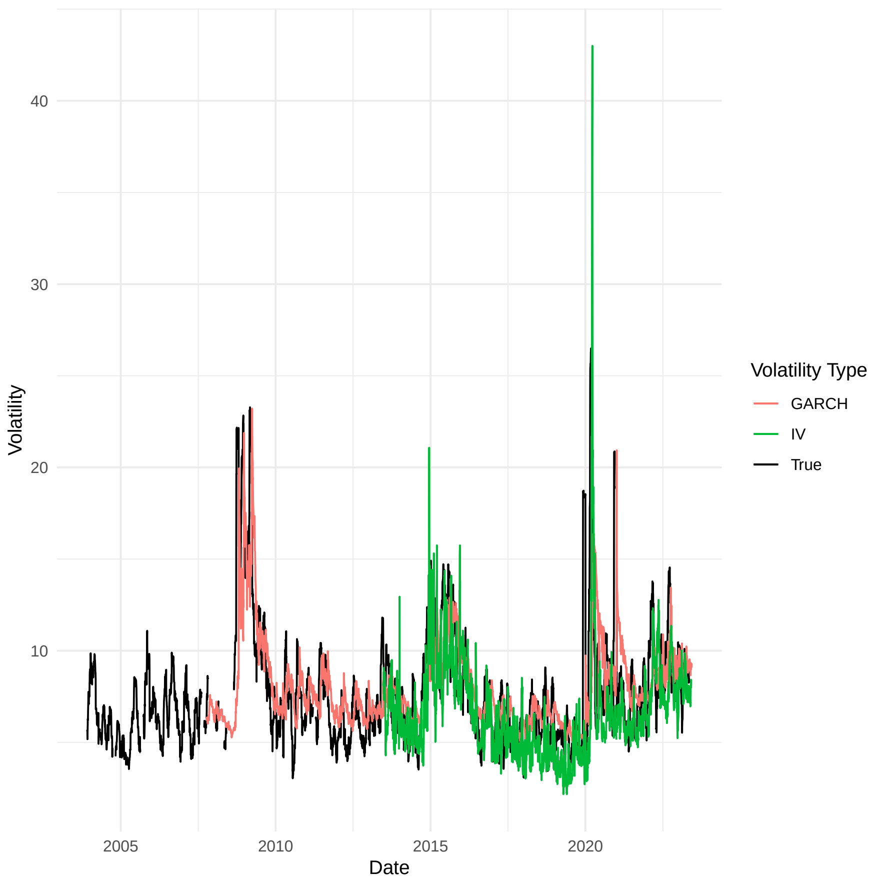

Figure 6: True, GARCH, and IV volatilities during the time period 2003-2023.

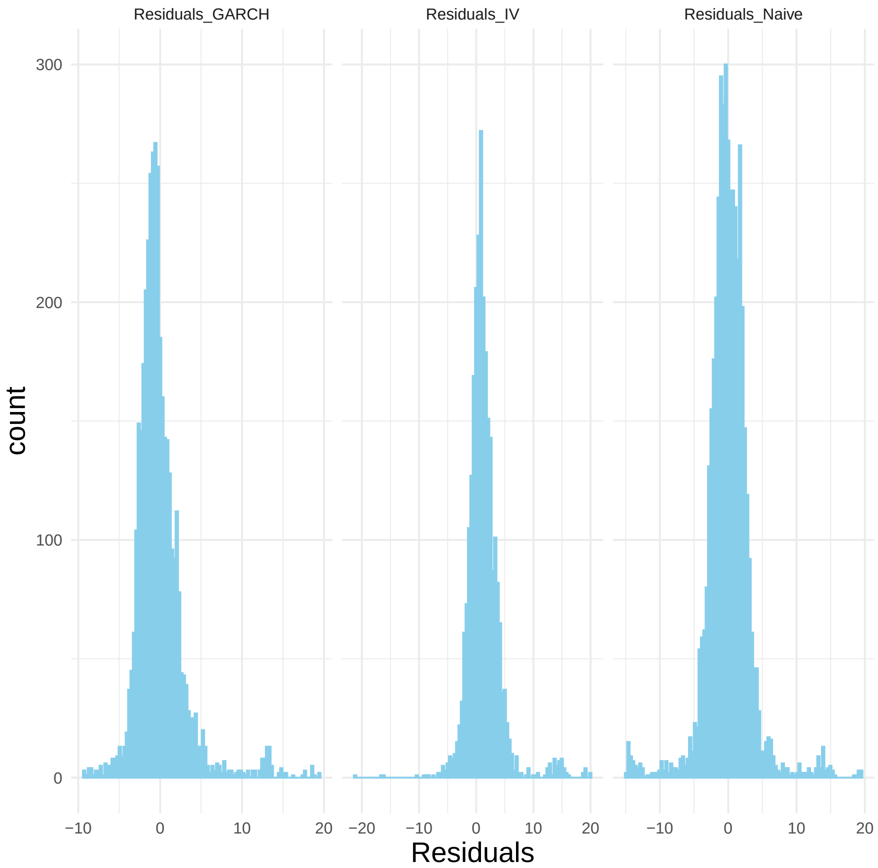

Figure 7: Histograms of residuals, for the three methods GARCH, OV and naive.

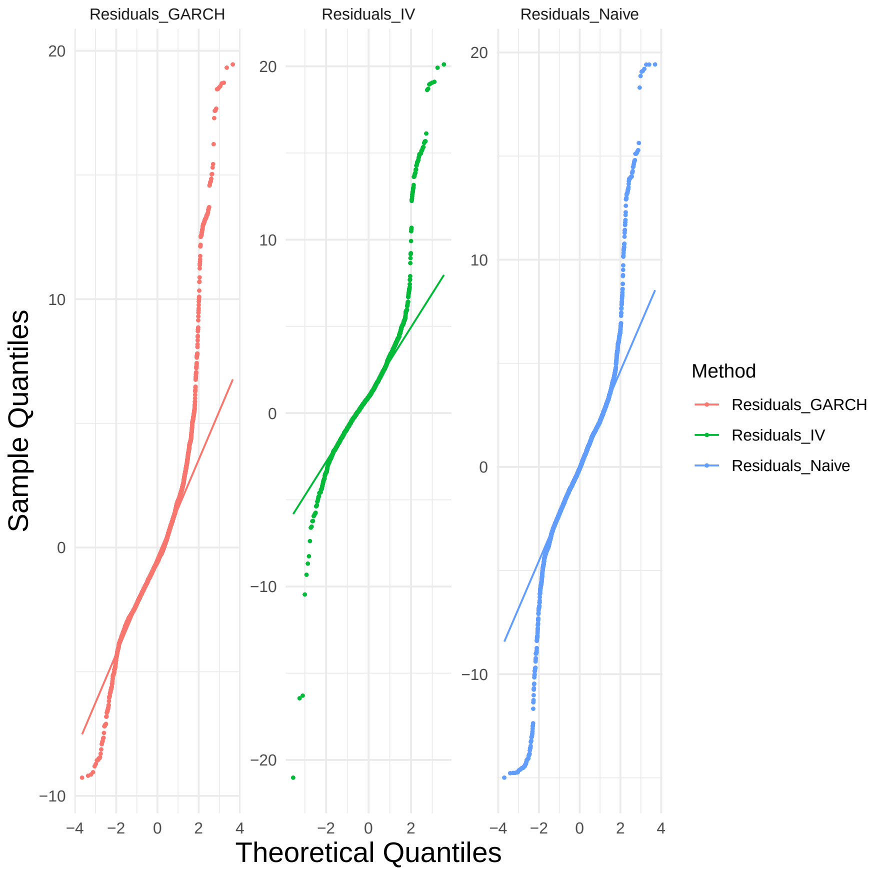

Figure 8: QQ-Plots of reiduals, for the three methods GARCH, IV and naive.

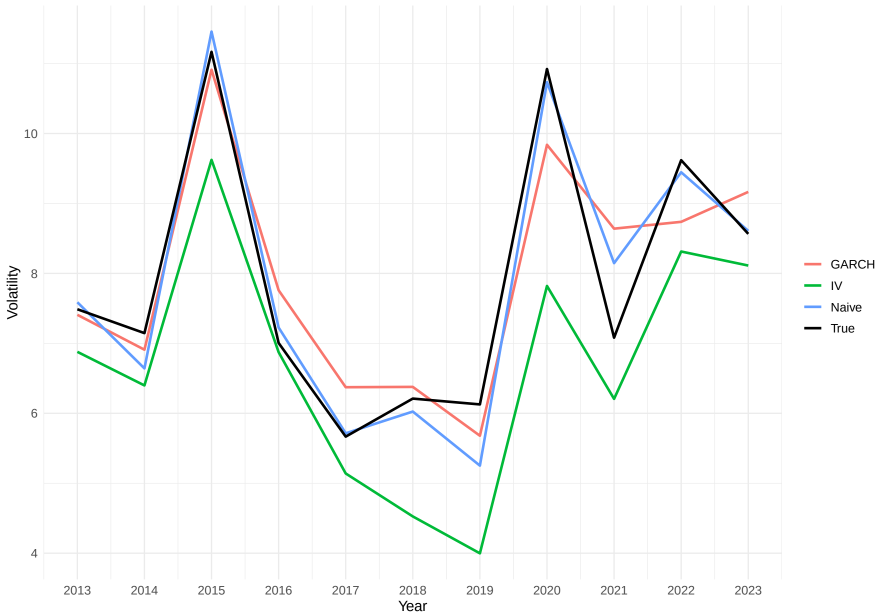

Figure 9: Mean of yearly volatility of the NOK/SEK rate from 2013 to 2023, for different methods.