Volatility and variance swaps

A comparison of quantitative models to calculate the fair volatility and variance strike

Johan Röring
June 8, 2017

Student
Master's Thesis, 30 Credits
Department of Mathematics
and Mathematical Statistics
Umeå University

Volatility and variance swaps
A comparison of quantitative models to calculate the fair volatility and variance strike

Johan Röring

Submitted in partial fulfillment of the requirements for the degree Master of Science in Industrial Engineering and Management with specialization in Risk Management

Department of Mathematics and Mathematical Statistics.
Umeå University

Supervisor: Åke Brännström
Examiner: Mats G Larson

Abstract

Volatility is a common risk measure in the field of finance that describes the magnitude of an asset's up and down movement. From only being a risk measure, volatility has become an asset class of its own and volatility derivatives enable traders to get an isolated exposure to an asset's volatility. Two kinds of volatility derivatives are volatility swaps and variance swaps.

The problem with volatility swaps and variance swaps is that they require estimations of the future variance and volatility, which are used as the strike price for a contract. This thesis will manage that difficulty and estimate strike prices with several different models. I will describe how the variance strike for a variance swap can be estimated with a theoretical replicating scheme and how the result can be manipulated to obtain the volatility strike, which is a technique that require Laplace transformations. The famous Black-Scholes model is described and how it can be used to estimate a volatility strike for volatility swaps. A new model that uses the Greeks vanna and vomma is described and put to the test. The thesis will also cover a couple of stochastic volatility models, Exponentially Weighted Moving Average (EWMA) and Generalized Autoregressive Conditional Heteroskedasticity (GARCH).

The models' estimations are compared to the realized volatility. A comparison of the models' performance over 2015 is made as well as a more extensive backtesting for Black-Scholes, EWMA and GARCH.

The GARCH model performs the best in the comparison and the model that uses vanna and vomma gives a good result. However, because of limited data, one can not fully conclude that the model that uses vanna and vomma can be used when calculating the fair volatility strike for a volatility swap.

Sammanfattning

Volatilitet är ett vanligt riskmått i finansbranschen som beskriver storleken på en tillgångs upp- och nedgångar i pris. Från att enbart vara ett riskmått så har volatilitet blivit ett eget tillgångsslag med volatilitetsderivat som möjliggör för investerare att få en isolerad exponering mot en tillgångs volatilitet. Två typer av volatilitetsderivat är volatilitesswappar och variansswappar.

Svårigheten med volatilitets- och variansswappar är hur strikepriset för dem ska beräknas. Den här uppsatsen hanterar den svårigheten och beräknar strikepriser med olika modeller. Jag kommer först undersöka en teoretiskt replikeringsmetod för att bestämma strikepriset för en variansswap och hur strikepriset för en volatilitetsswap kan tas fram från resultatet, en teknik som kräver Laplacetransformationer. Den kända modellen Black-Scholes beskrivs och hur den kan avändas till att estimera strikepriser för volatilitetsswappar. En helt ny modell som använder sig av vanna och vomma, greker från Black-Scholes modell, beskrivs och testas. Uppsatsen täcker även in de två stokastiska volatilitetsmodellerna Exponentially Weighted Moving Average (EWMA) och Generalized Autoregressive Conditional Heteroskedasticity (GARCH).

Modellernas volatilitetsestimat jämförs med den realiserade volatiliteten. En jämförelese mellan modellernas resultat över data från 2015 är gjord. För Black-Scholes, EWMA och GARCH så innehåller resultatet även en lång backtesting.

GARCH-modellen presterar bäst under jämförelsen och modellen som använder sig av vanna och vomma ger ett bra resultat. På grund av begränsningar i mängden data så kan det inte säkerställas till fullo att modellen med vanna och vomma fungerar när strikepriset för en volatilitetsswap ska beräknas.

Acknowledgements

First of all, I would like to express my greatest gratitude to my supervisor, Åke Brännström, at the department of Mathematics and Mathematical Statistics at Umeå University. His invested time with comments, support and clarifications regarding this thesis has been vital for completing it. I would also like to thank Markus Ådahl for guiding me into what led to this project and for his courses in the Risk Management specialization, which gave me the foundations required for this thesis. Thank you!

Contents

1 Introduction **1**
1.1 Background . . . 1
1.2 Problem statement . . . 2
1.3 Approach . . . 2
1.4 Outline . . . 3

2 Volatility and variance swaps **4**
2.1 What is volatility and variance? . . . 4
2.1.1 Volatility clustering . . . 5
2.2 Swaps . . . 7
2.2.1 A variance swap contract . . . 7
2.2.2 A volatility swap contract . . . 8
2.3 Usage of volatility and variance swaps . . . 8
2.3.1 Speculation . . . 9
2.3.2 Hedging . . . 9

3 Valuation using deterministic volatility models **10**
3.1 Deriving the fair variance strike with a replication scheme . . . 10
3.1.1 Assumptions . . . 10
3.1.2 Replication with a discrete set of options . . . 12
3.1.3 Derivation of the payoff function and how it can be approximated . . . 12
3.1.4 Numerical example . . . 15
3.2 Deriving the fair volatility strike using Laplace transformations . . . 16
3.3 Black-Scholes model and implied volatility . . . 18
3.3.1 Assumptions . . . 18
3.3.2 The Black-Scholes equation and formulas . . . 19
3.3.3 Volatility smile . . . 20
3.4 Vanna-Vomma model . . . 21
3.4.1 Greeks . . . 21
3.4.2 Approach . . . 22
3.4.3 Motivation of the model . . . 23

4 Valuation using stochastic volatility models **25**
4.1 Exponentially Weighted Moving Average . . . 25
4.1.1 Derivation of the model . . . 25
4.2 Generalized autoregressive conditional heteroskedasticity . . . 27

5 Method **28**
5.1 Data . . . 28
5.2 Implementation . . . 28
5.2.1 Black-Scholes model . . . 28
5.2.2 Vanna-Vomma model . . . 28
5.2.3 EWMA . . . 29
5.2.4 GARCH . . . 29

6 Results **30**
6.1 Comparison of the models . . . 30
6.2 Backtesting . . . 31

7 Discussion **33**
7.1 Conclusions and discussions of the results 33
7.2 Limitations and extensions 34
7.3 Outlook for volatility derivatives 35

Appendices **38**
A Tables of the comparison of the models from done by data from year 2015 . . . 38
B Backtesting for BS, EWMA and GARCH from 2008 when T=6 months 39
C Backtesting for EWMA and GARCH from 1995. 40

1 Introduction

This introductory chapter first explains the background of the thesis in Section 1.1. Section 1.2 explains why volatility derivatives exists and what models that will be used when valuing volatility and variance swaps. A problem statement is formulated in Section 1.3 as well as an approach for completing the thesis.

1.1 Background

The financial markets have evolved significantly over the last decades. It now consists of many complex derivatives such as exotic options, swaps, warrants and futures (Hull, 2012). During the 2008 financial crisis, shortcomings in risk management and the supervision of banks and financial institutions were exposed. As a result of poor risk management policies and ineffective supervision, the investment bank Lehman Brothers went bankrupt and caused instability across the global financial system (Adu-Gyamfi, 2016). One key factor of a solid risk management system for a financial institution is to value the derivatives in their portfolios accurately.

Volatility is a common risk measure in the field of finance that describes the magnitude of an asset's up and down movement. It is measured as the standard deviation of logarithmic returns and the variance is simply the variance of the returns, or volatility squared. From only being a risk measure, volatility has become an asset class of its own and volatility derivatives enable traders to get an isolated exposure to an asset's volatility. Volatility derivatives is a way for traders to generate profits by speculating on future realized volatility/variance of an asset if they sense to know something about the near future. As shown later, volatility derivatives can also be used in a hedging strategy for avoiding future losses.

A difficulty with volatility swaps and variance swaps is how to calculate the fair volatility/variance strike. This thesis will manage that difficulty on how to accurately calculate the volatility strike and variance strike. The calculations can be done using several different models. I will first look at a theoretical replicating scheme for estimating the variance strike for a variance swap. In theory, the replication require an infinite number of European put and call options but as later shown, the replication can be approximated with a finite number of options. How the conversion from obtaining the volatility strike from an estimated variance strike follows from the replicating scheme, which is a technique that require Laplace transformations because of a convexity error in the payoff for the variance swap. The famous Black-Scholes model which can be used to generate a closed-form formula for options is described and how it can be used together with market prices of European options to estimate a volatility strike for volatility swaps. The thesis will also cover a couple of stochastic volatility models, Exponentially Weighted Moving Average (EWMA) and Generalized Autoregressive Conditional Heteroskedasticity (GARCH). When evaluating these models, they are backtested for a long period of time.

Recently, a new model has come to life that uses vanna and vomma, which are Greeks from the Black-Scholes model, to approximate the fair volatility strike. The model is brand new which is why is does not even have a name yet, but it will hereafter in the thesis be referred to as the Vanna-Vomma model (VV model for short). Because it is new, the accuracy of this model has not been put to the test, besides from a comparison to the stochastic volatility model Heston, in the article that describes it (Rolloos & Arslan, 2017). One goal of this thesis is therefore to conclude if the model can be used or not when calculating the volatility strike for a volatility swap. A summary of the different models and a brief explanation of them are described in Table

1 below.

To the extent of the author's knowledge, there are no studies that makes a thorough comparison of the models. Comparing the models' performances can be complicated because of the different ways of measuring them. The estimated fair volatility/variance can be compared to the realized volatility/variance systematically over a time period with market data. The result can be used to observe which model that has the lowest mean deviation from the realized volatility/variance. Another measure is to investigate the distribution of the deviations and find which model that has the smallest extreme outcomes, i.e. the distribution with the thinnest tail. How user friendly the model is, measured in implementation difficulty and time complexity, will also effect the model's overall valuation. The purpose of this thesis is to, beside from examine the new Vanna-Vomma model, apply and evaluate several models for calculating the fair volatility/variance strike and test their performance based on the several measures.

Table 1: *Overview of the models.*

Model	**Description**	**Ref.**
Black-Scholes (BS)	Estimates the fair volatility/variance strike from the implied volatility of option prices.	1,2,3
Replicating scheme	Estimates the fair variance strike with a discrete set of European options.	4,5
Vanna-Vomma model (VV)	Uses the greeks Vanna and Vomma from the BS model to derive a formula to approximate the fair volatility strike.	6
EWMA	Uses historical returns to forecast volatility/variance. Gives recent observations greater weight when forecasting and the weights descends exponentially.	7,8
GARCH	Uses historical returns to forecast volatility/variance. It is an autoregressive model, ie. it depends on its own previous values.	7,8

1: Black & Scholes, 1973. 2: Flemming, 1998. 3: Christensen & Prabhala, 1998. 4: Demeterfi et. al, 1999. 5: Broadie & Jain, 2008. 6: Rolloos & Arslan, 2017. 7: Danielsson, 2011. 8: Alexander, 2008.

1.2 Problem statement

- Which model has the best performance when calculating the fair variance/volatility strike?
- Can the VV-model be used when calculating the fair volatility strike for a volatility swap?

1.3 Approach

When evaluating the models, the comparisons are done with the models' estimations of volatility and realized volatility. Each model is implemented to calculate volatility strikes and the estimations are then compared to realized volatility of historical data. The calculations in the VV-model are made from European put and call options that are written on the Standard & Poor's 500 stock index. The parameter estimations for EWMA and GARCH are made from daily returns from the same stock index. The realized volatility is calculated from Standard & Poor's 500. The parameter estimations for all formulas, as well as the EWMA and GARCH simulations are made in Matlab.

The models' performance is measured on the mean value of the differences from the estimations and the realized volatility. The standard deviation of the differences is another measure that is used and compared between the models. The tails of the distribution of the differences between estimations and realized volatility is measured by the the 1st and 99th percentile of the differences. The 1st percentile is denoted as the lower tail and the 99th percentile is denoted as the upper tail.

1.4 Outline

The rest of the thesis that follows from this introduction begin with a section that defines volatility and variance together with a description of volatility swaps and variance swaps. The description of volatility and variance swaps explains how they are structured and why financial institutions trade them.

Two sections that describe the different models are presented afterwards in Section 3 and Section 4, where Section 3 describe deterministic volatility models and Section 4 describes stochastic volatility models. What data that is used and how the models are implemented are explained in Section 5. The results when the models' estimations are compared to the realized volatility are presented in Section 6.

Ending this thesis, some discussions of the result are presented in Section 7, which covers conclusions, limitations and possible extensions and finally an outlook for volatility derivatives.

2 Volatility and variance swaps

This chapter gives the reader an overview of volatility and variance swaps. Section 2.1 defines volatility and variance and describes how they are calculated. The phenomenon known as volatility clustering is explained and proven by using historical returns of the Standard & Poor's 500. With the proven fact that volatility moves in clusters, it is explained that the volatility in the Standard & Poor's 500, and other stock indexes, can be slightly predictable.

An overview of swaps in general is described in section 2.2 together with the structure and the components for a variance swap and a volatility swap. In the end of this section, the reader gets explanations of the motives to trade volatility and variance swaps.

2.1 What is volatility and variance?

Volatility is often used as a risk measure for an asset. An asset with high volatility has larger movements of the return compared to an asset with lower volatility, and is therefore riskier to hold as an investor. Working with a discrete sample of n observations in asset prices, volatility is defined as the standard deviation of the logarithmic returns with the assumption that the average daily return is zero.

We are interested in calculating a realized volatility for an underlying asset of a swap contract with maturity T years. The discrete annualized volatility is denoted as $\sigma_{\text{d}}(0, T, n)$, where the subscript indicates that the sampling is discrete. To calculate the annual realized volatility over the interval $[0, T]$ with n observations with equal length, the following formula can be used

$$\sigma_{\text{d}}(0, T, n) = \sqrt{\frac{AF}{n-1} \sum_{i=0}^{n-1} \left(\log \left(\frac{S_{i+1}}{S_i} \right) \right)^2}, \tag{1}$$

where S_i is the price of the asset at time i. AF is an annualization factor and is defined as n/T. It has the purpose to make the calculated realized volatility measured as an annual volatility. For instance, when calculating the realized volatility with daily observations for a volatility swap with a maturity of one year, the annualization factor is 252, as the number of trading days in one year is 252.

Another way of calculating the volatility is to view it as a continuous sample of fluctuations. The continuous volatility is often used as a way of describing the realized volatility for an asset in a swap contract (Brockhaus & Long, 2000). The continuous volatility over $[0, T]$ is

$$\sigma_{\text{c}}(0, T) = \sqrt{\frac{1}{T} \int_0^T \sigma_u^2 \text{d}u}. \tag{2}$$

The discrete volatility approaches the continuous volatility as the number of observations, n, approaches infinity (Broadie & Jain, 2008)

$$\sigma_{\text{c}}(0, T) = \lim_{n \to \infty} \sigma_{\text{d}}(0, T, n). \tag{3}$$

Variance is another statistical measure of how much the asset's returns deviate from its mean, and it is the squared volatility. The formula for calculating the realized variance for a variance swap with maturity T,

$$\sigma_{\mathrm{d}}^2(0,T,n) = \frac{AF}{n-1}\sum_{i=0}^{n-1}\left(\log\left(\frac{S_{i+1}}{S_i}\right)\right)^2, \tag{4}$$

is Equation (1) squared.

Variance can also take the form as a continuous sample of fluctuations and is often used as a way of describing the realized variance for an asset in swap contract (Brockhaus & Long, 2000). For a variance swap with maturity T, the continuous realized variance of the underlying asset is denoted by $\sigma_{\mathrm{c}}^2(0,T)$ and defined as

$$\sigma_{\mathrm{c}}^2(0,T) = \frac{1}{T}\int_0^T \sigma_u^2 \mathrm{d}u. \tag{5}$$

2.1.1 Volatility clustering

Research of financial returns has shown an interesting fact regarding its volatility. One characteristic that can be seen in most financial returns is that the volatilities of financial returns tends to cluster together. This phenomenon is called *Volatility clustering* and is one of the stylized facts of financial returns (Danielsson, 2011).

The Chicago Board Options Exchange (CBOE) manage a Volatility Index, denoted VIX, that measure the market's expectation of volatility over a 30 day period from observed option prices. It is a widely used measure of market risk and is also known as the ”investor fear gauge”, ”fear index” or ”risk index”. Since its introduction in 1993, it has increased the interest of volatility derivatives and CBOE now offer as much as 25 different volatility products for investors to trade.

Studying Figure 1, which illustrate the evolution of the VIX from 1990 to 2016, the volatility clustering can easily be observed. During the years between 1991 and prior to the burst of the dot-com bubble in 1999, the volatility levels stayed fairly low, but then increased and had a constant higher level for the following four years approximately. Afterwards, another period of low volatility took place during 2003 to 2007. The following years showed an enormous spike in the volatility levels during the financial crisis, which was the start of another cluster with higher volatility.

Another way of illustrating volatility clusters is by using an autocorrelation function (ACF) on the returns. The ACF measures how correlated a one day return is with returns from previous days. Volatility does not consider if the returns are negative of positive. To measure the autocorrelation independently of the direction of the return, the squared returns can be investigated instead and measure if they are correlated with previous squared returns. Figure 2 shows an ACF plot of daily S&P500 squared returns with lags (number of previous days) on the x-axis. The figure clearly shows how the squared return have a correlation to the squared returns that occurred in the recent days. A high volatility today will most likely result in a high volatility tomorrow and if the volatility is low today, it is likely to be low tomorrow. The correlation decreases exponentially with the number of lags. For example, the volatility today will correlate with yesterday's volatility but it will have no correlation with the volatility 1000 days ago.

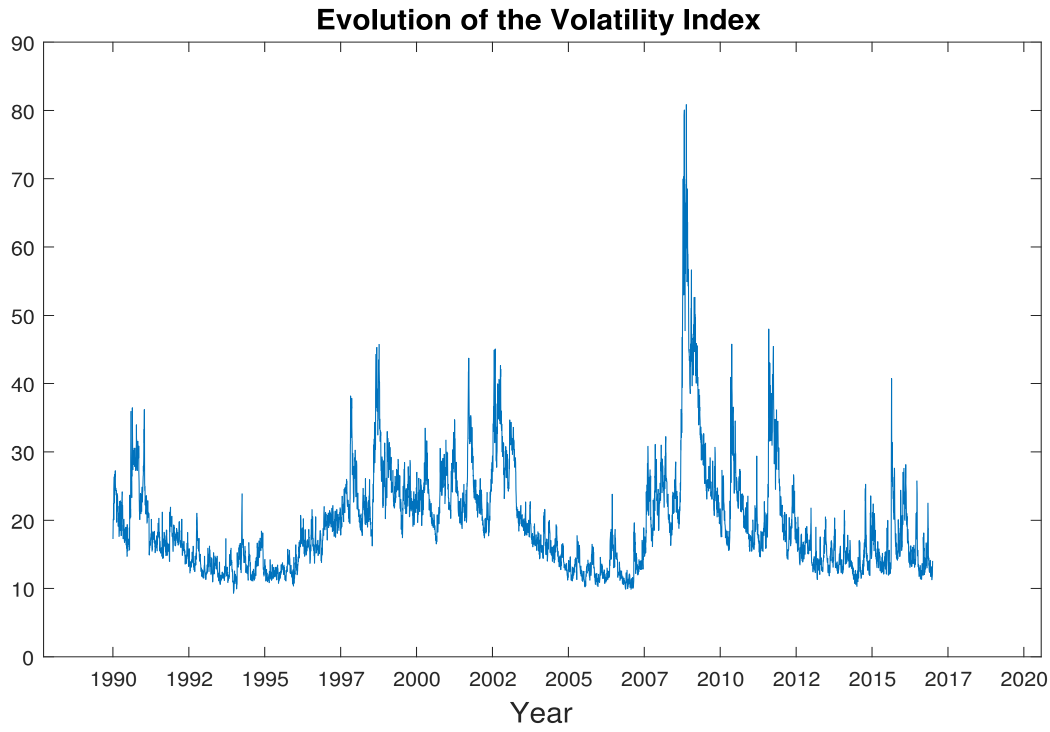

Figure 1: *Evolution of the Volatility Index. The data was retrieved from Yahoo Finance on 2017-03-10.*

Volatility clusters implies that the volatility in the near future is slightly predictable as it autocorrelates with a few number of lags. This is a fact that a lot of models for forecasting volatility takes into account and will be important when valuing volatility and variance swaps.

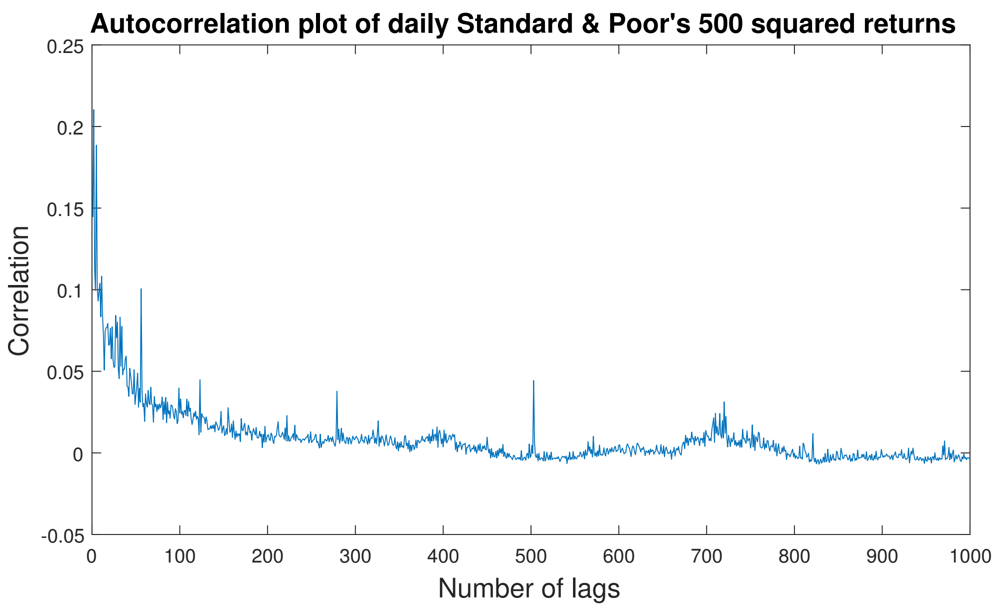

Figure 2: *Autocorrelation plot of daily Standard & Poor's 500 squared returns from 1950 to 2016. The data was retrieved from Yahoo Finance on 2017-03-10.*

2.2 Swaps

A swap is an agreement between two parties to exchange cash flows in the future. The first swap contract was introduced in the beginning of the 1980s and the market for swaps has since then grown rapidly (Hull, 2012). In the financial industry, there exist a lot of different swaps but the two most common are interest rate swaps and currency swaps. There are a couple of variants of interest rate swaps but the most common are the plain vanilla interest rate swap (Hull, 2012). In a plain vanilla interest rate swap one of the parties agrees to pay a cash flow that is equal to a predetermined fixed rate of a notional amount and in return, it receives interest at a floating rate of the same notional amount. The floating rate is usually the London Interbank Offered Rate (LIBOR). In a currency swap, the parties exchange cash flows in different currencies. Other examples of swap contracts are credit default swaps, currency swaps, compounding swaps and equity swaps (Hull, 2012). The development of new swap contracts is described by John C. Hull as:

"Swaps are limited only by the imagination of financial engineers and the desire of corporate treasurers and fund managers for exotic structures" (2012, page 175).

Variance swaps and volatility swaps are thus only two varieties in a wide spectrum of swaps. Swaps are traded Over-The-Counter and are categorized as OTC derivatives, which implies that they are traded between financial institutions or companies (Hull, 2012). Variance swaps and volatility swaps can be written on single stocks, stock indexes or on other assets.

2.2.1 A variance swap contract

A variance swap contract consist of three main parts, the realized variance, denoted $\sigma_{\text{d}}^2(0, T, n)$, the fair variance strike, denoted K_{var}, and a notional amount, denoted N_{var}. The notional amount is agreed by the two counterparties when entering a swap. When trading variance swaps, it is common to define the notional amount in terms of volatility that is expressed as a vega notional. The vega notional is the profit or loss for every 1% change in volatility (Bossu, Strasser & Guichard, 2005). The notional amount for a variance swap is,

$$N_{\text{var}} = \frac{N_{\text{vega}}}{2 \times \sqrt{K_{\text{var}}}}. \tag{6}$$

The realized variance is described in Section 2.1 and in Equation (4). It is the variance that has occurred in the assets returns on the interval $[0, T]$ or during the lifespan of the contract. The fair variance strike is predetermined in the beginning of the contract. It is set to be equal to the expected future realized variance over the interval $[0, T]$. Assuming that the variance is calculated discretely, the variance strike is chosen such that

$$K_{\text{var}} = E_0[\sigma_{\text{d}}^2(0, T, n)]. \tag{7}$$

The fair variance strike is also commonly referred to as the variance strike price, despite the fact that it is a level of variance and not a price. The strike price and the realized variance are both quoted in annual terms. It is also common that the variance strike price is quoted as a volatility level squared, $K_{\text{var}} = (25\%)^2$ for example (Demeterfi et al., 1999).

The two counterparties exchange cash flows at the end of the contract, as illustrated in Figure 3. Counterparty 1 pays the notional amount multiplied with the variance strike price and receives the notional amount multiplied with the realized variance from Counterparty 2. Thus, the payoff

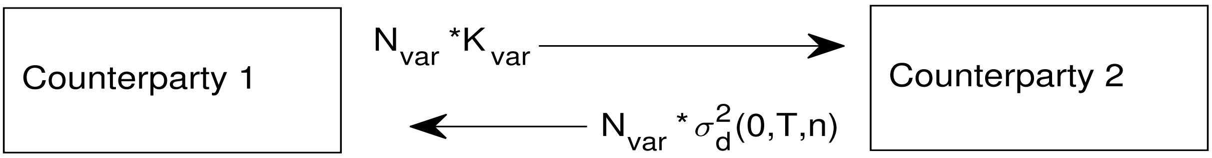

Figure 3: *Illustration of the exchange of cash flows for the two counterparties in a variance swap*

of the variance swap at maturity is, for Counterparty 1 in Figure 3, the notional amount multiplied with the difference between the variance strike and the realized variance. That means that the expected payoff of a variance swap is zero at initiation

$$\text{payoff} = N_{\text{var}} \times (\sigma_{\text{d}}^2(0, T, n) - K_{\text{var}}). \tag{8}$$

2.2.2 A volatility swap contract

The structure of a volatility swap is very similar to a variance swap. It also have a notional amount, where the notional is expressed as the vega amount N_{vega}, but the volatility swap uses the realized volatility over the interval $[0, T]$ instead of using the realized variance. How to calculate the realized volatility is described in Section 2.1 and in Equation (1).

The fair volatility strike, K_{vol}, is chosen in the same way as for the variance swap. It is set to be equal to the expected future realized volatility over $[0, T]$ (Rolloos & Arslan, 2017),

$$K_{\text{vol}} = E_0[\sigma_{\text{d}}(0, T, n)]. \tag{9}$$

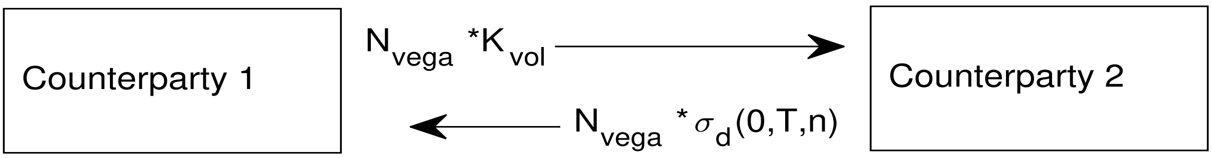

Figure 4: *Illustration of the exchange of cash flows for the two counterparties in a volatility swap*

The two counterparties in a volatility swap exchange cash flows in the same way as for the variance swap. The cash flows that occur in a volatility swap is illustrated in Figure 4. For Counterparty 1 in Figure 4, the volatility swap has the following payoff at expiry date T,

$$\text{payoff} = N_{\text{vega}} \times (\sigma_{\text{d}}(0, T, n) - K_{\text{vol}}). \tag{10}$$

2.3 Usage of volatility and variance swaps

Why trade with volatility? Stock investors trade stocks when they think that they know the direction of the stock market or of the individual stocks. Bond investors act in the same way, as they believe to know the direction of future interest rates. Derivatives have been invented as tools to generate extra profits, but also as a way of protecting capital and hedging of portfolios (Hull, 2012). Volatility and variance swaps are no different than other derivatives and enables financial institutions and banks to speculate on future volatility or variance and to hedge their portfolios to protect capital from losses.

2.3.1 Speculation

Volatility traders may have some idea of what the future volatility levels will be and can therefore use volatility swaps and variance swaps to speculate and generate profits. The investor can also believe that the current volatility levels, or that the expectation of future volatility, are incorrect making volatility swaps a good way to make money on that error.

Happenings like the release of companies' annual reports, upcoming elections and other political situations are common examples that may result in increasing volatility in the financial markets.

2.3.2 Hedging

Having volatility derivatives in a portfolio is a good diversification strategy to reduce the market risk. During financial turmoil and difficult times in the finance industry, the volatility levels tends to increase. Volatility and financial stock returns are therefore negatively correlated, which makes usage of volatility swaps and variance swaps a good way to reduce losses and for protection of capital. Figure 5 illustrates the evolution of the Volatility Index (VIX) together with the Standard & Poor's 500 stock index (S&P500). The negative correlation is especially noticeable during the 2008 financial crisis, where the value of the S&P500 fell substantially and the level of the VIX rose to an all time high.

It might be very difficult to know beforehand that a market crash is about to emerge. However, investors who could sense that the financial crisis was coming their way and held a long position in some volatility swaps and/or variance swaps, would have reduced their losses significantly.

Another way of reducing losses in a market crash is to by put options. The upside of using volatility swaps or variance swaps rather than put options is that if the market instead rises, the volatility and variance swaps can still generate a profit but the put options will be worth zero.

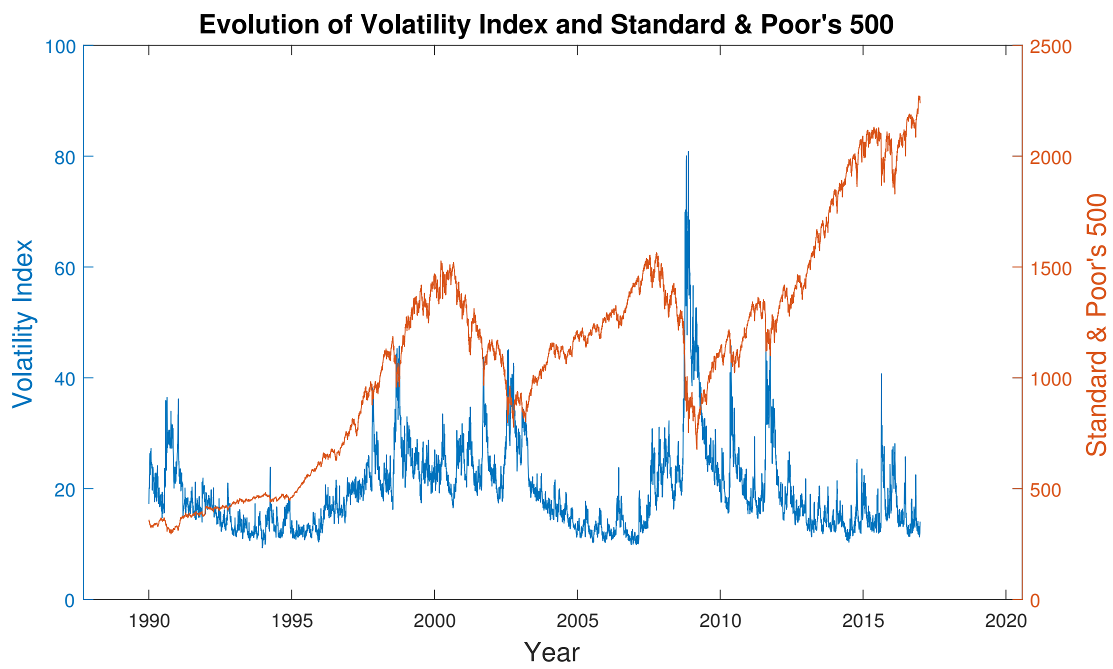

Figure 5: *Evolution of the Volatility Index and Standard & Poor's 500. The data was retrieved from Yahoo Finance on 2017-03-10.*

3 Valuation using deterministic volatility models

This section covers deterministic volatility models for calculating the variance strike for a variance swap and for calculating the volatility strike for a volatility swap. How the variance strike can be approximated by a portfolio of a discrete set European options is described in Section 3.1 and Section 3.2 describes a way to approximate the volatility strike from a variance strike, a technique that require Laplace transformations.

Section 3.3 covers the Black-Scholes model. The model is firstly described together with the assumptions that are made in the model. With the assumptions and model in place, an equation for pricing European call and put options are derived, which is used for calculating the implied volatility from option prices in the market. The implied volatility is an expectation of future volatility. The connection to the volatility indexes VIX, VXV and VXMT provided by CBOE are also described in this section as well as how they can be used as estimations of volatility.

The Vanna-Vomma model is closing this section. Vanna and vomma are Greeks from the Black-Scholes model and are defined in Subsection 3.4.1. Some flaws in the Black-Scholes model, regarding the implied volatility, is presented and how the Vanna-Vomma model handles that problem. The approach to approximate the volatility strike using vanna and vomma is described in Subsection 3.4.2 and the motivation of why it works theoretically is presented in Section 3.4.3.

3.1 Deriving the fair variance strike with a replication scheme

3.1.1 Assumptions

A variance swap contract is straightforward and its payoff is simple to understand with its three parts. The notional amount does not require any calculations since it is only a number that the counterparties agrees to and the realized variance can be calculated with Equation (4). The difficulty is how to accurately calculate the variance strike, K_{var}. Deriving the fair variance strike can be made without complex models by using a replicating portfolio scheme. The fair variance strike can in that way be determined with a portfolio that replicates the variance swap contract. (Demeterfi et al., 1999).

The replicating portfolio must be the same as the variance strike by arguments of an arbitrage free market. To replicate the swap, the portfolio need to consist of a static long position in a forward contract on the underlying asset and short position in a log-contract that is dynamically re-hedged. A log contract is a theoretical exotic option that depends on the logarithm of the underlying asset's price. The log contract is not traded, but its payoff can be replicated using a range of European call and put options with different strike prices (Demeterfi et al., 1999).

An assumption about the underlying asset of the replicating portfolio has to be made. The assumption is that the asset has similar characteristics to a Geometric Brownian Motion (GBM),

$$\mathrm{d}S_t = \mu(t, ...)S_t\mathrm{d}t + \sigma(t, ...)S_t\mathrm{d}W_t, \tag{11}$$

with the difference that the drift term μ and the continuously-sampled volatility σ are arbitrary functions of time and other parameters, compared to being constants in a GBM (Demeterfi et al., 1999). The stochastic part of the equation is determined by the Wiener process, W_t, which together with the diffusion term, σ, make up the deviations from the expected return. The Wiener process has the following properties (Björk, 2009):

1. $W_0 = 0$.
2. The process has independent increments. Thus, if $r < s \leq t < u$ then $W_u - W_t$ and $W_s - W_r$ are independent stochastic variables.
3. For $s < t$, the stochastic variable $W_t - W_s$ has the Gaussian distribution $N[0,\sqrt{t-s}]$, i.e. it is normally distributed with mean zero and standard deviation $\sqrt{t-s}$.
4. W has continuous trajectories.

The asset is assumed to pay no dividends for simplicity. With the assumption in Equation (11), Demeterfi et al. (1999) show that the fair variance strike price is given by the equation

$$\begin{aligned} K_{\text{var}} = \frac{2}{T}\Bigg[& rT - \left(\frac{S_0}{S_*}e^{rT} - 1\right) - \log\left(\frac{S_*}{S_0}\right) \\ & + e^{rT}\left(\int_0^{S_*}\frac{1}{K^2}P(K)\mathrm{d}K + \int_{S_*}^{\infty}\frac{1}{K^2}C(K)\mathrm{d}K\right)\Bigg]. \end{aligned} \tag{12}$$

$P(K)$ and $C(K)$ respectively denotes the fair values for European put and call options that are written on the underlying asset described above, with strike price K and maturity T. The integrals in Equation (12) sum up an infinite number of European put and call options with continuous strike spectra and r is the risk-free discount rate. The options are written on the same underlying asset as the variance swap whose strike is approximated with the replication. The maturity of the options are the same as the maturity of the variance swap. S_* define the moneyness, or ratio between the underlying assets price and the strike price, boundary between the put and call options. The moneyness boundary can be seen as the approximate at-the-money (ATM) forward stock level (Demeterfi et. al., 1999). ATM is the moneyness where the strike price is equal to the price of the underlying asset. For simplicity, S_* can be set equal to S_0 which gives the simplification

$$K_{\text{var}} = \frac{2}{T}\left[1 + rT - e^{rT} + e^{rT}\left(\int_0^{S_0}\frac{1}{K^2}P(K)\mathrm{d}K + \int_{S_0}^{\infty}\frac{1}{K^2}C(K)\mathrm{d}K\right)\right]. \tag{13}$$

With the assumption that $S_* = S_0$, the portfolio consist of one ATM call option, one ATM put option and the rest of the options are out-of-the-money (OTM).

3.1.2 Replication with a discrete set of options

Because there are only a finite number of available options in the market which have a discrete set of strikes, the hypothetical portfolio implied by the integrals in Equation (12) require an approximation by a portfolio of finite traded options. As shown below, the payoff at maturity for the hypothetical portfolio in Equation (12) is

$$f(S_T) = \frac{2}{T}\left[\frac{S_T - S_*}{S_*} - \log\left(\frac{S_T}{S_*}\right)\right]. \tag{14}$$

This is also the payoff of a portfolio with two assets, a future on the underlying asset S_t with strike price S_* and a log contract on S_*, both with maturity T. The market price of these portfolios, if they were traded on the market, would be the same by argument of an arbitrage-free market. In practice, neither of the two portfolios are traded but as shown shortly, the payoff function in Equation (14) can be approximated by a finite number of traded options. This will be the replicating portfolio and because the payoff at maturity for the hypothetical option portfolio can be replicated with a discrete set of options, their current market values are the same and will provide an estimate of the strike price in Equation (12).

If the present value of the portfolio with a finite number of options is denoted Π_{CP}, and is substituted with the hypothetical portfolio in Equation (12), the approximation for K_{var} is

$$K_{\text{var}} \approx \frac{2}{T}\left[rT - \left(\frac{S_0}{S_*}e^{rT} - 1\right) - \log\left(\frac{S_*}{S_0}\right)\right] + e^{rT}\Pi_{\text{CP}}, \tag{15}$$

or when $S_* = S_0$

$$K_{\text{var}} \approx \frac{2}{T}\left[1 + rT - e^{rT}\right] + e^{rT}\Pi_{\text{CP}}. \tag{16}$$

3.1.3 Derivation of the payoff function and how it can be approximated

To determine that Equation (14) is indeed the payoff at maturity of the hypothetical portfolio, a derivation is made with a simple example. It is given that put options have the payoff at maturity of $\max(K - S_T, 0)$ and call options have the payoff at maturity of $\max(S_T - K, 0)$. Assume that the stock price at maturity, S_T, is in the interval $(0, S_*)$. The call options have zero value and the payoff only depends on the put options, resulting in the following payoff for Π_{CP}

$$
\begin{aligned}
f(S_T) &= \frac{2}{T}\left[\int_{S_T}^{S_*} \frac{1}{K^2}(K - S_T)\mathrm{d}K + 0\right] = \frac{2}{T}\left[\int_{S_T}^{S_*} \frac{1}{K} - \frac{S_T}{K^2}\mathrm{d}K\right] \\
&= \frac{2}{T}\left[\left[\log(K) + \frac{S_T}{K}\right]_{S_T}^{S_*}\right] = \frac{2}{T}\left[\log(S_*) - \log(S_T) + \frac{S_T}{S_*} - \frac{S_T}{S_T}\right] \\
&= \frac{2}{T}\left[\frac{S_T - S_*}{S_*} - \log\left(\frac{S_T}{S_*}\right)\right],
\end{aligned}
\tag{17}
$$

which is the same as Equation (14). If we investigate a different case where the stock price at maturity is instead in the interval (S_*, ∞), the same result is obtained. In this case, the put options have zero value instead as in the previous example and the payoff only depends on the call options

$$
\begin{aligned}
f(S_T) &= \frac{2}{T}\left[0 + \int_{S_*}^{S_T} \frac{1}{K^2}(S_T - K)\mathrm{d}K\right] = \frac{2}{T}\left[\int_{S_*}^{S_T} \frac{S_T}{K^2} - \frac{1}{K}\mathrm{d}K\right] \\
&= \frac{2}{T}\left[\left[-\frac{S_T}{K} - \log(K)\right]_{S_*}^{S_T}\right] = \frac{2}{T}\left[-\frac{S_T}{S_T} + \frac{S_T}{S_*} - \log(S_T) + \log(S_*)\right] \\
&= \frac{2}{T}\left[\frac{S_T - S_*}{S_*} - \log\left(\frac{S_T}{S_*}\right)\right].
\end{aligned}
\tag{18}
$$

To complete the argument, it only remains to show how the payoff in Equation (14) can be replicated by a finite set of European call and put options. Assuming that you can trade European call options with strikes $K_0^c = S_* = S_0 < K_1^c < K_2^c < ...$ and European put options with strikes $K_0^p = S_* = S_0 > K_1^p > K_2^p > ...$ The strike prices' subscript indicates the individual number of the option and the superscript indicates whether it is the strike price for a put or for a call option. Using these options with individual weights, the payoff, $f(S_T)$, can be approximated with a piece-wise linear function. The first segment to the right of S_* is the same as the payoff of a call option with strike K_0. The weight of that option is determined by the slope

$$
w_0^c(K_0^c) = \frac{f(K_1^c) - f(K_0^c)}{K_1^c - K_0^c}. \tag{19}
$$

Figure (6) illustrate the linear approximation of the payoff curve, where the slope between K_0 and K_1^c is the weight of the call option with strike K_0. The slope of the segment between the strike prices K_1^c and K_2^c is steeper than the slope of the first segment, which can be explained by the additional option that is now in-the-money and increases the payoff. The payoff of the portfolio consisting of two options is

$$
w_0^c(K_0^c)(S_T - K_0) + w_1^c(K_1^c)(S_T - K_1), \tag{20}
$$

and when deriving it with respect to S_T and setting it equal to the slope of the segment it yields the expression

$$
w_0^c(K_0^c) + w_1^c(K_1^c) = \frac{f(K_2^c) - f(K_1^c)}{K_2^c - K_1^c}. \tag{21}
$$

Since we are interested in calculating the weight $w_1^c(K_1^c)$, we need to subtract $w_0^c(K_0^c)$ from the left hand side of the expression. This yields the final expression for calculating the weight

$$w_1^c(K_1^c) = \frac{f(K_2^c) - f(K_1^c)}{K_2^c - K_1^c} - w_0^c(K_0^c). \tag{22}$$

Another way of understanding the equation for determining the weight of the second option is to think that the payoff already consists of $w_0^c(K_0^c)$ which need to be subtracted. Continuing this method for all the options, the entire payoff curve for $f(S_T)$ can be approximated. The individual weights of each option can generally be determined by

$$w_n^c(K_n^c) = \frac{f(K_{n+1}^c) - f(K_n^c)}{K_{n+1}^c - K_n^c} - \sum_{i=0}^{n-1} w_i^c(K_i^c), \tag{23}$$

$$w_n^p(K_n^p) = \frac{f(K_{n+1}^p) - f(K_n^p)}{K_n^p - K_{n+1}^p} - \sum_{i=0}^{n-1} w_i^p(K_i^p). \tag{24}$$

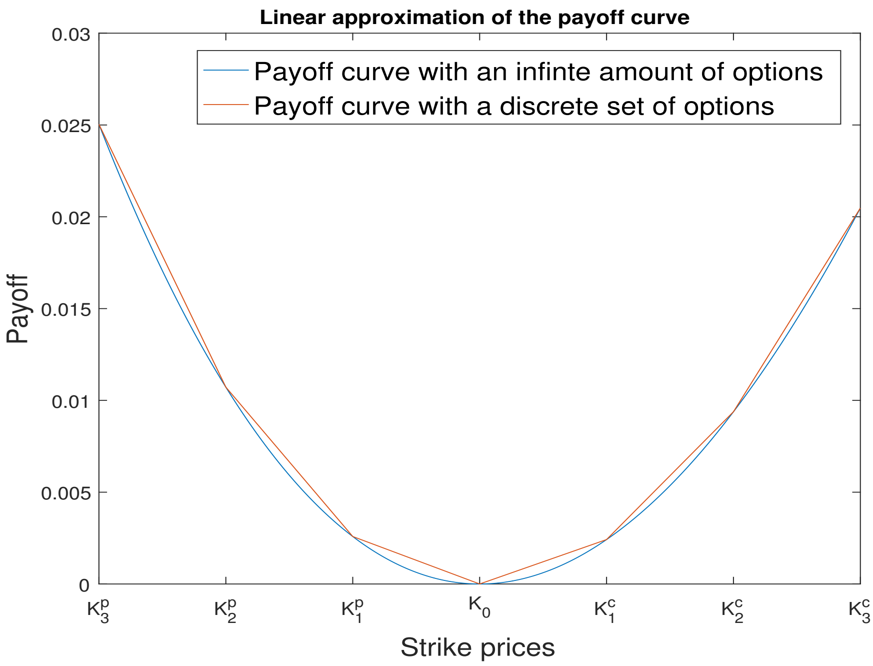

Figure 6: *Linear approximation of the payoff curve with a discrete set of options.*

Since the payoff at maturity for the option portfolio now is replicated with a discrete set of options, they have the same present value. With the weights for every individual option determined, the value of the option portfolio is obtained by

$$\Pi_{\text{CP}} = \sum_{i=0} P(K_i^p) w_i^p(K_i^p) + \sum_{i=0} C(K_i^c) w_i^c(K_i^c). \tag{25}$$

With the replicating portfolio defined one can use Equations (15), (23), (24) and (25) and a discrete set of European call and put options to calculate K_{var}.

3.1.4 Numerical example

Consider an example where we can trade options with strike prices that are uniformly spaced by five points and have a range between 50 and 110. Assume that the initial underlying stock price S_0 is 80, the risk free interest rate r is 5%, the dividend yield is zero and the maturity T for the swap is 3 months. The ATM implied volatility is assumed to be 22%, and have a smile (described in more detail in Subsection 3.3.3), causing the implied volatility to decrease with one percentage point for every 5 point increase in the strike price. The discrete set of options are illustrated in Table 2.

Table 2: *Example of a portfolio consisting of European options used for calculating the variance strike with the replication scheme.*

Option	Strike	Volatility (%)	Weight	BS price	Contribution
Put	50	28	16.08	0.0006	0.0102
Put	55	27	13.28	0.0054	0.0713
Put	60	26	11.15	0.0319	0.3552
Put	65	25	9.50	0.1407	1.3363
Put	70	24	8.18	0.4829	3.9524
Put	75	23	7.13	1.3293	9.4741
Put	80	22	3.26	3.0127	9.8264
Call	80	22	3.00	4.0065	12.0219
Call	85	21	5.55	1.8140	10.0603
Call	90	20	4.95	0.6342	3.1365
Call	95	19	4.44	0.1589	0.7052
Call	100	18	4.01	0.0260	0.1040
Call	105	17	3.63	0.0025	0.0089
Call	110	16	3.31	0.0001	0.0004
					Total: 51.0631

The weights in Table 2 are calculated using Equations (23) and (24) and the option values are determined using Black-Scholes formula for put and call options. The contribution of every individual option is its weight multiplied with its BS price. The sum of all contributions make up the total cost of the portfolio and is the value for Π_{CP} according to Equation (25). One can notice that the options that have a strike price close to the underlying stock price are the ones that contributes most to the total cost of the portfolio. The contributions then decrease as the options are more out-of-the-money and the reason for that is mostly because of the decreasing option value. Compared to the put options, the contributions decrease more rapidly for the call options as they become more and more out-of-the-money. This effect occur because the weight and the option price both decreases as the call options move out of the money, compared to the put options where the weights increase instead.

With the value of the option portfolio determined, one can use Equation (16), since $S_* = S_0$ in this case, to estimate K_{var}. For the example above, the fair variance strike is $(22.74\%)^2$ which is to be compared to the at-the-money implied volatility of 22%. The approximation is adequate and slightly overestimate the true value. The overestimation is expected as the approximation of Π_{CP} will always overestimate the true value, causing an overestimation of K_{var}.

When applying this replication scheme in practice, it causes imperfections since there are not an infinite amount of options in the market to accurately replicate the log contract. Using a discrete set of options will always overestimate the true theoretical value because of the linear approximation of the convex payoff function (Demeterfi et al., 1999). The scheme also require continuous purchasing of options which will be expensive in practice because of transaction costs and bid-ask spreads. Another fact that can cause imperfections in the estimated value for K_{var} is that the calculations are made with the assumption that no jumps occur in the underlying assets price movements, which may not replicate an assets price movement real life.

3.2 Deriving the fair volatility strike using Laplace transformations

Since volatility is the square root of variance, an approach that seem feasible up front is to use the square root of the calculated fair variance strike with the replication scheme,

$$K_{\text{vol}} = \sqrt{K_{\text{var}}}. \tag{26}$$

When doing so for deriving the fair value of a volatility strike price will cause an error since the convexity of the square root function is neglected. By calculating the payoffs for a variance swap and volatility swap with Equations (8) and (10), the payoff convexity in realized volatility for the variance swap is observed.

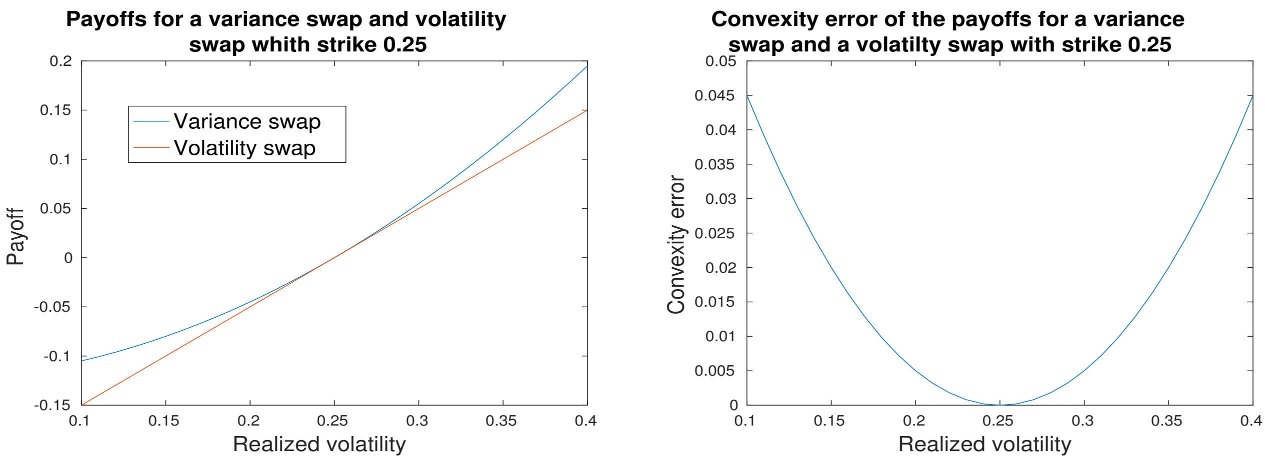

Figure 7: *Convexity error for a variance swap with strike* 0.25^2 *and* $N_{vega} = 1$.

Figure 7 illustrate the payoffs for a variance swap and a volatility swap with strike 0.25 when the realized volatility is ranged from 0.10 to 0.40. The payoff for the volatility swap is linear with the realized volatility but the payoff for the variance swap is not. The convexity of the payoff function for the variance swap is easily observed in the left side of the figure. The right side of the figure show the magnitude of the convexity error as the realized volatility deviates from the variance strike, quoted in volatility points. For the vega notional $N_{\text{vega}} = 1$ is the corresponding notional for the variance swap, according to Equation (6), $N_{\text{var}} = \frac{1}{2\times 0.25} = 2$.

The variance strike and the convexity of the square root function can however be used for creating an upper bound for the volatility strike (Broadie & Jain, 2008). As a consequence of Jensen's inequality,

$$E\left[\sqrt{X}\right] \leq \sqrt{E\left[X\right]}, \tag{27}$$

and substituting the random variable X for the continuous realized variance in (27) we get the upper bound for K_{vol} (Rolloos & Arslan, 2017),

$$K_{\text{vol}} = E\left[\sqrt{\frac{1}{T}\int_0^T \sigma_u^2 \mathrm{d}u}\right] \leq \sqrt{E\left[\frac{1}{T}\int_0^T \sigma_u^2 \mathrm{d}u\right]} = \sqrt{K_{\text{var}}}. \tag{28}$$

To deal with the convexity error, Brockhaus and Long (2000) have derived a convexity correction term with the use of a second order Taylor expansion on the square root function and expectations under the risk-neutral measure. Starting out by defining a square root function F as

$$F(x) = \sqrt{x}, \tag{29}$$

which has the first and second order derivatives

$$F'(x) = \frac{1}{2\sqrt{x}}, \tag{30}$$
$$F''(x) = -\frac{1}{4\sqrt{x}^3}. \tag{31}$$

Performing a Taylor-Series expansion for F around x_0 we obtain

$$\begin{aligned} F(x) &\approx F(x_0) + F'(x_0)(x - x_0) + \frac{1}{2}F''(x_0)(x - x_0)^2 \\ &\approx x_0^{1/2} + \frac{x - x_0}{2\sqrt{x_0}} - \frac{1}{8}\frac{(x - x_0)^2}{\sqrt{x_0}^3} \\ &\approx \frac{x + x_0}{2\sqrt{x_0}} - \frac{(x - x_0)^2}{8\sqrt{x_0}^3}, \end{aligned} \tag{32}$$

and when choosing $x = X$ and $x_0 = E[X]$

$$\sqrt{X} \approx \frac{X + E[X]}{2\sqrt{E[X]}} - \frac{(X - E[X])^2}{8\sqrt{E[X]}^3}. \tag{33}$$

Taking expectations on both sides yields

$$E[\sqrt{X}] \approx \frac{E[X] + E[X]}{2\sqrt{E[X]}} - \frac{E[(X - E[X])^2]}{8\sqrt{E[X]}^3}, \tag{34}$$

simplifying to

$$\sqrt{E[X]} - E[\sqrt{X}] \approx \frac{\text{Var}(X)}{8\sqrt{E[X]}^3}. \tag{35}$$

Using the definitions for volatility strike and variance strike from Equations (26), (27) and (28) the approximated convexity error has the form (Brockhaus & Long, 2000),

$$\sqrt{K_{\text{var}}} - K_{\text{vol}} \approx \frac{\text{Var}(\sigma_c^2(0, T))}{8\sqrt{K_{\text{var}}}^3}. \tag{36}$$

The magnitude of the convexity error depends on which model that is used to calculate the strike prices. For example, when estimating the volatility strike with the Heston model or with the Merton Jump Diffusion model, two models that are beyond the extent of this thesis, the convexity error approximation in Equation (36) will not be accurate (Broadie & Jain, 2008). The poor approximation will occur as a consequence of the fact that the error term will consist of a Taylor expansion of the third and fourth order as well. The higher order Taylor expansions makes the approximation a lot more complex and not very applicable.

The solution to the problem is to use a Laplace transformation (Broadie & Jain, 2008). This approach presents a way to solve the volatility strike price using the variance strike price by expressing the square root function as (Schürger, 2002)

$$\sqrt{X} = \frac{1}{2\sqrt{\pi}} \int_0^\infty \frac{1 - e^{-sX}}{s^{\frac{3}{2}}} \mathrm{d}s. \tag{37}$$

Using Fubini's theorem which makes it possible to switch the order of integration and evaluating (37) with expectations on both sides of the equal sign, the expression evolves into

$$E\left[\sqrt{X}\right] = \frac{1}{2\sqrt{\pi}} \int_0^\infty \frac{1 - E\left[e^{-sX}\right]}{s^{\frac{3}{2}}} \mathrm{d}s. \tag{38}$$

Substituting the random variable X in Equation (38) with the continuous realized variance we obtain a solution formula for calculating the volatility strike. The substitution gives

$$K_{\text{vol}} = E\left[\sqrt{\sigma_{\text{c}}^2(0,T)}\right] = \frac{1}{2\sqrt{\pi}} \int_0^\infty \frac{1 - E\left[e^{-s\sigma_{\text{c}}^2(0,T)}\right]}{s^{\frac{3}{2}}} \mathrm{d}s. \tag{39}$$

Or equivalently when assuming discrete realized variance,

$$K_{\text{vol}} = E\left[\sqrt{\sigma_{\text{d}}^2(0,T,n)}\right] = \frac{1}{2\sqrt{\pi}} \int_0^\infty \frac{1 - E\left[e^{-s\sigma_{\text{d}}^2(0,T,n)}\right]}{s^{\frac{3}{2}}} \mathrm{d}s. \tag{40}$$

To solve these equations and determine the volatility strike from the variance strike, one will have to use numerical integration techniques.

3.3 Black-Scholes model and implied volatility

3.3.1 Assumptions

It is possible to receive a volatility forecast from option prices that are observed in the market for European options. Using the Black-Scholes model to derive an explicit formula for calculating the option prices, the expected volatility can be received through a process of algebraic manipulation. The process consist of choosing a volatility that when used in the formula match the observed market price of the option. The resulting volatility is referred to as the implied volatility.

The Black-Scholes model was introduced in 1973 and has since then been widely used in the finance industry for pricing derivatives (Black & Scholes, 1973). The model include assumptions about the underlying asset and for the market. The underlying asset, S_t, is assumed to follow a random walk, or a geometric Brownian motion (GBM) and thus satisfying the following stochastic differential equation

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \tag{41}$$

where the drift term, μ, and the diffusion term, σ, are constants and quoted in annual terms. The drift is the annual expected return for the asset meaning that μdt is the expected return over a infinitesimal period of time. The Wiener process W_t is described in Subsection 3.1.1.

Other assumptions for the model are that there exist a bank account B with deterministic and constant interest rate r, there are zero transaction costs, that S_t can be traded continuously in any quantity required. Dividends are assumed to be paid continuously with yield q. With the assumption that the stock price follow a GBM, as described in Equation (41), it is log normally distributed,

$$\log(S_t) \sim N\left[\log(S_t) + \left(\mu - \frac{1}{2}\sigma^2\right)(T-t), \sigma\sqrt{(T-t)}\right]. \tag{42}$$

3.3.2 The Black-Scholes equation and formulas

With the assumptions and dynamic of the underlying asset described, we can continue to derive a formula for pricing derivatives. The payoff for the derivatives depends on the evolution of the underlying asset. The derivative to be priced is given the notation $V(S,t)$ which is a function whose value depends on time and the underlying asset's price. A risk-neutral portfolio can be created using an option with value $V(S,t)$ and a quantity of the asset S_t. Assuming that the portfolio consist of a long position of the option and a short position of Δ quantities of the asset. Using arguments of no-arbitrage, letting $\Delta = \frac{\partial V(S,t)}{\partial S_t}$ and applying Itô's lemma on the portfolio dynamics, the following partial differential (PDE) equation appear. For full derivation, the reader is referred to (Björk, 2009) and (Black & Scholes, 1973)

$$\frac{\partial V(S,t)}{\partial t} + rS_t\frac{\partial V(S,t)}{\partial S_t} + \frac{1}{2}\sigma^2 S_t^2\frac{\partial^2 V(S,t)}{\partial S_t^2} = rV(S,t). \tag{43}$$

The PDE (43) is called the Black-Scholes equation which describes the price of the option over time (Björk, 2009). When pricing a derivative written on an underlying asset with dynamics described in Equation (41), the payoff function has to be a solution to the Black-Scholes equation. Fortunately, this is the case for European call and put options. The payoff at maturity for a European call option is

$$C(S,T) = \max(S_T - K, 0), \tag{44}$$

where S_T is the price of the underlying asset at maturity, T and K is the strike price. For the European put option, the payoff at maturity is

$$P(S,T) = \max(K - S_T, 0). \tag{45}$$

Solving the Black-Scholes partial differential equation using the Feynman-Kac theorem and risk-neutral expectations of the possible outcomes for the options we get formulas for pricing both call and put options. The formulas are called the Black-Scholes formulas and are well-known and often used in the finance industry (Rolloos & Arslan, 2017).

The price of a European call option is given by

$$C_{\text{BS}}(S_t, K, \sigma, \tau, r, q) = S_t e^{-q\tau} N(d_1) - K e^{-r\tau} N(d_2), \tag{46}$$

and the price for a European put option is given by

$$P_{\text{BS}}(S_t, K, \sigma, \tau, r, q) = K e^{-r\tau} N(-d_2) - S_t e^{-q\tau} N(-d_1), \tag{47}$$

where $\tau = T - t$ is the time until maturity, σ is the volatility of the underlying asset and q is the dividend yield. $N(x)$ is the standard normal cumulative distribution function defined by

$$N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{z^2}{2}} \mathrm{d}z, \tag{48}$$

and the parameters d_1 and d_2 are defined as

$$d_1 = \frac{\log\left(\frac{S_t}{K}\right) + \left(r - q + \frac{\sigma^2}{2}\right)\tau}{\sigma\sqrt{\tau}}, \tag{49}$$

$$d_2 = d_1 - \sigma\sqrt{\tau}. \tag{50}$$

Using the Black-Scholes formulas in Equations (46) and (47) on observed market prices for European call and put options it is possible to obtain a volatility forecast. The forecast is referred to as the implied volatility, as it is the volatility that is implied by the market and option prices. The implied volatility is extracted by finding the volatility in the Black-Scholes formulas that match the market price of the option. The implied volatility is a market expectation of future realized volatility and can therefore be used as a tool for valuing volatility swaps.

The volatility index, VIX, mentioned in Subsection 2.1.1 is derived from the implied volatility from the Black-Scholes model using European options that are written on the Standard & Poor's 500 index. The calculations are made from options with a near term maturity, making the values of the index a 30 day forecast of volatility (CBOE, 2014). CBOE also have other indexes that corresponds to implied volatility of options with longer maturity. The index VXV is a volatility index that measure the 3 month implied volatility and can therefore be used as a 3 month volatility forecast. CBOE also provide an index that measure the 6 month implied volatility, which is called VXMT and can be used as a forecast of 6 month volatility.

3.3.3 Volatility smile

Since the volatility term in the Black-Scholes model is constant, European options that are written on the same underlying asset should have the same implied volatility regardless of the strike price and maturity (Alexander, 2008). When the implied volatility is calculated from market prices of options one should obtain a flat surface of the volatility level, under the assumption that the Black-Scholes model gives the accurate value of all the options.

If the implied volatility from market values of traded options with different strike prices and maturities are plotted in a three dimensional graph, a surface appear that is not flat. This phenomenon is called the implied volatility smile, as the implied volatility follows the shape of a smile.

Figure 8 illustrate the volatility smile phenomenon. The implied volatility surface should be a

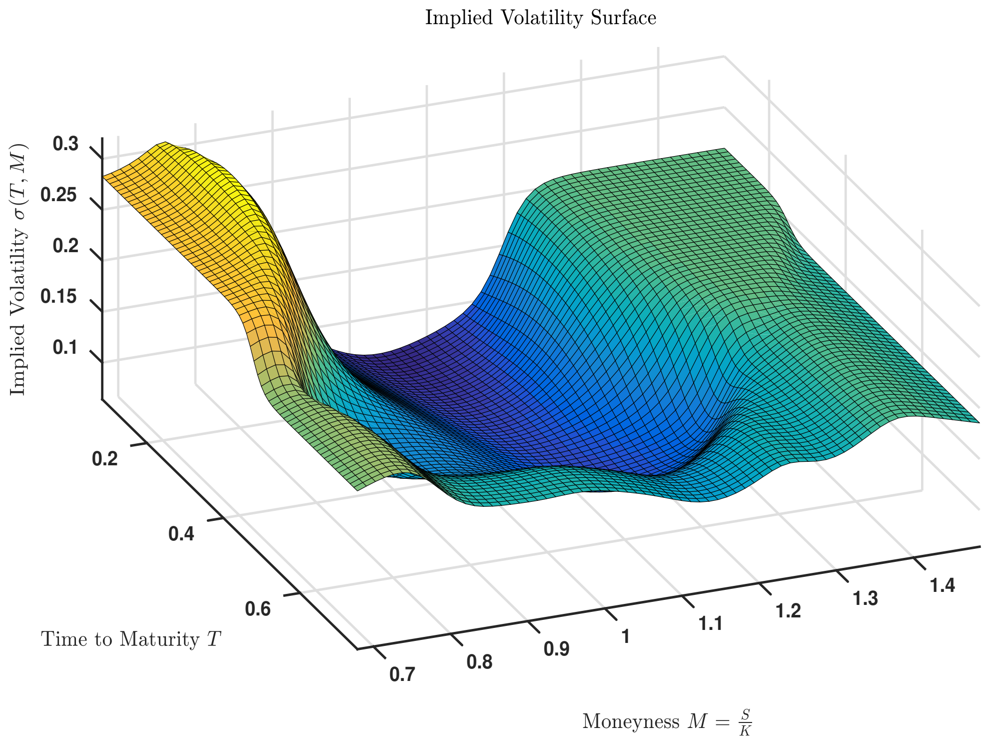

Figure 8: *Implied volatility surface from call options that are written on the Standard & Poor's 500 on 2017-04-19.*

flat plane in the graph, but it is not. The surface is obtained from market prices of about eight thousand call options that are written on the Standard & Poor's 500 on 2017-04-19. The option data is downloaded from CBOEs website.

3.4 Vanna-Vomma model

The Vanna-Vomma model proposed by Rolloos and Arslan (2017) uses Greeks from the Black-Scholes model to determine the volatility strike for a volatility swap. As described in section 3.3.3, an option with a certain maturity have an implied volatility that depends on the strike price because of the volatility smile. Estimations of the volatility strike done by calculating the implied volatility of an arbitrary chosen option, with the same maturity as the volatility swap, will differ depending on the strike price of the option. To know in advance which strike price that corresponds to the "correct" implied volatility is difficult. This is where the usage of the Greeks vanna and vomma comes to the rescue, as they can determine the strike price of which option to use to estimate the volatility strike price.

The model is very new and has therefore not been used that frequently and not tested enough to conclude if the results are accurate. To derive the resulting formula for calculating the volatility strike, a stochastic volatility model is included in order to motivate that the formula is correct and applicable to reality. To start the description of the model, the required Greeks from the Black-Scholes model are introduced.

3.4.1 Greeks

Greeks are sensitivities of how the value of an option change with the parameters of the underlying asset. The Greeks needed for this model are delta (Δ), vega (ν), vanna (va_{BS}) and vomma

(vo_{BS}) (Rolloos & Arslan, 2017):

$$\Delta_{\text{call}} = \frac{\partial C(S,T)}{\partial S} = e^{-qT} N(d_1), \tag{51}$$

$$\Delta_{\text{put}} = \frac{\partial P(S,T)}{\partial S} = -e^{-qT} N(-d_1), \tag{52}$$

$$\nu_{\text{BS}} = \frac{\partial C(S,T)}{\partial \sigma} = \frac{\partial P(S,T)}{\partial \sigma} = Se^{-qT} n(d_1)\sqrt{T}, \tag{53}$$

$$va_{\text{BS}} = \frac{\partial^2 C(S,T)}{\partial S \partial \sigma} = \frac{\partial^2 P(S,T)}{\partial S \partial \sigma} = -\frac{e^{-qT} d_2}{\sigma} n(d_1), \tag{54}$$

$$vo_{\text{BS}} = \frac{\partial^2 C(S,T)}{\partial \sigma^2} = \frac{\partial^2 P(S,T)}{\partial \sigma^2} = \frac{d_1 d_2}{\sigma} \nu. \tag{55}$$

Delta measure how sensitive the option price is to the underlying asset's price. Vega is another sensitivity that measure how much the option price change with the underlying asset's volatility. Vanna and vomma are second order sensitivities. Vanna is the second order derivative of the price of the option, with respect to the underlying asset's price and to the underlying asset's volatility. It can also be viewed as the sensitivity of the option delta with respect to volatility, or equivalently, the sensitivity of the option vega with respect to the underlying asset's price. Vomma is the second derivative of the option price with respect to the underlying asset's volatility.

Noticeable is that delta for both call and put options are expressed using the standard normal cumulative distribution function, described in Equation (48) but vega and vanna instead are expressed using the standard normal probability density function,

$$n(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}. \tag{56}$$

The Greeks, their behavior and usage are interesting (Haug, 2003) but this model focuses on vanna and vomma. Inspecting the formulas for vanna and vomma one can notice that they both are proportional to d_2 and that they will both be zero when $d_2 = 0$. This fact will be vital for deriving the result.

3.4.2 Approach

Implied volatility can be viewed as a function of the option strike price, $\sigma(K)$. Using market data, the implied volatility can be determined for any traded option with strike price K. By interpolating, one can approximate $\sigma(K)$ for any K within the traded range of strike prices.

As explained in the next section, we want to extract the implied volatility corresponding to the strike for which both vanna and vomma are zero, that is, when $d_2 = 0$. One can note that d_2 is an increasing function of K, and there exists a unique K such that $d_2 = 0$. If we let K_{d_2} be the strike price when $d_2 = 0$, let σ_{d_2} be the corresponding implied volatility, use the formula for d_2, (50), and solve for K we get

$$K_{d_2} = S_t e^{\left(r-q-\frac{1}{2}\sigma_{d_2}^2\right)T}. \tag{57}$$

If a unique solution, K_{d_2}, exists, the implied volatility $\sigma(K_{d_2})$ gives the estimated volatility strike for a volatility swap, that is $K_{\text{vol}} = \sigma(K_{d_2})$. Thus, to calculate the volatility strike, all one needs to do is to solve Equation (57) and find the corresponding implied volatility for that particular strike price.

An alternative method is to plot the vanna or vomma of all available quoted options and find the point for which the implied volatility corresponds to a vanna or vomma that is equal to zero (Rolloos & Arslan, 2017).

3.4.3 Motivation of the model

To motivate the accuracy of the Vanna-Vomma model one can present a stochastic volatility framework to illustrate how the result satisfies assumptions about stochastic volatility. In a stochastic volatility framework, the dynamics of the underlying asset price process and its volatility can be written as

$$\mathrm{d}S_t = rS_t\mathrm{d}t - \sigma_t S_t \mathrm{d}W_t^s, \tag{58}$$
$$\mathrm{d}\sigma_t = a(\sigma_t,t)\mathrm{d}t + b(\sigma_t,t)\mathrm{d}W_t^\sigma, \tag{59}$$

where $\mathrm{d}W_t^s$ and $\mathrm{d}W_t^\sigma$ are correlated Weiner processes

$$\mathrm{d}W_t^s \mathrm{d}W_t^\sigma = \rho \mathrm{d}t. \tag{60}$$

The functions $a(\sigma_t,t)$ and $b(\sigma_t,t)$ and their parameters are calibrated so that the price of European options under the stochastic volatility framework are equal to the market prices. Suppose that $C_{\text{SV}}(S_t,t,K,T,\sigma_t)$ denotes the price of a European call option under the stochastic volatility framework (58)-(60) where the parameters has been calibrated to market prices and $C_{\text{mkt}}(S_t,t,K,T)$ denote the market price of the option. Then, by definition of implied volatility,

$$C_{\text{SV}}(S_t,t,K,T,\sigma_t) = C_{\text{mkt}}(S_t,t,K,T) = C_{\text{BS}}(S_t,t,K,T,\sigma). \tag{61}$$

Based on the work done by Hull and White (1987), Romano and Touzi (1997) and Willard (1997), a connection between the stochastic volatility price and the Black-Scholes price was established by Rolloos and Arslan (2017),

$$C_{\text{SV}}(S_t,t,K,T,\sigma_t) = E_t\left[C_{\text{BS}}(S_t M(t,T,\rho),t,K,T,\sigma_{\text{c}}(t,T)\sqrt{1-\rho^2})\right], \tag{62}$$

where

$$M(t,T,\rho) = e^{\frac{\rho^2}{2}\int_t^T \sigma_u^2 \mathrm{d}u + \rho\int_t^T \sigma_u \mathrm{d}W_u^\sigma}, \tag{63}$$

and $\sigma_{\text{c}}(t,T)$ is the continuous volatility on the period $[t,T]$, recall equation (2). Substituting the price for the option with stochastic volatility with the price using the Black-Scholes model in (62) we get,

$$C_{\text{BS}}(S_t,t,K,T,\sigma) = E_t\left[C_{\text{BS}}(S_t M(t,T,\rho),t,K,T,\sigma_{\text{c}}(t,T)\sqrt{1-\rho^2})\right]. \tag{64}$$

Taylor expanding equation (64) around the correlation parameter ρ leads to,

$$\begin{aligned} C_{\text{BS}}(S_t,t,K,T,\sigma) \approx{}& E_t\left[C_{\text{BS}}(S_t,t,K,T,\sigma_{\text{c}}(t,T))\right] \\ &+ \rho\frac{\partial}{\partial\rho}E_t\left[C_{\text{BS}}(S_t M(t,T,\rho),t,K,T,\sigma_{\text{c}}(t,T)\sqrt{1-\rho^2})\right]_{\rho=0} + O(\rho^2), \end{aligned} \tag{65}$$

leading to,

$$\begin{aligned}C_{\text{BS}}(S_t,t,K,T,\sigma) &\approx E_t\left[C_{\text{BS}}(S_t,t,K,T,\sigma_{\text{c}}(t,T))\right] \\ &+ \rho S_t E_t\left[\Delta_{\text{call}}(S_t,t,K,T,\sigma_{\text{c}}(t,T))\int_t^T \sigma_u \mathrm{d}W_u\right] + O(\rho^2). \end{aligned} \quad (66)$$

Continuing the derivation of the result, the specific choice $K = K_{d_2}$ for which $\sigma = \sigma_{d_2}$ are made in Equation (66),

$$\begin{aligned}C_{\text{BS}}(S_t,t,K_{d_2},T,\sigma_{d_2}) &\approx E_t\left[C_{\text{BS}}(S_t,t,K_{d_2},T,\sigma_{\text{c}}(t,T))\right] \\ &+ \rho S_t E_t\left[\Delta_{\text{call}}(S_t,t,K_{d_2},T,\sigma_{\text{c}}(t,T))\int_t^T \sigma_u \mathrm{d}W_u\right] + O(\rho^2), \end{aligned} \quad (67)$$

and then Taylor expanding C_{BS} and Δ_{call} on the right-hand side around σ_{d_2} gives:

$$\begin{aligned}C_{\text{BS}}(S_t,t,K_{d_2},T,\sigma_{\text{c}}(t,T)) \approx\ & C_{\text{BS}}(S_t,t,K_{d_2},T,\sigma_{d_2}) \\ &+ \nu_{\text{BS}}(S_t,t,K_{d_2},T,\sigma_{d_2})(\sigma_{\text{c}}(t,T)-\sigma_{d_2}) \\ &+ \frac{1}{2}vo_{\text{BS}}(S_t,t,K_{d_2},T,\sigma_{d_2})(\sigma_{\text{c}}(t,T)-\sigma_{d_2})^2, \end{aligned} \quad (68)$$

respectively,

$$\begin{aligned}\Delta_{\text{call}}(S_t,t,K_{d_2},T,\sigma_{\text{c}}(t,T)) \approx\ & \Delta_{\text{call}}(S_t,t,K_{d_2},T,\sigma_{d_2}) \\ &+ va_{\text{BS}}(S_t,t,K_{d_2},T,\sigma_{d_2})(\sigma_{\text{c}}(t,T)-\sigma_{d_2}). \end{aligned} \quad (69)$$

Using the fact that all quantities that does not depend on $\sigma_{\text{c}}(t,T)$ can be put outside the expectation in Equation (67) and that

$$va_{\text{BS}}(S_t,t,K_{d_2},T,\sigma_{d_2}) = vo_{\text{BS}}(S_t,t,K_{d_2},T,\sigma_{d_2}) = 0, \quad (70)$$

Equation (67) is simplified to

$$\begin{aligned}C_{\text{BS}}(S_t,t,K_{d_2},T,\sigma_{d_2}) \approx\ & C_{\text{BS}}(S_t,t,K_{d_2},T,\sigma_{d_2}) \\ &+ \nu_{\text{BS}}(S_t,t,K_{d_2},T,\sigma_{d_2})E_t\left[(\sigma_{\text{c}}(t,T)-\sigma_{d_2})\right]. \end{aligned} \quad (71)$$

This can however only be the case if

$$K_{\text{vol}} = E_t\left[\sigma_{\text{c}}(t,T)\right] \approx \sigma_{d_2}, \quad (72)$$

and thus is the formula for the volatility strike fully derived and motivated. Despite the model's straight forward approach and that the calculations does not require any heavy computations, it still consider assumptions about stochastic volatility in it's estimations of the volatility strike.

4 Valuation using stochastic volatility models

This section covers stochastic volatility models that can be used when forecasting volatility. The section first covers the Exponentially weighted moving average (EWMA) model. Afterwards, the generalized autoregressive conditional heteroskedasticity (GARCH) model is described and derived. EWMA and GARCH do not require any option data as they both use historical prices of the same asset as the one that the forecast is made for. In this thesis, it is historical prices on the Standard & Poor's 500 stock index.

4.1 Exponentially Weighted Moving Average

EWMA uses recently past asset returns to estimate the future volatility levels. It is an average of a certain number of past squared returns where the individual weights of each observed squared return decrease exponentially, giving the most recent observations highest contribution to the forecast. The number of past observations in the asset price is called estimation window and is denoted W_E. As described in Subsection 2.1.1, volatility levels tends to move in clusters, which is why the letting the individual weight of each observation exponentially decline into the past is a good idea. The basis of deriving the EWMA model is by calculating the conditional variance at time t (Danielsson, 2011) with the formula

$$\sigma_t^2 = \frac{1-\lambda}{\lambda(1-\lambda^{W_E})} \sum_{i=1}^{W_E} \lambda^i y_{t-i}^2. \tag{73}$$

The model was first introduced by J.P. Morgan by the name RiskMetrics where they used $\lambda = 0.94$ for a daily sample of returns, which is the λ of choice in this thesis (J.P. Morgan, 1996). When $\lambda = 0.94$, λ^{W_E} approach zero very quickly as the estimation window increase, and as described shortly, the formula above can be rewritten to an explicit formula that is used in this thesis for forecasting volatility. The EWMA model is defined as

$$\sigma_t^2 = (1-\lambda)y_{t-1}^2 + \lambda\sigma_{t-1}^2, \tag{74}$$

where $0 < \lambda < 1$ is in this case a persistence parameter to describe how persistent recent variance is to the estimation of the variance for the next period. $(1-\lambda)$ is a reaction parameter and describes the asset's returns impact of the estimated variance.

4.1.1 Derivation of the model

Since the exponential λ^{W_E} decrease rapidly, one can approximate Equation (73) by its limit if we let W_E approach infinity (Danielsson, 2011)

$$\sigma_t^2 = \frac{1-\lambda}{\lambda} \sum_{i=1}^{\infty} \lambda^i y_{t-i}^2. \tag{75}$$

For a W_E that is large enough, but not necessarily very large as shortly explained, the weights beyond the estimation window, λ^n, are extremely small and can be neglected for all $n \geq W_E$. Figure 9 illustrate the rapid decline for the weights as the estimation window increase and shows that the estimation window does not have to be very large for the weights to approach zero. For example, by using an estimation window of 60 days, the last weight λ^{W_E} is 0.0244 and the weights beyond 60 days continues to decrease.

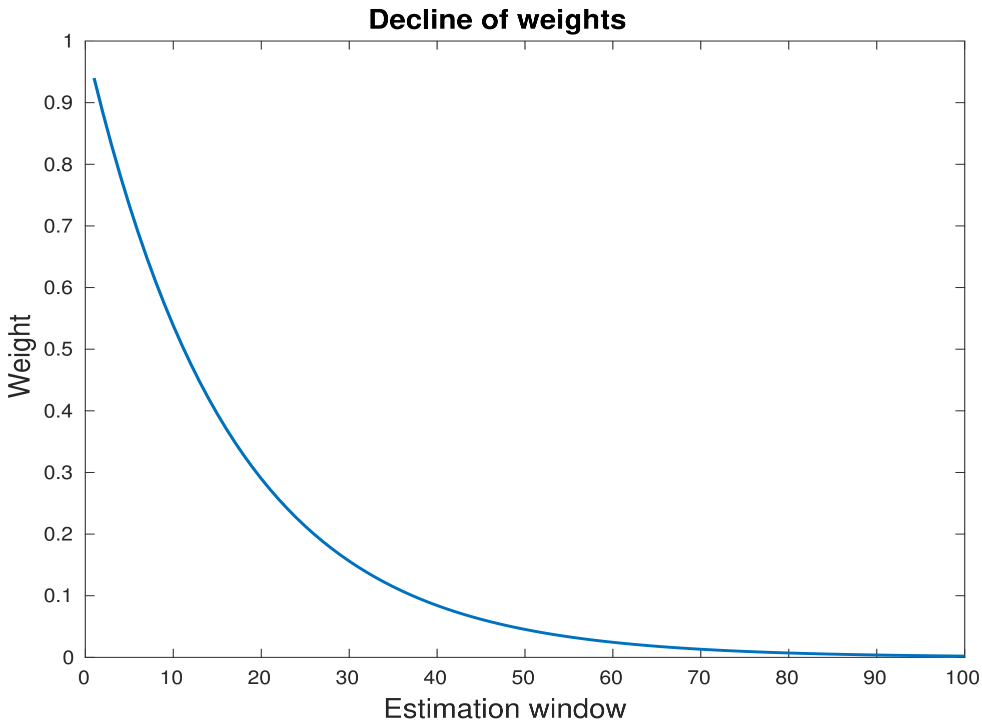

Figure 9: *The decline of weights as the estimation window increase when* $\lambda = 0.94$.

Hence, for $i > n$, the terms in the summation can be considered equal to zero. Taking out the first term of the summation, Equation (75) turns into

$$\sigma_t^2 = (1-\lambda)y_{t-1}^2 + \frac{1-\lambda}{\lambda}\sum_{i=2}^{\infty}\lambda^i y_{t-i}^2, \tag{76}$$

and if the indices are rearranged in the summation and for the λs before it, the equation becomes

$$\sigma_t^2 = (1-\lambda)y_{t-1}^2 + (1-\lambda)\sum_{i=1}^{\infty}\lambda^i y_{t-1-i}^2. \tag{77}$$

Looking at Equation (73) one can see that the expression above is very similar to the conditional variance at time $t-1$, which can be rearranged to

$$\frac{\lambda(1-\lambda^{W_E})}{1-\lambda}\sigma_{t-1}^2 = \sum_{i=1}^{W_E}\lambda^i y_{t-1-i}^2. \tag{78}$$

Substituting the summations and this equality while considering $\lambda^{W_E} = 0$ for a W_E large enough we get the EWMA formula described in Equation (74).

The formula for EWMA gives a one period forecast only. We need to forecast volatility and variance over several periods (days) which require simulation of the asset's returns. Since the EWMA forecasts most likely differ because of the simulated asset return at each time step, a Monte Carlo simulation of each EWMA forecast can be made to get a more reliable result. The procedure for calculating the volatility strike and variance strike with Monte Carlo simulations and EWMA is later described in Subsection 5.2.3.

4.2 Generalized autoregressive conditional heteroskedasticity

Generalized autoregressive conditional heteroskedasticity (GARCH) is a model that is a bit more complex than EWMA. GARCH and EWMA look similar but the main difference is the parameter estimations that are made in GARCH. The general notation for this model is GARCH(p,q) and is defined as (Danielsson, 2011)

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i Y_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2. \tag{79}$$

This thesis investigates the performance of GARCH(1,1),

$$\sigma_t^2 = \omega + \alpha Y_{t-1}^2 + \beta \sigma_{t-1}^2. \tag{80}$$

Beside from the constant ω, the GARCH(1,1) model is the same as EWMA where the names of the parameters α and β corresponds to the EWMA parameters $(1-\lambda)$ and λ, respectively.

The parameters in the GARCH model have a couple of restrictions (Danielsson, 2011)

- $\omega, \alpha, \beta > 0$: To ensure that all forecasts are positive
- $\alpha + \beta < 1$: To ensure covariance stationarity

With the parameter estimation in place, the forecasting of volatility and variance with GARCH(1,1) are done in the same fashion as for the EWMA model, with a Monte Carlo simulation for every estimation.

5 Method

In this section, the methods for calculating the volatility strike for the different models are explained. The data that has been used is firstly presented in Section 5.1. Section 5.2 covers the implementation of the models, which explains what necessary assumptions that are made and how the simulations for EWMA and GARCH are done. The simulations for EWMA and GARCH require Monte Carlo simulations.

5.1 Data

The data for S&P500 and the volatility indexes VIX, VXV and VXMT used in this thesis are daily closing prices. They are downloaded from Yahoo Finance into matlab with the function hist_stock_data, which is available for download at www.mathworks.com. VXMT have data available which goes back to 2008-01-07 and that is the start of the backtesting for the models Black-Scholes, EWMA and GARCH.

The option data required for the Vanna-Vomma model is downloaded from CBOE:s service livevol, consisting of European call and put options written on S&P500 and initiated during 2015. The options are filtered by removing the options that are non-traded (usually the ones that are far out-of-the-money) since they have big bid-ask spreads and their prices are therefore not reliable.

The maturities for the options vary with each day as the expiration dates occur the third Saturday each month. The option data for a particular day have expiration dates for the following three months, and then there are options with quarterly expirations. The options with quarterly expirations are filtered out. The data set also consist of so called Long-Term Equity Anticipation Securities (LEAPs), which are also filtered out when using the option data.

The Vanna-Vomma model require a risk free interest rate when calculating the implied volatility. The risk free interest that is used is the 1 year US Treasury yield. The interest rate is downloaded from www.treasury.gov and consists of daily interest rates during 2015.

5.2 Implementation

5.2.1 Black-Scholes model

The closing prices for the volatility indexes are downloaded to matlab and compared to the realized volatility.

5.2.2 Vanna-Vomma model

To get an estimation of the volatility strike with the Vanna-Vomma model, equation (57) is solved for the filtered set of options, that is, the strike price for the option corresponding to a vanna that is equal to zero. The implied volatility for the option with that particular strike is the estimate for the volatility strike.

The comparison of the models are done with maturities of 1, 2 and 3 months. Since the options only expire once a month, we get an estimation for each maturity once a month. For example, the options quoted in January will expire the third Saturday of February, March and April. The exact dates for the expirations vary from month to month. In February does the options expire on the twentieth and it is the same for the options expiring in March. For April however does the options expire on the seventeenth. The estimations of the volatility strike are thus done on options quoted on January 20 for maturities 1 and 2 months and on January 17 for the swap with maturity of 3 months.

If no solution is found where vomma is zero, the option with the strike price corresponding to the vomma that is closest to zero is chosen instead.

The interest rate that is used when calculating the implied volatility is the 1 year US Treasury yield observed on the particular day that the implied volatility is calculated.

5.2.3 EWMA

The EWMA estimations are done with Monte Carlo simulation because of it's stochastic nature when forecasting the volatility. The procedure of using N Monte Carlo simulations for calculating the annual volatility and variance forecast of n days is as following:

- Start by calculating the initial squared log return y_{t-1}^2 and variance σ_{t-1}^2.
- For every Monte Carlo simulation N:
  - At $t = 1$:
    - Use equation (74) to forecast the variance, $\sigma_1^2 = (1-\lambda)y_{t-1}^2 + \lambda\sigma_{t-1}^2$.
    - Take the square root for volatility, σ_1.
    - Preparation for the next period: Calculate the log return Y_1 by multiplying the forecast volatility with an independent and identically distributed (IID) random variable with mean 0 and variance 1, Z_t. The distribution of Z_t is assumed to be normal but Student-t distribution is another that is frequently used (Danielsson, 2011). Y_1= $\sigma_1 Z_1$.
  - At $t = 2$:
    - Input the log return and variance calculated in the previous step in equation (74) to forecast the variance at $t = 2$: $\sigma_2^2 = (1-\lambda)Y_1^2 + \lambda\sigma_1^2$.
    - Take the square root for volatility, σ_2.
    - Do the same as the time step before: Y_2= $\sigma_2 Z_2$.
  - Continue this process until $t = n$.
- Let $\boldsymbol{\Sigma(n,N)}$ and $\boldsymbol{\Sigma^2(n,N)}$ respectively denote $n \times N$ matrices of the n daily forecasts of volatility and variance by N simulations. By calculating the mean of $\boldsymbol{\Sigma(n,N)}$ and $\boldsymbol{\Sigma^2(n,N)}$ column wise, the average daily volatility and variance for every simulation $\boldsymbol{N}$ is obtained. To get the values in annual terms, the volatilities are multiplied with $\sqrt{252}$ and the variance is multiplied with 252. The vectors of the average annual volatility and variance for every simulation N can be denoted by $\boldsymbol{\bar{\sigma}}$ and $\boldsymbol{\bar{\sigma}^2}$, respectively. To complete the calculations, the mean of $\boldsymbol{\bar{\sigma}}$ and $\boldsymbol{\bar{\sigma}^2}$ are respectively calculated, which results in the estimations of K_{vol} and K_{var}

In this thesis, 10 000 Monte Carlo simulations are made for each estimation.

5.2.4 GARCH

GARCH use the same procedure as the EWMA when forecasting the volatility with Monte Carlo simulations. The difference is the parameter estimations that are done daily before each group of Monte Carlo simulations. The parameters ω, α and β are calibrated with a maximum likelihood estimation from 10 years of previous logarithmic returns and are updated daily for every new estimation. The long estimation window is chosen in order to get a stable estimation.

6 Results

This section covers the results from the different models' estimations of volatility strikes compared to the realized volatility. Firstly, a comparison of the models' performance over 2015 are made. To test the models' performance, they are measured in a number of performance measures. The measures are the mean value, standard deviation and upper and lower tail of the difference of the estimations from the realized volatility. For the Vanna-Vomma model, the results are made from options with maturities 1, 2 and 3 months that are initiated during the year of 2015 and the estimates from the other models are calculated with returns from S&P500 during the same time periods as for the estimations with the Vanna-Vomma model.

To get a more robust valuation of the models' performance, the comparison of performance with data from 2015 is followed by a section where the BS model, EWMA model and GARCH model are backtested. The backtesting is done for a maturity of 6 months and over the period 2008-01-08 to 2016-12-31. The Vanna-Vomma model is left out because of limitations of historical options data. To get even more material for the comparisons, an even longer backtesting have been made for EWMA and GARCH when the maturities are 1, 3 and 12 months. The results from the long backtesting is presented in Appendix C.

6.1 Comparison of the models

The comparison of the models Vanna-Vomma, Black-Scholes, EWMA and GARCH are made from the data set from year 2015. It covers the performance measures mean value, standard deviation and upper and lower tails from the distribution of the deviations of the models' estimations of volatility from the realized volatility. The realized volatility are calculated on the S&P500. The upper and lower tail are the 1th and 99th percentile of the differences of the estimations from the realized volatility, respectively.

Figure 10 illustrates the performance of the individual models for the maturities 1, 2 and 3 months.

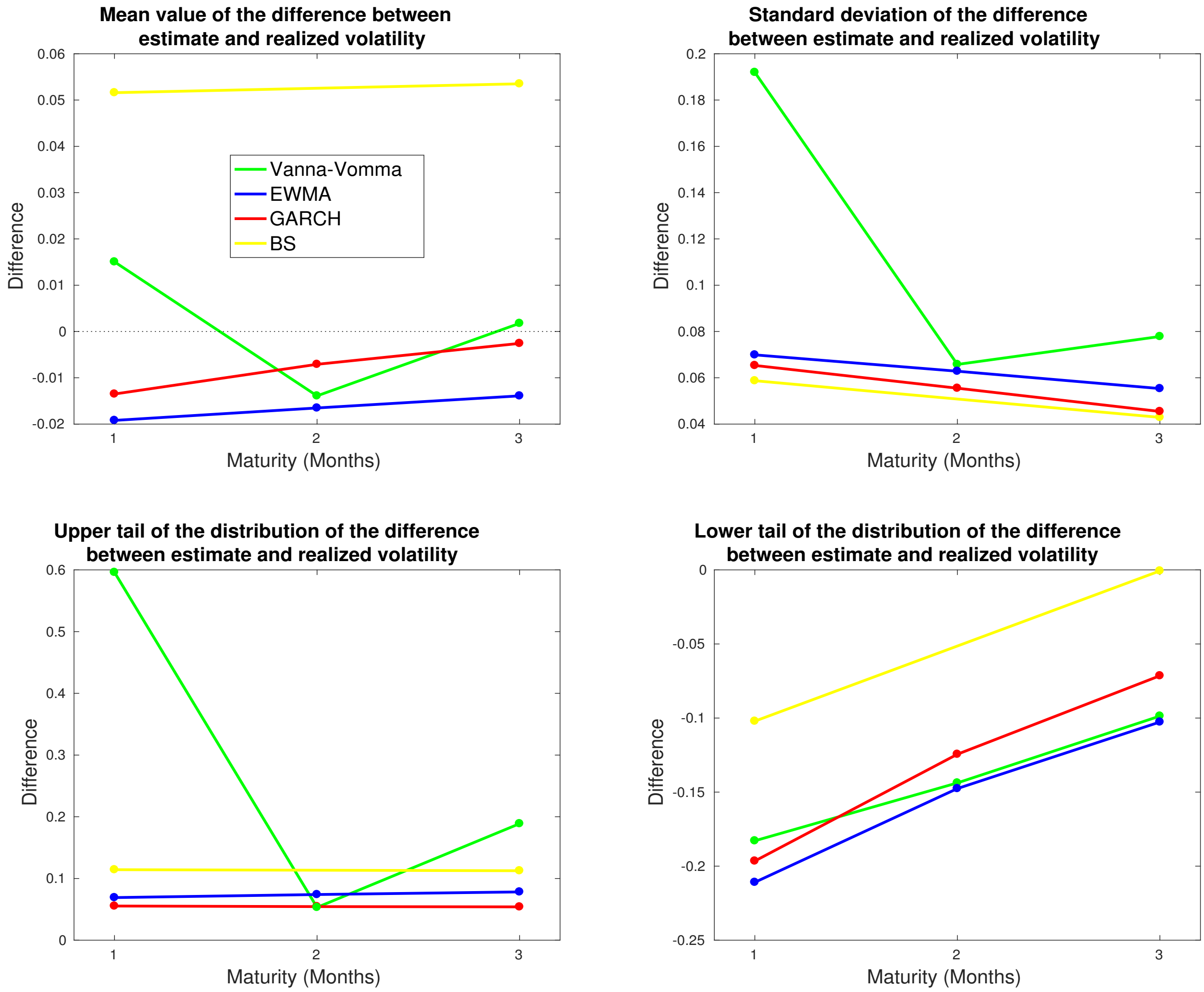

__Figure 10:__ A comparison of the models' performance based on mean value, standard deviation, and upper and lower tails of the differences between estimates and realized volatility.

6.2 Backtesting

To get at more extensive comparison with more data and longer maturity, the Black-Scholes model, EWMA and GARCH are backtested from 2008 to the end of 2016 when the maturity is 6 months. For the BS model is the VXMT index used as the 6 month forecast of volatility. Figure 11 illustrate the evolution of the forecasts together with a histogram of each models forecasts' deviations from the realized volatility. The histograms are fitted to a normal distribution, which are the black lines in the histograms.

Appendix C consist of an even more extensive backtesting, which is done for EWMA and GARCH from 1995-01-01 to 2016-12-31 and for the maturities 1, 3 and 12 months.

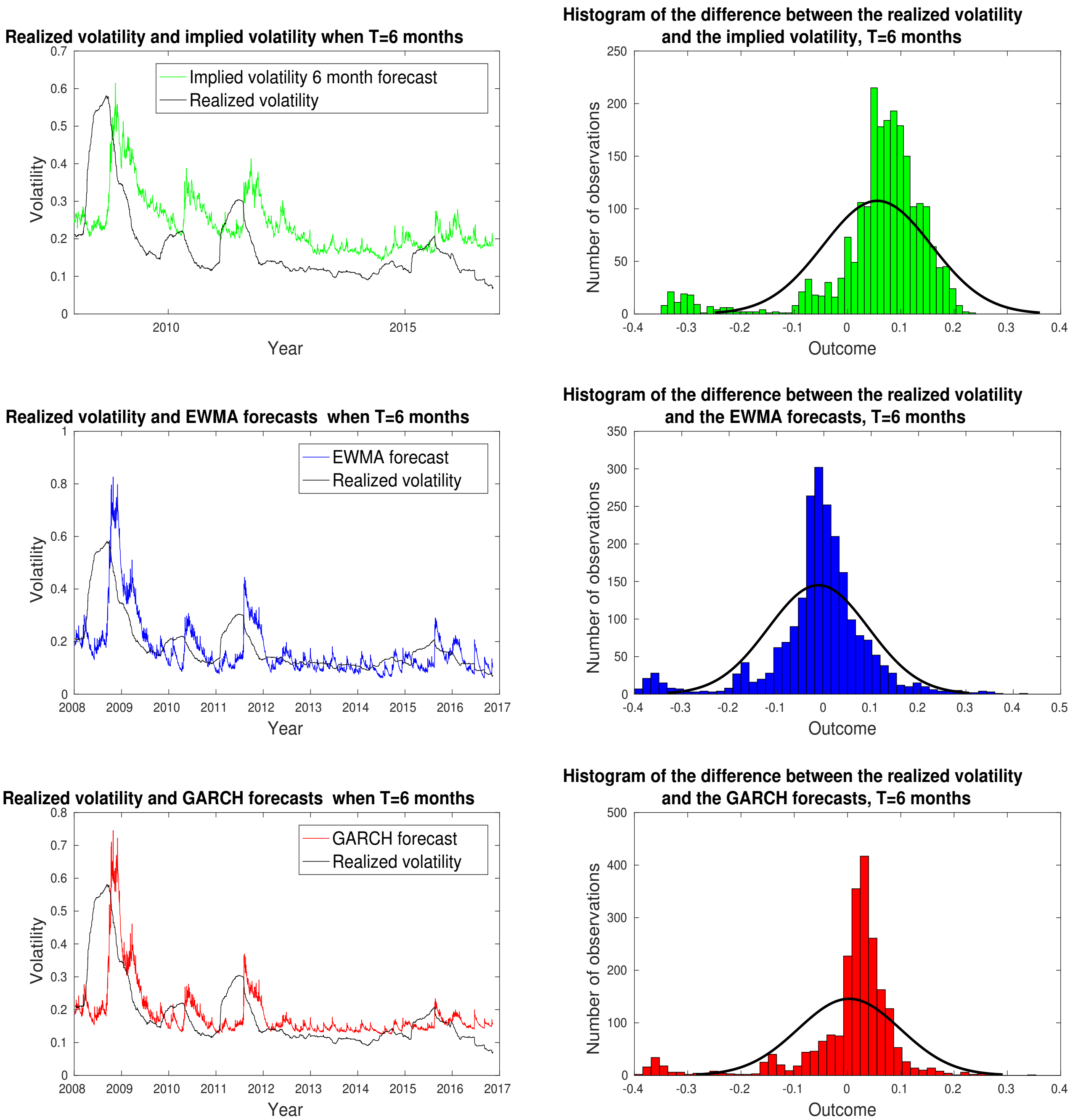

Figure 11: *The evolution of the realized volatility and the volatility forecasts of the Black-Scholes model, EWMA and GARCH.*

7 Discussion

Volatility derivatives are securities that enable traders to speculate on future volatility levels and to hedge their portfolios. Accurate pricing of derivatives are crucial for keeping the financial markets stable. In this thesis, I have investigated four salient methods for predicting future volatility and hence the strike price of volatility derivatives. I have also shown how the strike price for a variance swap can be approximated with a replicating scheme. Taken together, my results show that volatility can be accurately estimated, especially with the stochastic volatility models EWMA and GARCH, and that the implied volatility calculated from market prices of options overestimate the future volatility. I have also shown that the use of vanna and vomma when calculating the volatility strike takes care of the problem with the volatility smile in the Black-Scholes model, which is illustrated in Figure 8.

Ending the discussion chapter, and this thesis, is a possible outlook of volatility derivatives and the methods of how they can be priced.

7.1 Conclusions and discussions of the results

From Figure 10 one can see how the GARCH model overall perform the best. For a swap with maturity of 1 and 2 months, the mean value of the difference between the GARCH estimations and the realized volatility is the lowest for all the models. When the maturity is 3 months, the mean value of the difference is close to zero, but it is beaten by the Vanna-Vomma model by a fraction. By investing the extreme outcomes by looking at the tails for the GARCH model, it is clear that it performs well. The upper tail is the lowest of all models for all maturities, except for the Vanna-Vomma model when the maturity is 2 months, where the upper tail for Vanna-Vomma is slightly lower. Investigating the lower tails for the models, the Black-Scholes model is the winner, followed by the Vanna-Vomma model when the maturity is 1 month and by the GARCH model when the maturity is 2 and 3 months.

When looking at the lower tail, the Black-Scholes model performs the best. However, that is because is tends to overestimate the future volatility, as the mean value of the difference between its estimates and the realized volatility is highly above zero. The BS model's overestimation can also be quantified by looking at the lower tail when the maturity is 3 months. The extreme value for the Black-Scholes model when the maturity is 3 months is -0.0007. So the most extreme outcome is almost zero, meaning that all the other estimations is likely to be above zero and hence an overestimation of the true value of the future realized volatility. The overestimations is easily observed in Figure 11 as the forecast of implied volatility is, beside from the beginning of the period, almost constantly above the realized volatility.

The overestimation of the implied volatility could have an explanation in the human psychology. Because of the financial crisis that occurred in the beginning of the backtesting period, the risk of that happening again in the close future could have been overestimated by the consensus of the financial market, which effects the option prices that the implied volatility is calculated from. The psychology of the human behavior in the field of finance is beyond this thesis but worth mentioning. Especially when comparing the estimations for the Black-Scholes model with EWMA and GARCH during the couple of years after the financial crisis. The estimations for EWMA and GARCH followed the realized volatility as it went down. However, the implied volatility did not follow the decline in the realized volatility and thus gave an overestimation.

The Vanna-Vomma model perform well beside from an extremely high outcome in the upper tail when the maturity is 1 month, but the mean value is still decent. For the maturity of 2 months, the mean value is the second best and when the maturity is 3 months, the mean value is the best compared to all the other models. The standard deviation of the differences is however not that good, as it is the highest for all the models and maturities.

The results for EWMA and GARCH are very similar but the results for GARCH are slightly better at all performance measures. This is expectable as the EWMA model only has one parameter, λ that is constant for every simulation but the parameters for GARCH, α, β and ω are estimated before every forecast of volatility. The parameter estimations for the GARCH model comes at a cost of time. The time required for the estimations can be observed in Tables 8, 9 and 10 in Appendix C, where the results from the long backtesting from 1995 is presented. The time to run the EWMA forecast takes about half the time compared to the GARCH forecasts.

One interesting note is the fact that for EWMA and GARCH does the estimations tend to get better as the maturities increase. Especially when looking at the mean value, standard deviation and the lower tail where the results get better when the maturities get longer. This might be because a severe change in the realized volatility in the short run is not captured by the models but in the long run, these changes evens out. The increase in performance as the maturity gets longer is however small, but noticeable.

As a conclusion, for the models described in this thesis, the best one for calculating the volatility strike price for a volatility swap is the GARCH model. The EWMA model can be used instead as the results are very similar but the upside is the easier implementation compared to the GARCH model.

The Vanna-Vomma model yields a very good result for calculating the volatility strike price. However, the lack of extensive data and one extreme outlier in the upper tail for the maturity of 1 month, one can not fully conclude that the model can be used for calculating the volatility strike price for a volatility swap.

7.2 Limitations and extensions

The first limitation is the lack of extensive option data. For the Vanna-Vomma model have only option data of one year been used. To test the performance of the model thoroughly one would preferably use option data from several years.

The result is limited to the models described in this thesis. However, there exist several other models that can be used to forecast volatility. One example is the Merton Jump Diffusion model. Because of assumptions in the Black-Scholes model such as a constant volatility, constant interest rate and that the underlying stock price follows a random walk (or geometric Brownian motion), the model is a simplification of the reality (Merton, 1975). To better reflect the reality of assets' price movements when estimating the fair volatility/variance strike, one can use a model with a jump process for the underlying asset's price, which is what the Merton Jump Diffusion model does.

Another model that can be used for further research is the Heston model which is the model that the Vanna-Vomma model is compared to in Rolloos and Arslans article (2017). For the Monte Carlo simulations when the volatility forecasting is made with EWMA and GARCH, the random variable multiplied with the past volatility is assumed to be normal distributed. An extension there is to compare the results if the random variable is instead Student-t distributed.

For the GARCH model there are several variants of extensions, such as Student-t GARCH, APARCH, EGARCH and GJR-GARCH (Danielsson, 2011), that also can be used for forecasting volatility and for further research.

Another example of extension that would be interesting to investigate is to estimate volatility strike prices when the swaps have longer maturities. The comparison is made of maturities of 1, 2 and 3 months and the backtesting is done on a maturity of 6 months. Appendix C consists of the evolution of the EWMA and GARCH forecast when the maturity is 12 months, but the performance of the forecasts with maturities even longer might also be interesting to investigate.

The forecast are done on the S&P500 but other indexes or assets could be used instead. It would be especially interesting to do a comparison of the models' performance with stock indexes and on other asset classes, such as gold, oil and other commodities and maybe even on currencies.

7.3 Outlook for volatility derivatives

Volatility is a common risk measure for an asset but it has evolved into a much greater use than to only measure the risk of a security. The usage and number of trades of volatility derivatives has since its introduction to the financial markets increased significantly. The usage of volatility derivatives and the quantity of trades in the future is likely to increase, which will require more and more accurate models for pricing them, if the financial markets is to remain stable and for avoiding another financial crisis.

Perhaps will one of the mentioned models in this thesis be the most widely used when estimating a volatility or variance strike. The Vanna-Vomma model has the major upside that is easy to implement and does not require any heavy computations, which is a factor that should be advantageous when the quantity of trades increase.

One can only speculate on how the development of volatility models will be in the future. With the theory and knowledge about financial markets today together with the technical development in computing power, the quantitative analyst of tomorrow will surely develop new volatility models. However, one thing is for certain, the place for volatility derivatives in the finance industry is here to stay.

References

Adu-Gyamfi, M. (2016). The Bankruptcy of Lehman Brothers: Causes, Effects and Lessons Learnt. *Journal of Insurance and Financial Management*, 1(4), 132-149.

Alexander, C. (2008). *Market Risk Analysis III - Pricing, Hedging and Trading Financial Instruments*. John Wiley & Sons, Ltd

Björk, T. (2009). *Arbitrage Theory in Continuous Time*. Oxford University Press.

Black, F. & Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 81(3), 637-654.

Bossu, F., Strasser, E. & Guichard, R. (2005). *Just What You Need To Know About Variance Swaps*. JP Morgan.

Broadie, M., Jain, A. (2008). The Effect of Jumps and Discrete Sampling on Volatility and Variance Swaps. *International Journal of Theoretical and Applied Finance*, 11(8), 761-797

Brockhaus, O. & Long, D. (2000). Volatility Swaps Made Simple. *Risk Magazine*, 2(1), 92-95.

CBOE. (2014). *The CBOE Volatility Index - VIX*.

Christensen, B. J., Prabhala, N. R. (1998). The relation between implied and realized volatility. *Journal of Financial Economics*, 50(2), 125-150.

Danielsson, J. (2011). *Financial Risk Forecasting*. John Wiley & Sons, Ltd.

Demeterfi, K., Derman, E., Kamal, M., & Zou, J. (1999). *More Than You Ever wanted To Know About Volatility Swaps*.

Flemming, J. (1998). The quality of market volatility forecasts implied by S&P 100 index option price. *Journal of Empirical Finance*, 5, 317-345.

Haug, E. G. (2003). Know Your Weapon, Part 1. *Wilmott Magazine* May, 49-57

Hull, J. C. (2012). *Options, Futures And Other Derivatives*. Eighth ed. Pearson.

Hull, J & White, A. (1987). The pricing of options on assets with stochastic volatility. *Journal of Finance* 42(2), 281-300.

J.P. Morgan. (1996). *RiskMetrics Technical Document*.

Rolloos, F. & Arslan, M. (2017). Taylor-Made Volatility Swaps. *Wilmott Magazine*, 2017(87), 56-61.

Romano, M. & Touzi, N. (1997). Contingent claims and market completeness in a stochastic volatility model. *Mathematical Finance*, 7, 399-410.

Schürger, K. (2002). *Laplace Transforms and Suprema of Stochastic Processes*. University of Bonn.

Willard, G. A. 1997. Calculating prices and sensitivities for path-independent derivative securities in multifactor models. *Journal of Derivatives*, 5(1), 45-61.

Appendices

A Tables of the comparison of the models from done by data from year 2015

The result for the Vanna-Vomma model using option data from 2015 are depicted in Table 3.

Table 3: *Vanna-Vomma results from 2015. Differences of the realized volatility from the estimations of volatility.*

T (Months)	Mean	Standard deviation	Upper tail	Lower tail
1	0.0150	0.1920	0.5961	-0.1829
2	-0.0139	0.0658	0.0532	-0.1438
3	0.0018	0.0778	0.1888	-0.0987

The result for the BS model using data of VIX and VXV from 2015 are depicted in Table 4.

Table 4: *Black-Scholes results from 2015. Differences of the realized volatility from the estimations of volatility.*

T (Months)	Mean	Standard deviation	Upper tail	Lower tail
1	0.0516	0.0587	0.1143	-0.1021
3	0.0535	0.0429	0.1126	-0.0007

The result for the EWMA model using data from 2015 are depicted in Table 5.

Table 5: *EWMA results from 2015. Differences of the realized volatility from the estimations of volatility.*

T (Months)	Mean	Standard deviation	Upper tail	Lower tail
1	-0.0192	0.0699	0.0691	-0.2109
2	-0.0165	0.0628	0.0740	-0.1476
3	-0.0139	0.0553	0.0782	-0.1027

Table 6 presents the result from the GARCH model of the year 2015.

Table 6: *GARCH results from 2015. Differences of the realized volatility from the estimations of volatility.*

T (Months)	Mean	Standard deviation	Upper tail	Lower tail
1	-0.0135	0.0653	0.0554	-0.1965
2	-0.0071	0.0555	0.0545	-0.1244
3	-0.0026	0.0454	0.0540	-0.0714

B Backtesting for BS, EWMA and GARCH from 2008 when T=6 months

Table 7: *Result from the backtesting from 2008. Differences of the realized volatility from the estimations of volatility.*

Model	**Mean**	**Standard deviation**	**Upper tail**	**Lower tail**	**Time to run (s)**
BS	0.0559	0.1017	0.1961	-0.3311	
EWMA	-0.0114	0.1060	0.2598	-0.3717	157.7411
GARCH	0.0032	0.0959	0.1769	-0.3656	289.2386

C Backtesting for EWMA and GARCH from 1995.

Table 8: *EWMA and GARCH backtesting from 1995 when T=1 month*

Model	Mean	Standard deviation	Upper tail	Lower tail	Time to run (s)
EWMA	0.0006	0.0674	0.1527	-0.2199	133.3611
GARCH	0.0056	0.0662	0.1488	-0.2216	447.9714

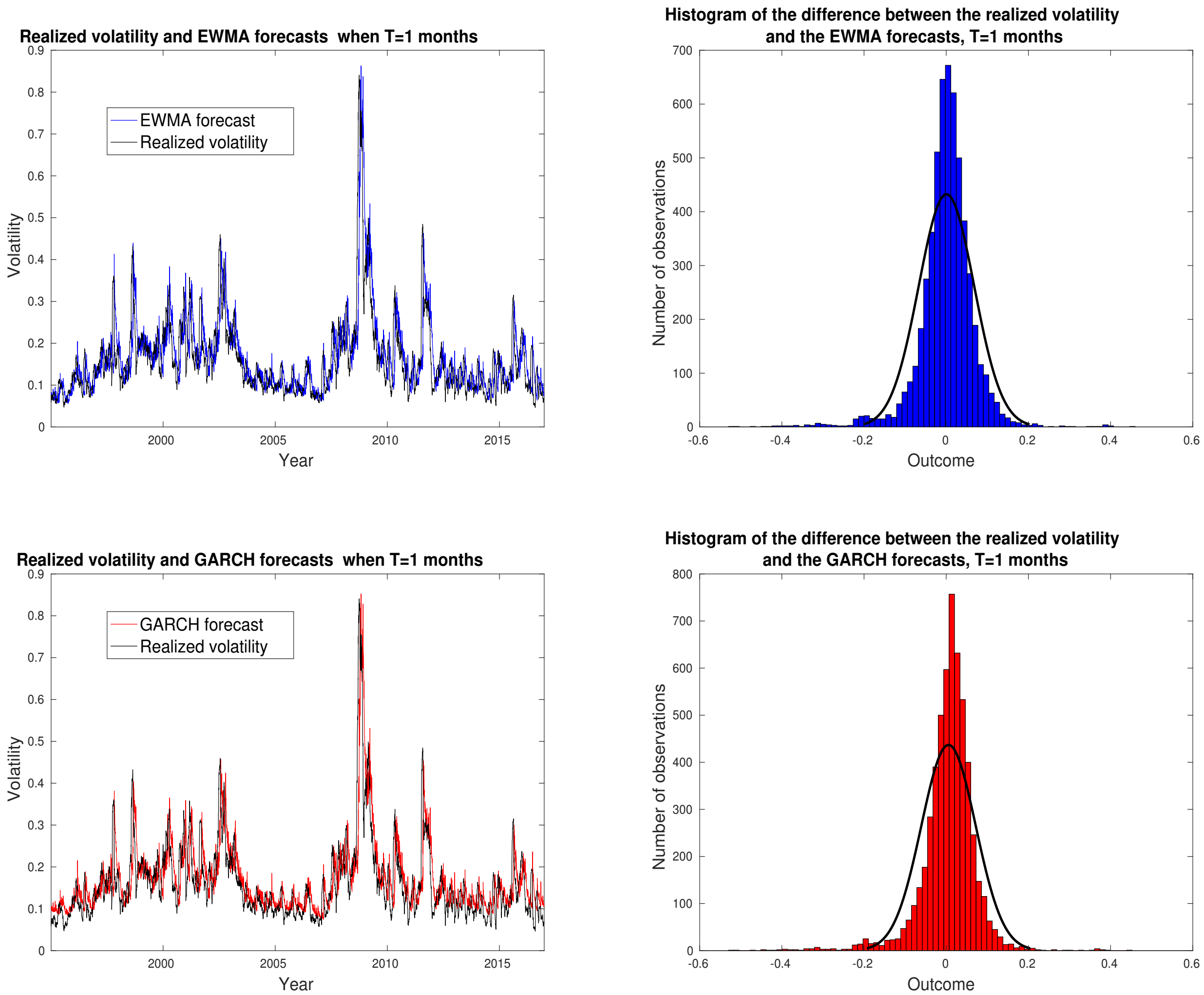

Figure 12: *Evolution of EWMA and GARCH forecasts together with the realized volatility and histograms of the differences between the forecasts and the realized volatility. T=1 month.*

Table 9: EWMA and GARCH backtesting from 1995 when T=3 months.

Model	Mean	Standard deviation	Upper tail	Lower tail	Time to run (s)
EWMA	-0.0065	0.0731	0.1655	-0.2292	238.3521
GARCH	0.0054	0.0703	0.1473	-0.2283	546.2230

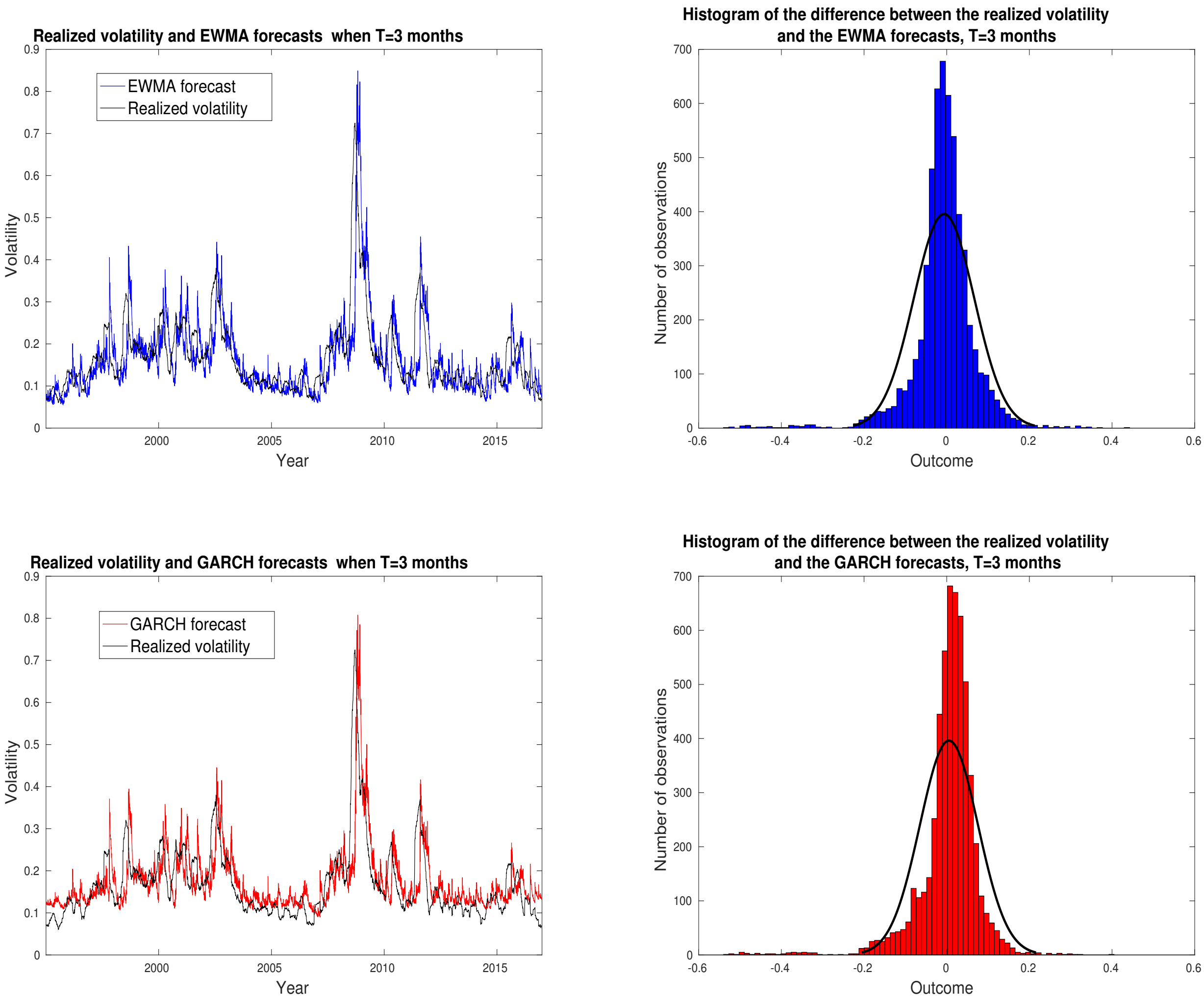

Figure 13: *Evolution of EWMA and GARCH forecasts together with the realized volatility and histograms of the differences between the forecasts and the realized volatility. T=3 months.*

Table 10: *EWMA and GARCH backtesting from 1995 when T=12 months.*

Model	Mean	Standard deviation	Upper tail	Lower tail	Time to run (s)
EWMA	-0.0253	0.0826	0.2188	-0.2761	668.5762
GARCH	0.0026	0.0743	0.1485	-0.2670	973.0355

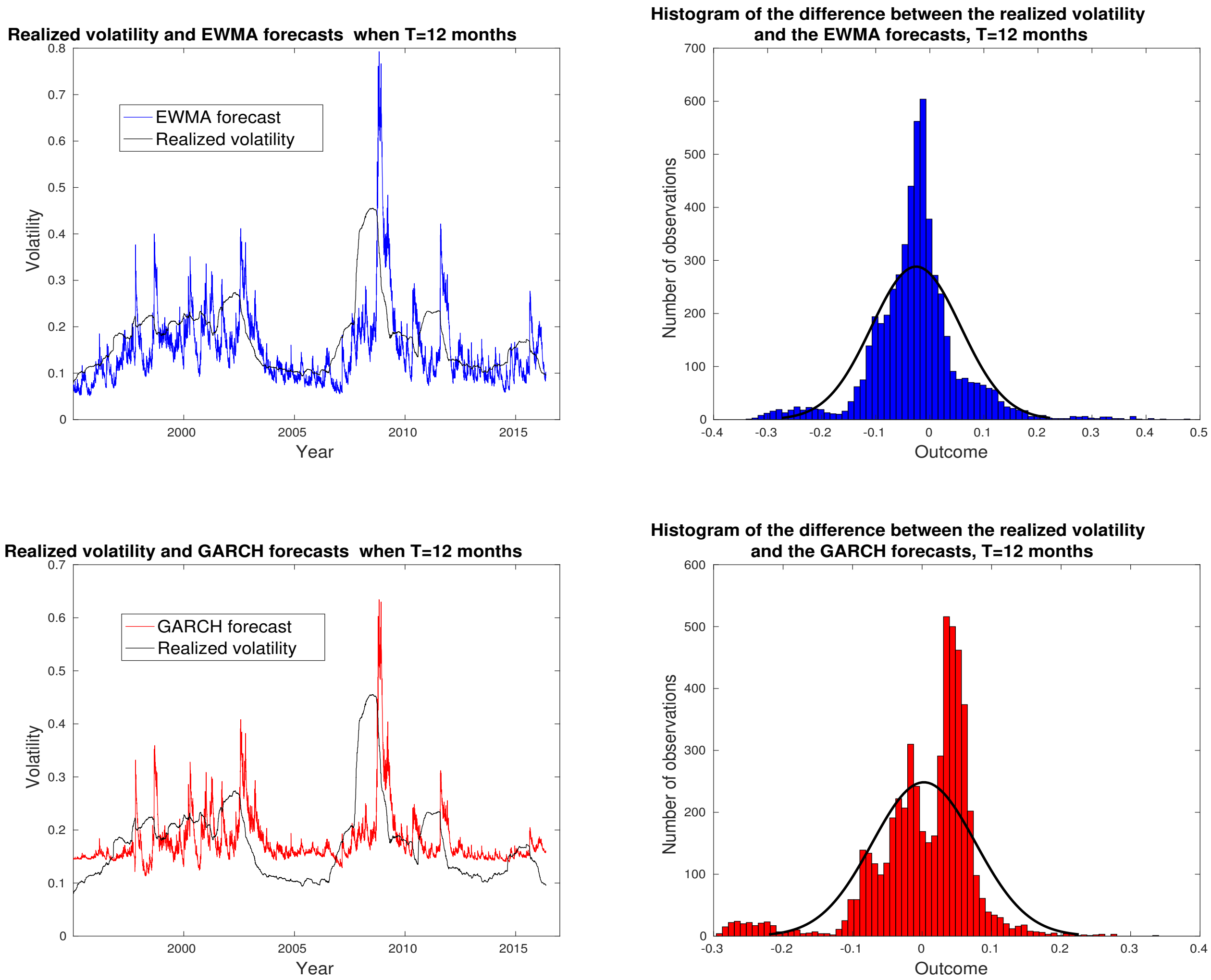

Figure 14: *Evolution of EWMA and GARCH forecasts together with the realized volatility and histograms of the differences between the forecasts and the realized volatility. T=12 months.*