Zeppelin Universität

Chair of Empirical Finance and Econometrics

The Information Content of VIX Volatility

Humboldt Research Project Project

in

Modul 11251 Forschungsprojekt

Name:	Sophia Charlotte Gläser
Student number:	15202284
Programme:	Corporate Management and Economics
Term:	Fall Semester 2018
Examinor:	Prof. Dr. Franziska Peter
Due date:	31.01.2019

Contents

List of Tables

List of Figures

Acronyms

BS Black-and-Scholes-Merton Model

CBOE Chicago Board of Options Exchange

SPX S&P 500

VIX Volatility index

Abstract

This paper investigates the information content of model-free implied volatility for daily realized volatility of the S&P 500, using the VIX from the Chicago Board of Options Exchange. In contrast to the Black-and-Scholes implied volatility, the VIX is not based on any specific option pricing model, therefore it provides a direct test of market efficiency and does not suffer from the joint hypothesis problem. Using the approach from an HAR-RV model as described by Corsi (2009) to account for the multi-scaling observed in financial data, the paper conducts an encompassing regression analysis to address the research question. The results show, that the VIX provides additional information compared to historic volatility, but is not able to subsume all the information contained in historic volatility. This results are robust to serial correlation and alternative estimation methods.

1 Introduction

Financial market volatility is of high interest for the financial sector. Asset return volatility is for example key input to the pricing of financial instruments like derivatives, or to risk measures such as the Value at Risk. Moreover, it gives information on the risk-return trade-off, which is a central question in portfolio allocation and managerial decision making.

However, as volatility is not directly observable it has to be estimated. Seeing its importance, it is not astonishing that considerable research has been devoted to the question, how volatility can be estimated and predicted. Whereas earlier approaches were mainly ARCH or stochastic volatility models, using historic volatility, there has recently been a growing interest in volatility implied from option price data (Bakanova 2010)[1].

[1]There are, of course, various other methods for volatility estimation and forecasting, such as nonparametric methods or neural networks based models. An encompassing overview of volatility estimation and forecasting can be found in Jiang, George J & Tian (2003).

Since options are contracts giving the holder the right to buy or sell an underlying asset at a specified date in the future, they are said to have a "forward-looking nature", meaning that they are supposed to be highly related to the market's expectation about the future volatility of the underlying asset over the remaining life of the option. Therefore, if market agents are rational (meaning that the market uses all available information to form its expectations about future price movements and volatility) markets are efficient and the model pricing the option is specified correctly, the volatility implied from option prices should be an unbiased and efficient estimator of future realized volatility (Bakanova 2010).

One popular approach for estimating implied volatility is the Black-and-Scholes-Merton Model (BS) implied volatility. The BS model is an option pricing model, using volatility as an input factor. By using observed option prices as the input and solving for volatility, it is possible to obtain a volatility measure that is widely believed to be "informationally superior to the historic volatility of the underlying asset" (Jiang, George J & Tian 2003, p.1305). Whereas early studies found it to be a biased forecast, the correction of several methodological problems (such as long enough time series, the inclusion of crisis periods and the use of non-overlapping samples) provided evidence, that BS implied volatility contains significant information content and is a more efficient forecast of realized volatility than historic volatility (ibid.).

However, the BS implied volatility has some specification problems. Firstly, it focuses on at-the-money options. At-the-money options are usually the once most actively traded and thus the most liquid ones, but this fails to include information contained in other options. Moreover, volatility estimation with the BS model includes the same assumptions as are made in the BS model itself, thus, tests based on the BS equation are joined tests of market efficiency and the option pricing model, and therefore suffer from model misspecification errors (ibid.).

That is why during the last years, implied volatility indexes which are not based on a pricing assumption have gained popularity. The idea was introduced by Britten-Jones and Neuberger (2000), extending the work of Derman and Kani (1994) Dupire et al. (1994), Dupire (1997) and Rubinstein (1994) on implied distributions. Contrary to the previously described BS implied volatility, their approach is derived directly from option prices and the no-arbitrage condition. Showing, that the risk-neutral return variance can be derived entirely from option price data, they provide a volatility measure which does not suffer from the joined hypothesis problem and moreover allows to incorporate not only at-the-money options. Several papers used the results from Britten-Jones and Neuberger (2000), extended their approach and tested the informational efficiency of the model-free implied volatility, such as Bakanova (2010), Taylor, Yadav and Zhang (2010) or Jiang, George J & Tian (2003). Jiang, George J & Tian (ibid.), for example, find that it subsumes all information contained in both BS and historic volatility using an OLS regression and variance data. However, in the OLS regression using volatility data and log specifications, the model-free implied volatility is not more significant than historic volatility. In 2003, Chicago Board of Options Exchange (CBOE) reacted to this overall confiming academic findings and changed the calculation of its Volatility index (VIX) from the BS to the model-free implied volatility. This paper will directly use the VIX index from CBOE and test the data for a 17-year time period, including the financial crisis around 2008, using both univariate and encompassing regression analysis. Moreover, the historic volatility information included in the model is extended. Previous research mainly used only the one-day-lagged realized volatility, whereas this paper uses the approach from the HAR-RV model from Corsi (2009), and includes equally weekly and monthly historic volatility averages in the regression.

2 Selected Estimation and Modelling Procedures for Volatility

As mentioned in section 1, volatility can not be directly observed and thus has to be estimated. This section briefly presents a concept for estimating realized volatility and an approach for modelling volatility, which is able to capture very well the stylized facts observed with financial data.

2.1 Estimating and Modeling Volatility using Historic Volatility

2.1.1 Estimating Realized Volatility

By definition "volatility seeks to capture the strength of the (unexpected) return variation over a given period of time" (Torben G Andersen, Tim Bollerslev and Diebold 2009, p.7). As terminology is not consistent in previous research, in this paper return variance is simply the second moment distribution characteristic, and return volatility its square root, the standard deviation.

With volatility being a latent variable, there are multiple concepts for volatility. According to Torben G Andersen, Tim Bollerslev and Diebold they can be grouped in (i) the *notional volatility* corresponding to the ex-post sample-path return variability over a fixed time interval, (ii) the ex-ante *expected volatility* over a fixed time interval or the (iii) the *instantaneous volatility* corresponding to the strength of the volatility process at a point in time. Fur the purpose of this paper the notional ex-post sample-path return variability is of importance. Volatility measures usually represent the average volatility over a discrete time period, as a continuous record of price data is not available, and even for very liquid markets, price data are distorted by micro-structure effects. It can however be shown, that under some assumptions the sum of squared high-frequency returns is a consistent estimator for the return variance. This results mainly builds on the work of (ibid.) and will only be briefly introduced here.

To begin with, it should be assumed that we have a continuous-time no-arbitrage setting. As return volatility aims to capture the strength of the unexpected return variation, one needs to define the component of a price change as opposed to an expected price movement. Torben G Andersen, Tim Bollerslev and Diebold (ibid.) show, that under certain assumptions the instantaneous return process can be decomposed into an expected return component, and a martingale innovation (in the discrete time setting this decomposition is more complex, for this paper it shall only be relevant that the martingale part is still the dominant contribution to the return variation over short intervals). Furthermore Torben G Andersen, Tim Bollerslev and Diebold (ibid.) show, that the cumulative sample path variability of this martingale component can be represented by a quadratic variation process, and they define the ex-post measured *notional variance* as the increment to this quadratic variation. Moreover they show, that this notional variance can be consistently estimated, using high-frequency returns or a large sample of returns, with the *realized variance*, defined as:

$$v^2(t,h;n) = \sum_{i=1}^{n} r(t-h+(i/n)\times h, h/n)^2 \tag{1}$$

over any fixed $[t-h,t], 0 < h$ time interval (Their proposition 5) This realized variance is simply the second sample moment of the return process, scaled by the number of observations n.

So in summary, the increment to the quadratic return variation which is the past variance, can be consistently and well approximated through the accumulation of squared high-frequency returns. Taking the square root of this realized variance, the *realized volatility* is obtained.

2.1.2 Modelling Volatility - the HAR-RV Model

Having introduced the concept of measuring realized volatility, there are multiple approaches that try to model the volatility process. Although volatility is not directly observable, it has some characteristics that are commonly examined and can be used to build volatility models.

These commonly observe stylized facts include for example high excess kurtosis for daily return series, and clustering of return variability, meaning that periods of large volatility seem to be followed by high volatility, and periods of low volatility seem to be followed by low volatility (Tsay 2005). Moreover, the autocorrelations of the square and absolute returns show a very strong persistence over long time periods. If return distributions with regard to different time horizons are observed, they show tail-crossover, thus the shape depends on the time scale. As the time scale increases, the return distribution slowly converges to the normal distribution, however very slowly. Moreover, financial data show evidence of scaling and multi-scaling Corsi (2009).

There are various modelling approaches for volatility. This paper will refer to the HAR-RV model, based on Corsi (2009), as it is able to reproduce the stylized facts described very well. The model presents a volatility cascade from low to high frequencies, in which not only daily lagged realized volatility, but also weekly and monthly volatility influence future realized volatility.

To start with, a quick note on the terminology should be made. The model assumes that prices follow the standard continuous-time process, represented by the stochastic differential equation

$$dp(t) = \mu(t)dt + \sigma(t)dW(t), \ 0 \leq t \leq T \tag{2}$$

where $p(t)$ is the logarithm of the instantaneous price, μ is a finite variation stochastic process, $W(t)$ standard Brownian motion and σ a stochastic process independent of $W(t)$. For this process, the variance is the *integrated variance*, which is integral of the instantaneous variance over the one-day interval

$$IV_t^{(d)} = \int_{t-1d}^{t} \sigma^2(w)dw, \tag{3}$$

and the corresponding volatility it's square root: $\sigma = \sqrt{IV_t^{(d)}}$. Torben G Andersen, Tim Bollerslev and Diebold (2009) show, that instead of the abstract martingale representation of the return decomposition that was described in 2.1.1, the continuous sample paths can also be represented using stochastic differential equations, as it is done here. Then the integrated variance equals the notional volatility, and can equally be estimated using the sum of squared returns. Thus, the HAR-RV model uses the following approximation for realized volatility over the one-day interval, which will be adopted in this paper:

$$RV_t^{(d)} = \sqrt{\sum_{j=0}^{M-1} r_{t-j\times\Delta}^2} \tag{4}$$

with $\Delta = 1d/M$ being the sampling frequency and $r_{t-j\times\Delta}^2$ defined as the continuously compounded Δ-frequency returns (with t the day and j the time within the day).

The idea behind the HAR-RV model is closely connected to the *Heterogeneous Market Hypothesis*, which describes the presence of heterogeneity across market participants and was presented by Müller et al. (1993). This view of financial markets grounds on the fractal model, introduced and applied to financial markets by Mandelbrot (1963). The approach of the fractal model is to analyse (time series) objects on different time scales and compare the obtained results. The argument is, that conventional time series analysis, focusing on regularly spaced observations, does not capture the real nature of the raw data, as the usual time choice for recording observations (e.g. a day) is arbitrary. Using this fractal approach and empirical finding of volatility characteristics[2]) gave

[2]Müller et al. (1993) observed that the decline of the return autocorrelation function is not exponential, as suggested for example by lower-order GARCH or ARCH models, but rather hyperbolic. Assuming that each of the distinct components has an exponential

rise to the heterogeneous market hypothesis. This hypothesis states that the market gives rise to heterogeneous trading behaviours since different market participants or components have different time horizons for their trading goals and for their consideration of past events. The time span has on the one end the high-frequency dealers such as market makers, in the middle medium term dealers and on the other end the low-frequency dealers, such as central banks or commercial organizations. Driven by this components, "the market is heterogeneous with a "fractal" structure of the participants' time horizon" (Müller et al. 1993, p.12).

Corsi (2009) adds to the observations of Müller et al. (1993), that volatility has an asymmetric behaviour of influence, meaning that volatility over longer time periods has a stronger influence on volatility observed over short periods than conversely. The pattern that emerges from this is a volatility cascade from low to high frequencies. To formalize the model the *latent partial volatility* $\tilde{\sigma}_t^{(\cdot)}$ is defined as the volatility generated by a certain market component. To account for short-term, medium-term and long-term traders, the time horizons of one day ($1d$), one week ($1w$) and one month ($1m$) are considered, and denoted by $\tilde{\sigma}_t^{(d)}$, $\tilde{\sigma}_t^{(m)}$ and $\tilde{\sigma}_t^{(w)}$. In the model, each of this volatility components corresponds to a market participant, that forms the expectation for the volatility of one period ahead based on both the observation of the current realized volatility according to the own time frame and on the expectation of the one horizon longer volatility. For example for the market participant with a daily horizon, the latent volatility would be

$$\tilde{\sigma}_{t+1d}^{d} = c^{(d)} + \Phi^{(d)} RV_t^{(d)} + \gamma^{(d)} E_t[\tilde{\sigma}_{t+1w}^{(w)}] + \tilde{w}_{t+1d}^{(d)},$$

decline with different time horizons, in sum comes close to a hyperbolic decline. Moreover, if market participants were homogeneous, volatility should be negatively correlated with market activity, as the price should converge to the "real value". However, they are positively correlated, which might be explained by the fact that actors react and execute in different market situations (ibid.).

where $\tilde{w}^{(d)}_{t+1d}$ is the return innovation. Substitution the latent volatilities of the different horizons and defining the latent volatility as the daily integrated volatility ($\sigma^{(d)}_t$), the cascade model can be written as

$$\sigma^{(d)}_{t+1d} = c + \beta^{(d)} RV^{(d)}_t + \beta^{(w)} RV^{(w)}_t + \beta^{(m)} RV^{(m)}_t + \tilde{w}^{(d)}_{t+1d}, \tag{5}$$

Observing the volatility data ex-post, $\sigma^{(d)}_{t+1d}$ can be written as the realized volatility, a functional form for time series representation can be written as:

$$RV^{(d)}_{t+1d} = c + \beta^{(d)} RV^{(d)}_t + \beta^{(w)} RV^{(w)}_t + \beta^{(m)} RV^{(m)}_t + w_{t+1d}, \tag{6}$$

where w_{t+1d} includes the innovation component and the measurement and estimation error. The realized volatility for the weekly and monthly aggregated periods are simply the rolling average over the respective periods. Using simulated return data, Corsi (2009) shows, that the HAR-RV simulated returns and volatility reproduce the stylized facts mentioned above very well. The data has not only the excess of kurtosis but also the tail cross-over, meaning that the fat tails get thinner as the aggregation level increases. The simulated data is also able to reproduce the long memory of the empirical data. Moreover, using OLS, in-sample and out-of-sample forecast, Corsi (2009) shows, that the past volatility components all have significant information content for one day ahead realized volatility and show good forecasting performance.

2.2 Estimating Volatility Using Option Price Data

2.2.1 The General Idea and Evolution of Implied Volatility

Apart from using historic data for volatility modelling, the forward-looking nature of options data can be used to augment the information set. This approach is termed *implied volatility*. The intuition behind implied volatility is, that option prices can be seen as reflecting market participants' expectations of the future movements of the underlying asset. Assuming that the market is efficient, as described by Malkiel and Fama (1970), and that the asset pricing model is correct, the implied volatility derived from the option prices should not only subsume all information contained in historic volatility but also be a more efficient forecast of future volatility (Jiang, George J & Tian 2003).

One example for implied volatility is the *BS implied volatility*, building on the BS asset pricing model as presented by Black and Scholes (1973), which uses volatility as one input for pricing options. Having the option prices available on the market, it is possible to extract a value for the expected volatility by inverting the theoretical asset pricing model.

Whereas previous research yielded ambiguous results concerning the informational efficiency and forecasting ability of BS implied volatility, more recent research made several methodological and data corrections and found evidence supporting the hypothesis that implied volatility has predictive power for future volatility. These corrections included using longer time series, adapting high-frequency asset returns or using non-overlapping samples (Jiang, George J & Tian 2003). Bent Jesper Christensen, Hansen and Prabhala (2001) found that under certain of these measurement errors (such as overlapping samples) statistical tests are no longer meaningful. Correcting for this measurement errors results in support for the hypothesis, that implied volatility subsumes all information contained in historic volatility.

However, even though there is evidence in support of BS implied volatility, there are several disadvantages of using the implied volatility with an asset pricing model. One disadvantage is, that the BS implied volatility relies mostly on information from at-the-money options, which are generally the most actively traded ones. This, however, fails to incorporate information contained in other options. Moreover and more importantly, testing for the BS implied volatility always means testing for market efficiency and the BS model jointly. Assumptions of the BS model include constant volatility, no transaction costs or taxes, no dividend before option maturity, no arbitrage, continuous trading, constant risk-free interest rate, divisible securities and no short sell (S.-h. Poon and Granger 2003). As it is not possible in this setting to test for market efficiency or the pricing assumptions separately, these tests are subject to model misspecification errors (Jiang, George J & Tian 2003). This is why an implied volatility that does not rely on an asset pricing model was introduced.

2.2.2 Model-Free Implied Volatility - the VIX

The idea of model free implied volatility was introduced by Britten-Jones and Neuberger (2000) who show that the risk-neutral realized volatiltiy can be derived from a set of options with matching expiration, thus extendending the approach of Derman and Kani (1994) Dupire et al. (1994), Dupire (1997) and Rubinstein (1994) on implied distributions. Contrary to BS implied volatility, no assumptions are made concerning the pricing process. Instead, a complete set of option prices is taken as given and as much information as possible is extracted about the underyling price process (Britten-Jones and Neuberger 2000).

Britten-Jones and Neuberger (ibid.) show in an approach resembling a binomial tree, that the probability of the stock price reaching any particular level and of a price move is determined by the initial set of option prices. Moreover they show that the risk-neutral expected sum of squared returns between to dates is given from the set of options expiring on these to dates (their proposition 2). This formula does not use any specific option pricing model to derive implied volatility and is only based on no-arbitrage conditions, therefore it solves the joint-hypothesis problem and test directly for market efficiency (Jiang, George J & Tian 2003).

Several paper used this result and compared the informational efficiency of model-free implied, BS and historic volatility. One paper that examines the information content of model-free implied volatility is Jiang, George J & Tian

(2003). They use the approach of Britten-Jones and Neuberger (2000), extend the formula to asset price processes with jumps and test the informational efficiency of this model-free implied volatility in comparison to both BS implied and historic volatility, using both univariate and encompassing regression analysis. Using option data from the S&P 500 for the model-free implied volatility calculation and 5-min returns to calculate daily realized volatility, they examine monthly non-overlapping samples of a 6-year sample period between June 1988 and December 1994. Their findings are, that the model-free implied volatility subsumes all the information contained in the BS implied volatility and past realized volatility. Another example is the paper of Bakanova (2010). With daily data from oil futures between November 1986 and December 2006, they evaluate the information content of model-free implied volatility in comparison to historic volatility using a monthly frequency. Using regression analysis, they come to the result that implied volatility subsumes the information contained in historical volatility. In contrast to these results is for example Taylor, Yadav and Zhang (2010). With individual stock data from 149 U.S. firms between January 1996 and December 1999, they find that for a one-day-ahead estimation, historic volatility outperforms model-free implied volatility- When however extending the prediction horizon, the option implied volatility is more informative.

One of the first implemented implied volatility index was the VIX from the CBOE. It is computed each trading day on a real-time basis, with data available dating back to January 1986. It's introduction in 1993 had the intention to provide both a benchmark for market volatility in the short term, and a volatility index on which futures and options could be written and traded. At this time the VIX was based on the BS option pricing model using S&P 100 options, as they where the most actively traded in the U.S. (Fleming, Ostdiek and Robert E. Whaley 1995). In 2003 the VIX adopted to the changes that took place in the options market. First, the S&p 500 option market superseded the S&P 100 options market as the most actively traded option market in the U.S., thus following 2003 the VIX was based on the S&P 500 options. Secondly, option index trading behavior changed. Whereas in the 1990s both call and put index options where equally important, over the years both out-of-the money and at-the-money puts gained popularity as they were bought by portfolio insurers. Thus the VIX also started to include out-of-the money options in it's calculation, bringing another advantage compared to BS implied volatility (Robert E Whaley 2008). Finally, the VIX calculation changed in 2003, adopting to the model-free implied volatility approach, which was in 2003 widely used by financial theorists, risk managers, and volatility traders alike (Chicago Board of Options Exchange 2009).

The VIX is constructed in the way that it eliminates mis-specification and "smile" effects thus making it an accurate measurement of implied volatility (Blair, S.-H. Poon and Taylor 2001). The formula for calculation is

$$VIX = \sqrt{\frac{2}{T}\sum_i \frac{\Delta K_i}{K_i^2} e^{RT} Q(K_i) - \frac{1}{T}(\frac{F}{K_0} - 1)^2} \times 100 \tag{7}$$

with T being time to expiration , F the forward index level, K_0 the first strike below the forward index level, K_i the strike of the i^{th} out-of-the-money option, ΔK_i the interval between the strike prices, R the risk-free rate and $Q(K_i)$ the midpoint of bid-ask spread for each option with strike K_i. For the calculation first the options used for calculation are selected. The VIX always uses near- and next-term put and call options, with more than 23 and less than 37 days to expiration. Each week, the options used for calculation roll over to new maturities, thus long-term options become next-term options, and in the following week fall out of the sample. Thus the VIX will always reflect an interpolation of the volatility between these two option maturities, using the midpoint bid-ask price as the transaction price are subject to bid-ask bounde (S.-h. Poon and Granger 2003). After the selection of the options for the calculation, the volatility for both the near- and next-term options is calculated. Out of this two volatility, the 30-day-weighted average is calculated, so that the VIX gives information about the expected volatiltiy over the next 30 days. (Chicago Board of Options Exchange 2009).

3 Methodology and Data

The aim of this paper is to use the results from Jiang, George J & Tian (2003), who justify both the general validity of the model-free implied volatility, and the decision of CBOE to modify the calculation of the VIX index. This paper will examine directly the information content of the VIX index compared to historic volatility, using a more recent and longer time period of 17 years, between January 2000 and December 2017. Moreover, more information regarding the historic volatility is included with the use of the HAR-RV model by Corsi (2009).

In this section, first the sample and the model is introduced, then the testable hypothesis are formalized. Finally the limitation of the model are briefly discussed.

3.1 Data and Calculation of Input Factors

The data used in this study comes from several sources. The realized variance of the S&P 500 index is obtained from the *Oxford-Man Institute's realized library V0.2* (2009). Daily VIX values are the closing values from CBOE. As of September 22, 2003 CBOE started to use the model-free implied volatility for the VIX, but calculated the prices with the new methodology ex-post, dating back until 1990. As the sample from this study covers both the period before and after the revised methodology, the data was taken both from the ex-post Chicago Board of Options Exchange (n.d.[b]) and the daily-updated Chicago Board of Options Exchange (n.d.[c]). The SPX closing values, used only for visualisation, are also taken from Chicago Board of Options Exchange (n.d.[a]). The sample consists of daily data in the period from January 2000 to December 2017.

The calculation of the daily realized variance is done by *Oxford-Man Institute's realized library V0.2* (2009) using the sum of squared 5-min high-frequency returns, as explained in section 2.1.1. The formula is given by

$$\sigma_t = \sum x_t^2 \tag{8}$$

with $x_t = X_t - X_{t-1}$ and X_t being the logarithm of the price at time t. Previous research has used different sampling frequencies for the calculation of the realized volatility. Whereas earlier studies used mainly daily returns, most recent studies argue in the favor of using intraday returns, as they have certain advantages over daily data (Jiang, George J & Tian 2003). For example Torben G. Andersen, Tim Bollerslev et al. (2003) point out that high-frequency returns help both for predicting again high-frequency returns, but also that they contain information for longer horizons, such as monthly or quarterly. Moreover and more importantly, Torben G. Andersen and T. Bollerslev (1998) show that realized volatility calculated using squared returns produces inaccuracies with daily returns. The realized variace data is cleaned by *Oxford-Man Institute's realized library V0.2* (2009). First, entries outside the timestamp when exchanges are open are deleted. Secondly, entries with the same time stamp are replaced with the median bid-ask price. Thirdly entries with a negative spread (as they violate the no-arbitrage condition) or extremely large spread (50 times the median of the day) are deleted. Lastly, entries for which the mid-quote deviated largely from the mean where deleted. For this paper, to obtain the realized volatility, the realized variance is simply squared. Moreover, for a more intuitively comparison to the VIX, the values are also multiplied by 100:

$$RV_t = \sqrt{\sum x_t^2} \times 100. \tag{9}$$

For the historic volatility the lagged realized volatility is used. As the approach from the HAR-RV model from section **??** is used, additionally to lagged daily realized volatility, the lagged weekly and lagged monthly volatility is needed.

They are computed using the rolling average over the respective time periods, thus weekly realized volatility over the last 5 trading days is calculated:

$$RV_t^{(w)} = \frac{1}{5}(RV_t^{(d)} + RV_{t-1d}^{(d)} + ... + RV_{t-4d}^{(d)}). \tag{10}$$

The monthly volatility is calculated alike, using a period of 20 days.

For the VIX the daily closing value is taken, as it contains the information from the whole day. The calculation is described in 2.2.2. To display the data in a more intuitive manner, the annualized VIX is divided by $\sqrt{252}$ as in Blair, S.-H. Poon and Taylor (2001) and Robert E Whaley (2008), in order to obtain an index with daily information content.

The S&P 500 together with the VIX is illustrated in 6 and the realized vola is added in 6. The summary statistics can be found in 2 and using logarithm in **??** in the appendix. The summary statistics show, that the mean of the VIX is for every time period a slightly higher than the mean of the realized variance for all time periods. From the graphics, too, it can be seen that usually the VIX is a little higher than the realized volatility. This is consistent with the findings previous research, e.g. Jiang, George J & Tian (2003). However, the graphics show, that during crisis periods, the realized volatility exceeds the VIX. For example, whereas both the realized volatility and the VIX are particularly high during the period of the crisis between 2008 and 2012, the realized variance is significantly higher. To account for this period, a dummy for the crisis 2008 - 2012 was included in the model.

Moreover the summary statistics show, that the skewness and kurtosis of the logarithmic specification is closer to the one of the normal distribution. Consequently, a regression based on the log volatilities is also specified and should be statistically better specified than those based on simple volatility.

The correlation matrix for the realized volatility, it's lagged specifications and the model-free implied volatility and it's lags can be found in **??**. Overall, the realized volatility is highly correlated with both the past realized volatility and the past VIX values. For comparison, also the weekly and monthly VIX are included, calculated in the same way as the weekly and monthly realized volatility. Whereas for the one day lag the correlation with the VIX is higher, for the weekly and monthly averages the correlation with the historic realized variance is higher.

3.2 Methodology: Linear Regression and HAR-RV model

Consistent with prior research, for example Jiang, George J & Tian (ibid.), Canina and Figlewski (1993) or Bent J Christensen and Prabhala (1998), both univariate and encompassing regression analysis is used to analyse the information content of volatility measures. However, contrary to previous research, the approach from the HAR-RV model, described by Corsi (2009) is integrated. Thus, not only one day lagged realized volatility is used as an explanatory variable in the regression, but also weekly and monthly realized volatility, computed as described in section 3.1.

In the univariate regression, the realized volatility is regressed once solely on the historic data, and once only on the VIX. For comparison, realized volatility is regressed on both historic data and the VIX in the encompassing regression analysis. Thus the encompassing regression analysis gives information about the relative importance of the volatility measures, and whether the VIX subsumes the information from the historic volatility (Jiang, George J & Tian 2003). For each explanatory variable, the value of the current day t is used, whereas for the explained variable, the one-day-ahead value $t + 1d$ is used. Like this, all the information available on day t is used to evaluate the one-day-ahead realized volatility. The three regressions are then given by:

$$RV_{t+1d} = c + \beta_t^{RV,d} RV_t^{(d)} + \beta_t^{RV,w} RV_t^{(w)} + \beta_t^{RV,m} RV_t^{(m)} + \beta^{crisis} crisis \tag{Reg1a}$$

$$RV_{t+1d} = c + \beta_t^{VIX} VIX_t + \beta^{crisis} crisis \tag{Reg2a}$$

$$\begin{aligned} RV_{t+1d} = c &+ \beta_t^{RV,d} RV_t^{(d)} + \beta_t^{RV,w} RV_t^{(w)} + \beta_t^{RV,m} RV_t^{(m)} + \beta_t^{VIX} VIX_t \\ &+ \beta^{crisis} crisis \end{aligned} \tag{Reg3a}$$

with RV_{t+1d} the realized volatility one-day-ahead, $RV_t^{(d)}$ the daily realized volatility, $RV_t^{(w)}$ the weekly realized volatility, $RV_t^{(m)}$ the monthly realized volatility, VIX the VIX closing value and $crisis$ the dummy variable, indicating one in the time of the financial crisis (2008-2012) and zero otherwise. The same regressions are specified with the logarithm for each variable.

$$\begin{aligned} \ln(RV_{t+1d}) = c &+ \beta_t^{RV,d} \ln(RV_t^{(d)}) + \beta_t^{RV,w} \ln(RV_t^{(w)}) \\ &+ \beta_t^{RV,m} \ln(RV_t^{(m)}) + \beta^{crisis} crisis \end{aligned} \tag{Reg1b}$$

$$ln(RV_{t+1d}) = c + \beta_t^{VIX} ln(VIX_t) + \beta^{crisis} crisis \tag{Reg2b}$$

$$\begin{aligned} ln(RV_{t+1d}) = c &+ \beta_t^{RV,d} ln(RV_t^{(d)}) + \beta_t^{RV,w} ln(RV_t^{(w)}) + \beta_t^{RV,m} ln(RV_t^{(m)}) \\ &+ \beta_t^{VIX} ln(VIX_t) + \beta^{crisis} crisis \end{aligned} \tag{Reg3b}$$

In all regression specifications the Newey-West covariance correction is used, to account for the possible presence of serial correlation in the data.

3.3 Research Question and Hypothesis

Model-free implied volatility uses the forward-looking nature of options in comparison to historic volatility and does not suffer from misspecification problems, moreover it aggregates information across all strike prices. It is therefore assumed, that it is informationally more efficient than historic volatility. BS implied volatility is not included, as previous research found that model-free implied volatility outperforms BS implied volatility.

The following hypothesis are tested:

Hypothesis 1 (H1): The VIX contains information about future realized volatility.

Hypothesis 2 (H2): The VIX has more explanatory value than the historic volatilities in estimating one-day-ahead realized volatility.

Hypothesis 3 (H3): The VIX adds explanatory value to the historic realized volatilities in estimating one-day-ahead realized volatility.

Hypothesis 4 (H4): The VIX incorporates all information regarding one-day-ahead realized volatility, the historic volatilities contain no information beyond what is already included in the VIX.

The hypothesis are formalized using the regressions in the following way:

For **H1**, the $H_0 : \beta^{VIX} = 0$ is will be tested in Reg1a and Reg1b, using t-tests. If the VIX contains information about future volatility, the slope estimate should be positive and significantly different from zero. For the **H2**, the adjusted R^2s and $AIC's$ should be compared in Reg1a to Reg2a and Reg1b to Reg2b. If the VIX can explain more variation of the one-day-ahead realized volatility than the historic volatilities, the R^2 and adjusted $R2$ should be larger in the second regression. However, for this paper the adjusted R^2 is preferred to the normal R^2, as it adds a penalty for adding additional independent variables, which could be misleading here because the first specification has more explanatory variables than the second one. For **H3**, the adjusted R^2s and $AIC's$ should be compared inReg2a to Reg3a and Reg2b to Reg3b. If the VIX adds explanatory value compared to the historic volatilities, the (adjusted) R^2 should increase in the third specification. Finally, for **H4**, the $H_0 : \beta^d = \beta^w = \beta^m = 0 \; and \; \beta^{VIX} = 1$ will be tested, using F-tests. If the estimates for the volatilities are zero and the estimate for the VIX is not, it should subsume all information contained in the historic volatilities.

3.4 Limitations

As volatility is stochastic, the ex-ante estimation will not equal the return volatility, as it is a measurement over an aggregated discrete time period (Torben G Andersen, Tim Bollerslev and Diebold 2009). The VIX might be flawed, as **jiang2007** showed. There might be a bias in estimated realized volatility due to autocorrelation in intraday returns (Jiang, George J & Tian 2003).

4 Results

This section presents the estimation results. The first subsection reports the results obtained with the regression analysis described in section 3.2. The second subsection provides the results of the robustness checks conducted.

4.1 Regression Analysis Results

In the following, the regression results are presented. The regression outputs can be found in table 4.1 for the logarithmic regression and 6 for the level regression, in the appendix. To make it easier to overview the results, only one regression table was included in this section. The logarithmic table is displayed, since as discussed in 3.1 it should be better specified, which is confirmed by it having a lower AIC than the level specification.

Firstly, if the VIX contains information about future realized volatility, the slope of the VIX in Reg1b and Reg2b should be positive and significantly different from zero. The $H_0 : \beta^{VIX} = 0$ can be rejected for both the level and the logarithmic regression specification, which implies that the VIX contains significant information for future volatility. Thus the **H1** can be confirmed.

Secondly, if the VIX has more explanatory value than the historic volatilities, the adjusted R^2 in the second regression with the VIX (Reg2a and Reg2b) should be larger than in the first regression with the historic volatilities (Reg1a and Reg1b). This not true for both the level and the logarithmic specification. Even though the R^2's are close, they are slightly larger in the first regression (0.726) than in the second (0.693). Thus, the **H2** can not be confirmed.

Thirdly, if the VIX adds explanatory value to the historic volatilities, the adjusted R^2 in the third regression (Reg3a and Reg3b) with the VIX included should be larger than in the first regression, containing only the historic volatilities. This is true for both specifications, in alignment with **H3**.

Finally, if the VIX subsumes all information contained in the historic volatilities, the historic volatilities should not be significantly different from zero, contrary to the VIX. Thus the $H_0 : \beta^d = \beta^w = \beta^m = 0 \; and \; \beta^{VIX} = 1$ should not be rejected. The results from the F-test can be found in the appendix in table 6 and **??**. The results show, that the H_0 can be rejected at the 0.001 significance level for both specifications. Thus, **H4** can not be confirmed.

Table 1: Logarithmic regression (whole sample)

	Dependent variable:		
	Realized Volatility		
	Reg1b	Reg2b	Reg3b
	(1)	(2)	(3)
Intercept	-0.043^{***}	-0.407^{***}	-0.187^{***}
	(0.007)	(0.017)	(0.014)
$ln(RV_t^d)$	0.344^{***}		0.263^{***}
	(0.027)		(0.025)
$ln(RV_t^w)$	0.395^{***}		0.284^{***}
	(0.035)		(0.035)
$ln(RV_t^m)$	0.208^{***}		0.014
	(0.024)		(0.029)
$crisis$	0.020^{*}	-0.224^{***}	-0.100^{***}
	(0.012)	(0.034)	(0.016)
$ln(VIX_t)$		1.472^{***}	0.648^{***}
		(0.048)	(0.045)
AIC	1874.2	2363.6	1551.6
Observations	4,434	4,434	4,434
R^2	0.726	0.694	0.745
Adjusted R^2	0.726	0.693	0.745
Residual Std. Error	0.299 (df = 4429)	0.316 (df = 4431)	0.288 (df = 4428)
Note:			*p<0.1; **p<0.05; ***p<0.01

4.2 Robustness Checks

Previous samples testing the information content of (model-free) implied volatility often used overlapping samples, meaning that the same option is used in several implied-volatility calculations. However, Bent J Christensen and

Prabhala (1998) showed, that the use of overlapping samples creates a telescopic overlap problem and thus standard statistical inferences are no longer valid. Therefore the same regression analysis was conducted using non-overlapping samples. Jiang, George J & Tian (2003) use monthly non-overlapping samples, using the first Wednesday of every month, since they calculate the implied volatility over a horizon on one month. The VIX however is calculated slightly differently. It contains near- and next-term options options between 23 and 37 days to maturity (which is always a Friday), and every week the options roll over to new maturities. For example, taking the second Tuesday in October, the near-term option expires in 24 days, and the next-term option in 31 days. One day later, the option that expires now in 30 days is the near-term option, and another option expiring in 37 days is the next-term option. This next-term option will, one week later, roll over to a near-term option and, one more week later, drop out of the calculation. Thus, an option can be included in the calculation for up to two weeks. Therefore, the regression is conducted with daily volatilities, but only for one value out of two weeks. As in Jiang, George J & Tian (ibid.), the values of Wednesday are used, for each second week.
The estimation results for the sample using non-overlapping data are summarized in 6 and 6.

5 Discussion

Similarity of the regression specifications is strong evidence

6 Appendix

Figures

Figure 1: SPX and VIX

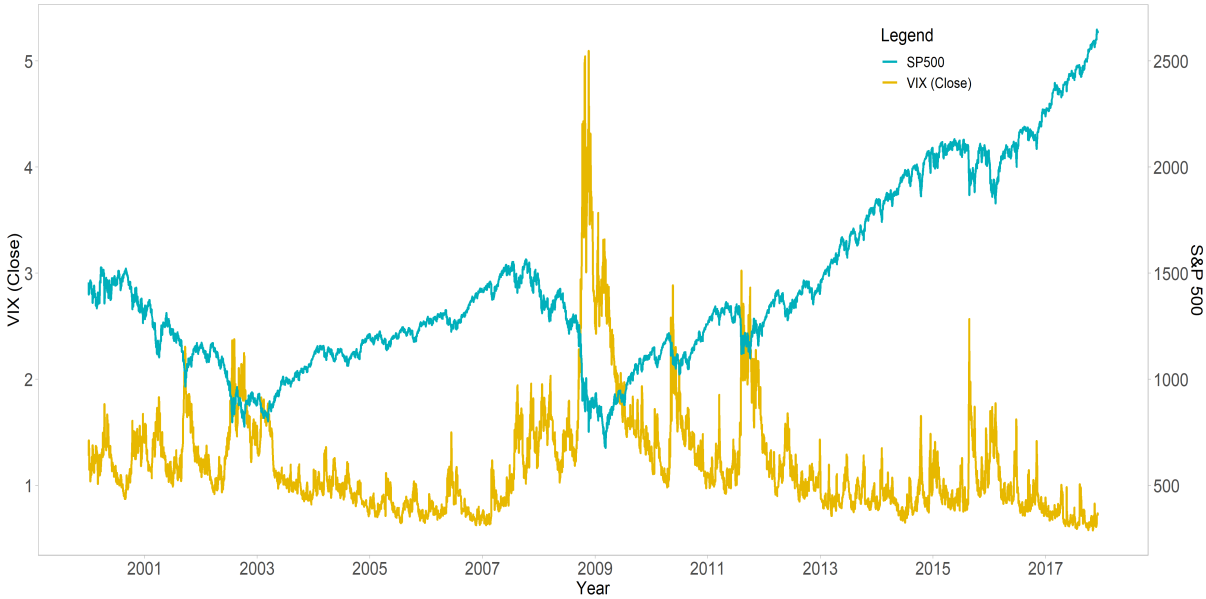
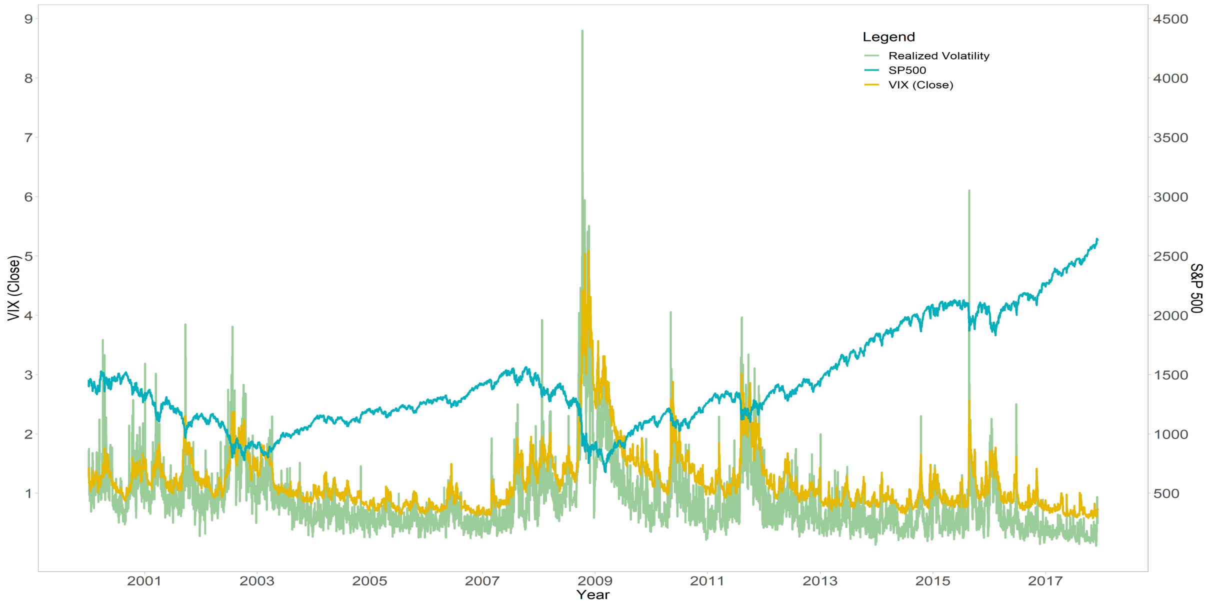


Figure 2: SPX and RV and VIX



Data Summary Statistics

Table 2: Summary statistics: Variables

Statistic	N	Min	Max	Mean	St. Dev.	Skewness	Kurtosis
All time periods							
RV	4,434	0.110	8.802	0.874	0.615	3.127	17.751
VIX	4,434	0.576	5.094	1.196	0.525	2.552	9.787
Weekly RV	4,434	0.183	5.586	0.874	0.556	2.757	12.451
Monthly RV	4,434	0.223	4.375	0.876	0.512	2.534	10.051
Weekly VIX	4,434	0.596	4.593	1.197	0.519	2.510	9.235
Monthly VIX	4,434	0.618	4.126	1.198	0.504	2.491	8.853
During crisis							
RV	1,252	0.213	8.802	1.171	0.834	2.713	11.399
VIX	1,259	0.847	5.094	1.623	0.695	1.941	4.232
Weekly RV	1,259	0.257	5.586	1.169	0.753	2.409	7.417
Monthly RV	1,259	0.423	4.375	1.172	0.689	2.178	5.448
Weekly VIX	1,259	0.885	4.593	1.623	0.684	1.885	3.753
Monthly VIX	1,259	0.946	4.126	1.625	0.662	1.837	3.311
Outside of crisis							
RV	3,229	0.110	6.109	0.762	0.454	2.252	10.798
VIX	3,251	0.576	2.566	1.034	0.307	1.206	1.365
Weekly RV	3,246	0.183	3.165	0.765	0.400	1.564	3.441
Monthly RV	3,229	0.223	2.367	0.764	0.363	1.308	1.687
Weekly VIX	3,246	0.596	2.188	1.034	0.300	1.157	1.047
Monthly VIX	3,229	0.618	2.014	1.033	0.285	1.1	0.78

Table 3: Summary statistics: Logarithm variables

Statistic	N	Min	Max	Mean	St. Dev.	Skewness	Kurtosis
All time periods							
RV	1,236	−9.664	0.777	−1.469	1.262	-1.369	3.185
VIX	2,528	−8.005	0.487	−1.562	1.190	-1.294	2.586
Weekly RV	4,434	0.183	5.586	0.874	0.556	2.757	12.451
Monthly RV	4,434	−1.501	1.476	−0.259	0.483	0.487	0.357
Weekly VIX	4,434	−0.518	1.525	0.111	0.350	0.933	1.117
Monthly VIX	4,434	−0.481	1.417	0.115	0.340	0.963	1.227
During crisis							
RV	1,252	0.213	8.802	1.171	0.834	2.713	11.399
VIX	1,259	0.847	5.094	1.623	0.695	1.941	4.232
Weekly RV	1,259	0.257	5.586	1.169	0.753	2.409	7.417
Monthly RV	1,259	0.423	4.375	1.172	0.689	2.178	5.448
Weekly VIX	1,259	0.885	4.593	1.623	0.684	1.885	3.753
Monthly VIX	1,259	0.946	4.126	1.625	0.662	1.837	3.311
Outside of crisis							
RV	3,229	0.110	6.109	0.762	0.454	2.252	10.798
VIX	3,251	0.576	2.566	1.034	0.307	1.206	1.365
Weekly RV	3,246	0.183	3.165	0.765	0.400	1.564	3.441
Monthly RV	3,229	0.223	2.367	0.764	0.363	1.308	1.687
Weekly VIX	3,246	0.596	2.188	1.034	0.300	1.157	1.047
Monthly VIX	3,229	0.618	2.014	1.033	0.285	1.100	0.780

Table 4: Correlation table

	RV	VIX	Daily RV	Weekly RV	Monthly RV	D. VIX	W.VIX	M. VIX
RV	1	0.837	0.806	0.823	0.769	0.819	0.781	0.698
VIX	0.837	1	0.822	0.883	0.903	0.981	0.969	0.926
Daily RV	0.806	0.822	1	0.890	0.794	0.837	0.803	0.710
Weekly RV	0.823	0.883	0.890	1	0.903	0.896	0.906	0.809
Monthly RV	0.769	0.903	0.794	0.903	1	0.910	0.932	0.931
Daily VIX	0.819	0.981	0.837	0.896	0.910	1	0.982	0.935
Weekly VIX	0.781	0.969	0.803	0.906	0.932	0.982	1	0.961
Monthly VIX	0.698	0.926	0.710	0.809	0.931	0.935	0.961	1

Regression Results with Robustness Checks

Table 5: Level regession (whole sample)

	Dependent variable:		
	Realized Volatility		
	Reg1a	Reg2a	Reg3a
	(1)	(2)	(3)
Intercept	0.045***	−0.324***	−0.169***
	(0.015)	(0.059)	(0.034)
RV_t^d	0.362***		0.256***
	(0.038)		(0.040)
RV_t^w	0.391***		0.286***
	(0.056)		(0.064)
RV_t^m	0.188***		−0.106**
	(0.036)		(0.050)
crisis	0.025*	−0.214***	−0.112***
	(0.013)	(0.035)	(0.021)
VIX_t		1.052***	0.579***
		(0.059)	(0.064)
AIC	2817.4	3104.2	2446
Observations	4,434	4,434	4,434
R^2	0.708	0.689	0.732
Adjusted R^2	0.708	0.688	0.732
Residual Std. Error	0.332 (df = 4429)	0.343 (df = 4431)	0.319 (df = 4428)
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 6: Level regression (non-overlapping sample)

	Dependent variable:		
	Realized Volatility		
	Reg1a	Reg2a	Reg3a
	(1)	(2)	(3)
Intercept	0.046	-0.342^{***}	-0.134^{**}
	(0.031)	(0.091)	(0.052)
RV_t^d	0.408^{***}		0.330^{***}
	(0.109)		(0.116)
RV_t^w	0.501^{***}		0.394^{***}
	(0.126)		(0.128)
RV_t^m	0.072		-0.132
	(0.079)		(0.083)
$crisis$	-0.022	-0.257^{***}	-0.131^{***}
	(0.029)	(0.058)	(0.035)
VIX_t		1.092^{***}	0.460^{***}
		(0.094)	(0.089)
AIC	192.5	281.8	166
Observations	456	456	456
R^2	0.757	0.701	0.771
Adjusted R^2	0.754	0.700	0.769
Residual Std. Error	0.297 (df = 451)	0.328 (df = 453)	0.288 (df = 450)
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 7: Logarithmic regression (non-overlapping sample)

	Dependent variable:		
	Realized Volatility		
	Reg1b	Reg2b	Reg3b
	(1)	(2)	(3)
Intercept	-0.001	-0.365^{***}	-0.156^{***}
	(0.018)	(0.022)	(0.029)
$ln(RV_t^d)$	0.346^{***}		0.262^{***}
	(0.064)		(0.062)
$ln(RV_t^w)$	0.408^{***}		0.289^{***}
	(0.078)		(0.076)
$ln(RV_t^m)$	0.171^{***}		-0.034
	(0.054)		(0.064)
crisis	-0.017	-0.266^{***}	-0.148^{***}
	(0.027)	(0.046)	(0.034)
$ln(VIX_t)$		1.478^{***}	0.700^{***}
		(0.060)	(0.104)
AIC	126.4	173.4	81.8
Observations	456	456	456
R^2	0.748	0.718	0.773
Adjusted R^2	0.746	0.717	0.770
Residual Std. Error	0.276 (df = 451)	0.291 (df = 453)	0.262 (df = 450)
Note:			*p<0.1; **p<0.05; ***p<0.01

F-test Results with Robustness Checks

Table 8: F-test Reg3a

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	4432	524.31				
2	4428	449.29	4	75.03	184.86	0.0000

Table 9: F-test Reg3b

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	4432	529.46				
2	4428	367.22	4	162.24	489.09	0.0000

Table 10: F-test Reg3a non-overlapping sample

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	4432	524.31				
2	4428	449.29	4	75.03	184.86	0.0000

Table 11: F-test Reg3b non-overlapping sample

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	4432	529.46				
2	4428	367.22	4	162.24	489.09	0.0000

References

Andersen, Torben G. and T. Bollerslev (1998). 'Answering the Critics: Yes, ARCH Models Do Provide Good Volatility Forecasts'. In: *International Economic Review* 39, pp. 885–905.

Andersen, Torben G., Tim Bollerslev et al. (2003). 'Modeling and forecasting realized volatility'. In: *Econometrica* 71.2, pp. 579–625.

Andersen, Torben G, Tim Bollerslev and Francis X Diebold (2009). 'Parametric and nonparametric volatility measurement'. In: *Handbook of Financial Econometrics: Tools and Techniques*. Ed. by Yacine Ait-Sahalia and Lars Hansen. Elsevier, pp. 67–137.

Bakanova, Asyl (2010). 'The information content of implied volatility in the crude oil market'. In: *Proceedings of the Brunel conference*. Citeseer, pp. 1–19.

Black, Fischer and Myron Scholes (1973). 'The pricing of options and corporate liabilities'. In: *Journal of political economy* 81.3, pp. 637–654.

Blair, Bevan J, Ser-Huang Poon and Stephen J Taylor (2001). 'Modelling S&P 100 volatility: The information content of stock returns'. In: *Journal of banking & finance* 25.9, pp. 1665–1679.

Britten-Jones, Mark and Anthony Neuberger (2000). 'Option prices, implied price processes, and stochastic volatility'. In: *The Journal of Finance* 55.2, pp. 839–866.

Canina, Linda and Stephen Figlewski (1993). 'The informational content of implied volatility'. In: *The Review of Financial Studies* 6.3, pp. 659–681.

Chicago Board of Options Exchange (n.d.[a]). *Historic daily prices, year = [2018*. `http://www.cboe.com/products/stock-index-options-spx-rut-msci-ftse/s-p-500-index-options/s-p-500-index/spx-historical-data`.

— (n.d.[b]). *Vix Data for 1990 - 2003, year = [2018.* `http://www.cboe.com/products/vix-index-volatility/vix-options-and-futures/vix-index/vix-historical-data`.

— (n.d.[c]). *VIX Data for 2004 to present, year = [2018.* `http://www.cboe.com/products/vix-index-volatility/vix-options-and-futures/vix-index/vix-historical-data`.

— (2009). *White Paper Cboe Volatility Index.* Tech. rep. Cboe Exchange, pp. 1–23.

Christensen, Bent J and Nagpurnanand R Prabhala (1998). 'The relation between implied and realized volatility1'. In: *Journal of financial economics* 50.2, pp. 125–150.

Christensen, Bent Jesper, Charlotte Strunk Hansen and Nagpurnanand R Prabhala (2001). 'The telescoping overlap problem in options data'. In: *Draft Paper*.

Corsi, Fulvio (2009). 'A Simple Approximate Long-Memory Model of Realized Volatility'. In: *Journal of Financial Econometrics* 7.2, pp. 174–196.

Day, Theodore E. and Craig M. Lewis (1992). 'Stock market volatility and the information content of stock index options'. In: *Journal of Econometrics* 52.1-2, pp. 267–287.

Derman, Emanuel and Iraj Kani (1994). 'Riding on a smile'. In: *Risk* 7.2, pp. 32–39.

Dupire, Bruno (1997). 'Pricing and hedging with smiles'. In: *Mathematics of derivative securities* 1.1, pp. 103–111.

Dupire, Bruno et al. (1994). 'Pricing with a smile'. In: *Risk* 7.1, pp. 18–20.

Fleming, Jeff, Barbara Ostdiek and Robert E. Whaley (1995). 'Predicting Stock Market Volatility: A New Measure'. In: *The Journal of Futures Markets* 15.3, pp. 265–302.

Jiang, George J & Tian, Yisong S. (2003). 'Model-Free Implied Volatility and Its Information Content'. In: *The Review of Financial Studies* 18.4.

Jorion, Philippe (1995). 'Predicting volatility in the foreign exchange market'. In: *The Journal of Finance* 50.2, pp. 507–528.

Lamoureux, Christopher G and William D Lastrapes (1993). 'Forecasting stock-return variance: Toward an understanding of stochastic implied volatilities'. In: *The Review of Financial Studies* 6.2, pp. 293–326.

Malkiel, Burton G and Eugene F Fama (1970). 'Efficient capital markets: A review of theory and empirical work'. In: *The journal of Finance* 25.2, pp. 383–417.

Mandelbrot, Benoit B (1963). 'The variation of certain speculative prices'. In: *The Journal of Business* 36, pp. 394–419.

Müller, Ulrich A et al. (1993). 'Fractals and intrinsic time: A challenge to econometricians'. In: *Unpublished manuscript, Olsen & Associates, Zürich.*

Oxford-Man Institute's realized library V0.2 (2009). Tech. rep. [Online; accessed 31-October-2018].

Poon, Ser-huang and Clive W.J. Granger (2003). 'Forecasting volatility in financial markets : A review'. In: *Journal of Economic Literature* 41.June, pp. 478–539.

Rubinstein, Mark (1994). 'Implied binomial trees'. In: *The Journal of Finance* 49.3, pp. 771–818.

Taylor, Stephen J, Pradeep K Yadav and Yuanyuan Zhang (2010). 'The information content of implied volatilities and model-free volatility expectations: Evidence from options written on individual stocks'. In: *Journal of Banking & Finance* 34.4, pp. 871–881.

Tsay, Ruey S (2005). *Analysis of financial time series.* Vol. 543. John Wiley & Sons.

Whaley, Robert E (2008). *Understanding VIX.* accessed 14.01.2019.

Declaration of Authorship

I hereby declare that the paper

"The Information Content of VIX Volatility"

is my own unaided work. All direct or indirect sources used are acknowledged as references. I am aware that the paper in digital form can be examined for the use of unauthorized aid and in order to determine whether the thesis as a whole or parts incorporated in it may be deemed as plagiarism. This paper was not previously presented to another examination board and has not been published.

Place, Date　　　　author's signature