UPPSALA
UNIVERSITET

U.U.D.M. Project Report 2015:15

Volatility Derivatives – Variance and Volatility Swaps

Joakim Marklund and Olle Karlsson

Examensarbete i matematik, 30 hp
Handledare och examinator: Erik Ekström
Juni 2015

Department of Mathematics
Uppsala University

Abstract

We give a comprehensive overview of volatility derivatives including the history behind it, the applications as well as pricing procedures in various models. Given these models we also apply market data to approach some empirical evidence, estimating and evaluating the performance of the models in the frame of variance and volatility swaps.

Contents

1 Introduction

In terms of finance volatility has always been considered a key measure. The development and growth of the financial market over the last centuries have caused the role of volatility to change. Instead of being just a component in pricing theory it has evolved into an asset class of its own. Volatility derivatives in general is a specialized financial tool which gives the opportunity to trade on the volatility of an underlying asset. Several derivatives have been created which puts great emphasis on volatility, providing a direct exposure to one of the most common measures of risk. Ever since the mid-1990s, securities such as variance/volatility swaps as well as futures and options have provided a good approach for investors to trade future realized variance or volatility against the current implied counterpart.

Who trades in volatility? For the same reason as stock investors tries to predict the movements of the stock market or bond investors think they can foresee the direction of interest rates, one may think they know something about the future volatility levels. If one thinks current volatility is high there are several derivatives in which one can take a position which profits if the volatility decreases.

1.1 Purpose and Goals

Our purpose for this thesis is to do an extensive study in the financial area known as volatility derivatives, and apply some of the most popular models to these derivatives. Mainly focused around the variance and volatility swap, we include the standard Black-Scholes model, the Merton Jump Diffusion model as well as Stochastic Volatility Models such as GARCH, Heston and 3/2 in order to give a comprehensive study on the evaluation and performance of these swaps.

The goal is to analyze and compare the different models in a volatility based setting. We will use empirical studies based on options as well as historical data on the famous S&P500 index and translate the calibrations into our models, where we can explicitly compare them against each other.

1.2 Structure of Thesis

In this paper we will present multiple trading tools which emphasises on volatility. We start out in chapter 2 by going over the history of volatility derivatives, giving a brief overview of the growth and development. Moving on we will in chapter 3 derive and present some of the easier contracts used when trading volatility, namely the variance and volatility swaps. We will also price and evaluate a variance swap in a semi model-independent setting.

Options play a significant role in finance, and will naturally have many useful purposes even in volatility trading. The famous model for option pricing which can generate closed-form formulas for options, The Black-Scholes model, have been shown to have certain drawbacks in the way they simplify the reality. One way to find better empirical support have been to allow jumps in the model, and we will in chapter 4 look into these effects and variants. To relax the assumption about constant volatility the usual answer is applying a stochastic volatility model. This popular area has a lot of different models which can be applied, and we will derive a couple of them. Thus we will in chapter 5 explore the realm of stochastic volatility models.

Looking at several models throughout the paper we will towards the end at chapter 6 and 7 present some empirical evidence in order to fully compare and evaluate the differences between the models alongside some relevant analysis. Finally we conclude in chapter 8 and 9 with our findings and reflect on some possible future research and extensions.

2 History

2.1 Option Pricing and Modelling

One of the main reasons that financial mathematics have become an interesting subject in the world today can be explained by two words, option pricing. The concept of buying or selling various commodities in order to make money off of it has always been a compelling aspect to humans. In our contrast the important question is whether or not the market can determine a unique price for every given option, and is this price explicitly computed? The ones to come up with a valid answer to this question was Nobel prize winners Robert Merton and Myron Scholes in the 1970s (together with Fischer Black, who unfortunately had passed away when Merton and Scholes were awarded the Nobel prize). Their model, the famous Black-Scholes model, is the market convention when referring to standard option valuations. This model can be put into an explicit formula to determine the price of European options.

Although innovative and very elegant at the time, the Black-Scholes-Merton model has been criticized because of its limitations and possible defects. The model in itself is made fairly simple, and only provides a fair approximation to the real world. But with simplicity comes not only easy implementations, but also certain drawbacks. Some of the assumptions made, the simplifications of reality, might not be fully accurate and the model does not always empirically support the market. Assumptions made such as efficient markets, liquidity or no transaction costs are hard to relax, but some of the assumptions made in the B-S model could be taken off by doing some extensions.

The continuity of the stock returns are aimed to behave in a nice continuous way dictated by the geometric Brownian motion, something that has been questionable. Empirically the stock returns have exhibited jumps, meaning that they over a small time period can experience drastic changes. To deal with this phenomenon the solution have been to introduce a jump process into the original model, where the most famous extension being the Merton Jump Diffusion model.

One assumption that have been questioned and criticized include that we have a constant volatility (σ). To be able to relax this assumption one often moves over to stochastic volatility models. The difference being that we no longer assumes constant volatility, but that it follows a random process given by some dynamics. Examples of such models that have been seen to give better empirical support include the Heston model or the GARCH (Generalized Autoregressive Conditional Heteroskedasticity) model.

Another important assumption of the B-S model is that we have Gaussian log-increments of the stock price. Studies over the years of stock prices have shown that the usage of the Gaussian model is incompatible due to evidence of heavy tails empirically, which suggests that it might be more reasonable to replace the Brownian motion with a more general family that removes the faulty assumptions, for example a Lévy process. [1]

2.2 Volatility Trading

Volatility derivatives started to appear in the late 20th century. At this time the contract that first saw its light was a variance swap, and were dealt at the UBS investment bank in Switzerland in 1993 by Michael Weber. The variance swaps were rather illiquid during the first years, but increased heavily after 1998 and has ever since the millennia become a credible trading tool. Thanks to the development of the replicating argument which involves using a portfolio of vanilla options to properly replicate a variance swap, the market for volatility grew even larger. [8]

Although the popularity of these volatility derivatives have evolved since its introduction and is nowadays seen as an asset class in it owns right and a tradable market instrument, there has been downfalls. The Wall Street Journal reported during the financial crisis in 2008 that the extreme volatility levels almost killed the market for volatility swaps. [37] The unforeseen extremes made traders unable to hedge volatility reliably. In this period of time the liquidity became very scarce and trading in this niched area diminished significantly and was reduced to "trade by appointment".

A famous volatility symbol is the Volatility Index (denoted VIX), a volatility index operated by the Chicago Board Options Exchange (CBOE). Introduced in 1993, this volatility index have been a popular measure of the implied volatility, and measures expectations of volatility over 30 day periods. It started off by replicating the one-month implied volatility on the S&P100 index, and expanded in 2003 to measure market expectations of volatility conveyed by the S&P500 stock index. The VIX index has since its introduction been considered the world's benchmark for stock market volatility. The introduction of this index has laid a good foundation for the development and interest in volatility products and speculating in volatility derivatives. Over the years the CBOE has launched a futures exchange (CFE) as well as allowing trades on VIX options to enlarge the family of volatility derivatives. The figure below maps the evolution of the VIX index, an index also known as the "fear index" or "risk index".

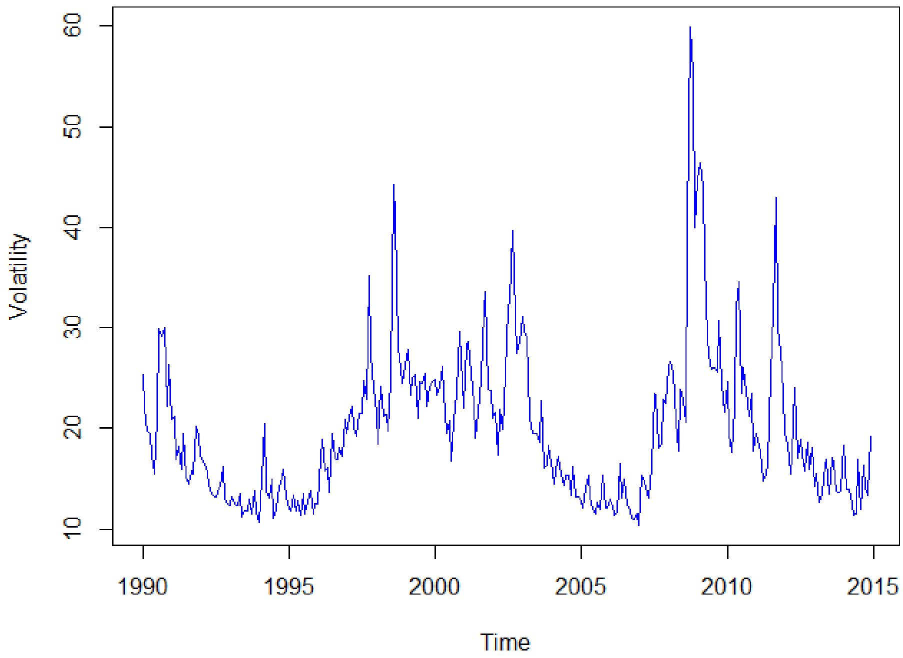

Figure 1: Historical evolution of the VIX index [10]

As time have passed and volatility have returned to normal levels, the developments to this niched area of finance has continued to grow. Over the past years even more types of volatility derivatives have emerged, and comparing to the standard swaps we can today observe several other tools such as gamma swaps, corridor variance swaps and conditional variance swaps etc.

3 Volatility Derivatives

Volatility derivatives are a special type of financial derivative where the payoff will depend on the volatility of an underlying asset. Some of the most common examples in this area are variance and volatility swaps or options written on the volatility index, VIX.

Trading in volatility derivatives have become popular in risk management due to its direct exposure to the volatility. Thus with these types of financial derivatives volatility is now a tradable market instrument. In the past traders would normally use a delta-hedged position to trade with volatility. This however does not perfectly trade with volatilities alone, since the return also is dependent of the underlying stock price. The introduction to variance and volatility swaps have provided a pure exposure to volatility, making it popular when interested in these circumstances. [26]

Why trade in volatility over regular financial derivatives? Volatility as a trading tool have several characteristics which makes trading it attractive. Although there is technically possible to achieve a variance swap payoff with vanilla options, a natural question becomes why should one be encouraged to trade in volatility compared to other financial tools. The easiest answer is that a these tools offers direct exposure to volatility, removing the path dependency issues that comes when delta-hedging an option. Trading in regular options the investor has to both be aware of the underlying price as well as how the volatility changes, something that a pure volatility trader doesn't. Some of the most common reasons are listed below:

- **Speculative position:** Here the investor takes a directional position in the volatility of an underlying. If one believes that current volatility levels are incorrect or that the expectation of future volatility is high/low one can utilize volatility derivatives to make a profit. Examples of such situations include political or financial turmoil due to current debt issues or belief in changes due to a forthcoming election.
- **Hedging position:** Looking at a hedging perspective there is a variety of industries that trades in volatility as part of their portfolio. Volatility is often found to be negatively correlated with a stock or index level, making hedging in volatility a good diversification strategy. During market crashes volatility is generally known to increase, an phenomenon knows as Black's leverage effect [4], which makes hedging in volatility during those times a means to reduce losses.

3.1 Swaps

Introduced in the 1980s, a swap is an agreement between two counterparts to exchange cash flows in the future. The agreement will define which dates the cash flows will be paid as well as how they will be calculated. One party agrees to pay a fixed amount to the other counterpart which in return accepts the agreement by paying a floating amount, an amount which depends on the underlying. A swap is an OTC product (Over-The-Counter), and the most common types of swaps are for example *interest rate swaps* and *currency swaps.* The focus of this thesis will be on the swaps written on variance and volatility. These swaps allows the investor a direct position in volatility or variance of a stock index or a specific stock. These type of swaps generally only have a single payment at expiration, which is not always the case when dealing in other swaps.

3.2 Variance Swaps

A forward contract where one counterparty agrees to pay a notional amount, call it N, times the difference between a fixed level of variance and a realized amount. For variance swaps the fixed level is usually called variance strike price. The realized variance is determined by calculating the variance of the assets return over the lifespan of the swap.

To calculate the realized variance we need a couple of variables, which includes the observation frequency of the price of the underlying, the annualization factor and the method of calculating the variance. It is important that the procedure of calculating these variables are clearly specified before entering a swap. For example whether the standard deviation of returns is calculated by assuming a zero mean or not. When dealing with discrete sampling we create a partition $0 = t_0 < t_1 < ... < t_n = T$ of the time interval $[0, T]$, thus creating n equal segments with lengths Δt ($t_i = iT/n$). Given all this information we define the realized variance in these types of contracts by the formula

$$V_d(0, n, T) = \frac{AF}{n-1} \sum_{i=0}^{n-1} \Big(\log \Big(\frac{S_{i+1}}{S_i} \Big) \Big)^2. \tag{1}$$

Here n is the number of return observations, S_i is the price of the asset at time t_i and AF stands for the annualization factor. AF, defined by n/T, would be 252 if the maturity of the swap were 1 year with daily sampling ($T = 1, n = 252$). Thus the variance swap payoff is defined as

$$(V_d(0, n, T) - K_{var}) \times N \tag{2}$$

where $V_d(0, n, T)$ is defined as above, N is the notional amount and K_{var} is the variance strike.

3.2.1 Example of a Variance Swap

To get a better grasp of how the variance swap works, we proceed with a simple made up example. Worth mentioning before we start is that the variance strike price is usually quoted in units of volatility squared, e.g. $(15\%)^2$.

Suppose that we (the trader) takes a long position in a variance swap where the strike price K_{var} is calculated to have a volatility of 15%, or in terms of variance $(15\%)^2 = 0.0225$. We say that the notional amount N is $1,000,000$ dollars. Over the life of the contract we can then use the formula above (1) and calculate the realized variance V, which given the parameters in our imaginary contract then turned out to be $(20\%)^2 = 0.04$. The various actions are illustrated in a picture below:

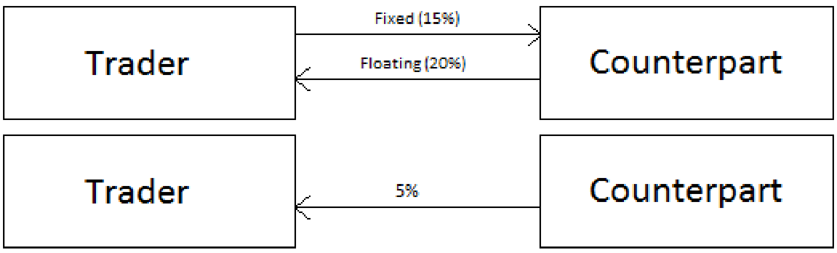

Figure 2: Example of a variance swap

As we can see we payed a fixed variance of $15\%^2$ at the start of the contract, and received at the end of the time period a realized variance of $20\%^2$. Thus using the payoff function (2) we see that the long position yielded a profit of $(0.04 - 0.0225) \times 1,000,000 = \$17,500$.

3.2.2 Pricing and Valuation

Going back to our example we notice that the question that arises is how to properly calculate the variance strike price. We already have a formula for the realized variance, and are looking for a proper way to accurately calculate K_{var}. In this section we present an approach to find a closed formula for the variance strike price which does not rely on any complex models and can be determined by pricing vanilla call and put options.

Considering a simple model similar to the classic Black-Scholes model, consisting of the same assumptions with the difference being the characteristics of the underlying. The only assumption made on the asset is that it is continuous and does not contain any jumps. Later on in this paper we will look into the effects and changes when we allow for jumps in the stock prices.

For now the asset takes the following characteristics:

$$dS_t = \mu\left(t, ...\right) S_t dt + \sigma\left(t, ...\right) S_t dW_t. \tag{3}$$

Here, we assume that the drift μ and the volatility σ are arbitrary functions of time and other parameters. These assumptions include, but are not limited to, models in which the volatility is a function of stock price and time only ($\sigma\left(t, S_t\right)$). While we denoted the discrete realized variance as $V_d(0, n, T)$ we will naturally let the continuous sampled realized variance be denoted by $V_c(0, T)$. This is a theoretical element representing the average combined fluctuations of the volatility and is defined as

$$V_c(0, T) = \frac{1}{T}\int_0^T \sigma^2\left(s, ...\right) ds. \tag{4}$$

The discrete realized variance defined in the previous section, or the so called floating leg of the swap, will in the limit approach the continuous realized variance. This means that we have the relationship

$$V_c(0, T) = \lim_{n\to\infty} V_d(0, n, T). \text{ [26]} \tag{5}$$

To justify this relation we begin by looking at $V_d\left(0, n, T\right)$ (recall equation (1)), and substituting the solution $S_t = S_0 \exp\left(\int_0^t \left(\mu - \frac{\sigma^2}{2}\right) ds + \int_0^t \sigma dW_s\right)$ into the discrete realized variance. Recall that we also have that $AF = n/T$. This yields that

$$\begin{aligned} &\lim_{n\to\infty} V_d(0, n, T) = \\ &\lim_{n\to\infty} \frac{n}{T\left(n-1\right)} \sum_{i=0}^{n-1} \left(\int_{t_i}^{t_{i+1}} \left(\mu - \frac{\sigma^2}{2}\right) ds + \int_{t_i}^{t_{i+1}} \sigma dW_s\right)^2. \end{aligned} \tag{6}$$

Here we consider σ and μ to be simple processes, recall that if they are not simple we can approximate them with simple processes. Calculating the square above and evaluating we arrive at

$$\lim_{n\to\infty} \frac{n}{T\left(n-1\right)} \sum_{i=0}^{n-1} \sigma^2 \left(W_{t_{i+1}} - W_{t_i}\right)^2 + R_i = V_c(0, T). \tag{7}$$

Here we see that under the limit the first sum converges to our definition of $V_c\left(0, T\right)$ and the rest term R_i converges to zero since it consists of terms with $\sum_{i=0}^{n-1} a_i \left(t_{i+1} - t_i\right)^2$ and $\sum_{i=0}^{n-1} b_i \left(t_{i+1} - t_i\right)\left(W_{t_{i+1}} - W_{t_i}\right)$ where a_i and b_i are some constants.

Having defined the necessary tools for this model we start by looking at the corresponding price process of a variance swap $\Pi\left(t, N\left(V_c - K_{var}\right)\right)$. This process are solved using the Feynman-Kac formula, which gives a solution of the form

$$\Pi\left(t, N\left(V_c - K_{var}\right)\right) = E_t^Q\left[e^{-r(T-t)} N\left(V_c\left(t, T\right) - K_{var}\right)\right]. \quad (8)$$

As we assume a lack of arbitrage in the market we can conclude that at the time of signing the net worth of the contract has to be zero. This assumption thus implies that the continuous variance strike price must satisfy the equation

$$K_{var} = E_0^Q\left[V_c\left(0, T\right)\right] = E_0^Q\left[\frac{1}{T}\int_0^T \sigma^2 dt\right]. \quad (9)$$

Thus we have found a valid approach of what the variance strike should be. This approach is a good idea for valuating the contract but does not however give a good interpretation in how to properly calculate it. In order to replicate the contract we then look into a method involving a rebalancing strategy of a stock as well as a log contract, a method which caused a big outburst in the area of volatility derivatives.

3.2.3 Replicating Approach: First Steps

To start off the replicating strategy we use Itô's lemma on a theoretical log contract of the underlying asset which yields the stochastic differential equation

$$d\left(\log S_t\right) = \left(\mu - \frac{\sigma^2}{2}\right) dt + \sigma dW_t. \quad (10)$$

In order to achieve a plausible result we combine $1/S_t$ contracts of the asset as well as being short in the aforementioned log contracts we arrive at the very useful equation

$$\frac{dS_t}{S_t} - d\left(\log S_t\right) = \frac{\sigma^2}{2} dt. \quad (11)$$

With the use of this technique we can see that the drift μ has been cancelled, which is what we were trying to obtain. Writing this as an integral equation from 0 to T multiplied with the constant $1/T$ and rearranging we get a new formula for the theoretical realized variance

$$\frac{1}{T}\int_0^T \sigma^2 dt = \frac{2}{T}\left(\int_0^T \frac{dS_t}{S_t} - \log\left(\frac{S_T}{S_0}\right)\right). \quad (12)$$

Recall here that the variance strike price K_{var} was equal to the expected value of the theoretical realized variance in equation (9). Thus to proceed we have to calculate the expected values of these contracts in order to get a solution formula for K_{var}. This means that our variance strike will have the form:

$$K_{var} = \frac{2}{T} E\left[\int_0^T \frac{dS_t}{S_t} - \log\left(\frac{S_T}{S_0}\right)\right]. \quad (13)$$

Thus to solve this we need to calculate the respective expected values. First we note that as we are working with a risk neutral measure the asset will take a slightly new form as

$$dS_t = rS_t dt + \sigma\left(t, ...\right) S_t dW_t. \tag{14}$$

Here r stands as usual for the risk free interest rate, and will be assumed constant. Given these characteristics we are now ready to calculate the expected values. For the first set of contracts we can easily evaluate this and through standard calculations we have

$$E_0^Q\left[\int_0^T \frac{dS_t}{S_t}\right] = rT. \tag{15}$$

Next up we need to replicate the log contract, which is a bit harder and can be done by statically using vanilla call and put options. We do this by letting f be a twice differentiable function which will represent a payoff of a European style path-independent derivative security. Such a function can be expressed in accordance of Breeden and Litzenberger (1978, [5]) as:

$$\begin{aligned} f(S_T) = {} & f(x) + f'(x)(S_T - x) \\ & + \int_0^x f''(K)(K - S_T)^+ dK + \int_x^\infty f''(K)(S_T - K)^+ dK. \end{aligned} \tag{16}$$

The formula above can be obtained by using integration by parts on

$$f(s) - f(x) = \int_x^s f'(t)dt = f'(x)(s - x) + \int_x^s f''(t)(s - t)dt. \tag{17}$$

Rewriting the last integral with an indicator function and dividing it into two different integrals makes us arrive at the required formula. Then the payoff function f could be replicated by holding different positions on a zero coupon bond (with face value $f(x)$), a forward contract with strike price x as well as call and put options. As we are interested in the zero value of the claim, we can express this in terms of standard European call and put options denoted $C(K)$ and $P(K)$ respectively with maturities at time T and strike price K. Calculating the expected values at time zero and discounting for our function $f(S_T)$ we have

$$\begin{aligned} E_0^Q[e^{-rT} f(S_T)] = {} & e^{-rT} f(x) + f'(x)[C(x) - P(x)] \\ & + \int_0^x f''(K)P(K)dK + \int_x^\infty f''(K)C(K)dK. \end{aligned} \tag{18}$$

With this equation we just need to transfer back the payoff function f into our log contract in order to fully replicate the variance swap. Letting $f(S_T) = \log(S_T/S_*)$ and $x = S_*$, where $S_* > 0$ is some arbitrary cutoff in order to separate the call and put options, as well as substituting this into equation (16) we get:

$$\log\left(\frac{S_T}{S_*}\right) = \frac{S_T - S_*}{S_*} - \int_0^{S_*} \frac{1}{K^2}(K - S_T)^+ dK - \int_{S_*}^{\infty} \frac{1}{K^2}(S_T - K)^+ dK. \quad (19)$$

One important note here is that the log contract above is not the exact same as the one we looked at in equation (12). It should be noted that we can rewrite the log payoff as

$$\log\frac{S_T}{S_0} = \log\frac{S_T}{S_*} + \log\frac{S_*}{S_0}. \quad (20)$$

The first term in the equation above is the one we have replicated in (19), and the second one is a constant term which is independent of the final stock price S_T. Given that we have replicated the first term and the second term is constant we are now ready to move forward and put everything together.

3.2.4 Replicating Approach: Final Steps

Up to this point we have seen how to handle both the rebalanced hedge of $1/S_t$ and the log contract. To finish off the replicating strategy we need to put together both those terms. What we have obtained without expectations is

$$\begin{aligned} V_c(0,T) = \frac{2}{T}\Bigg[& \int_0^T \frac{dS_t}{S_t} - \frac{S_T - S_*}{S_*} - \log\left(\frac{S_*}{S_0}\right) + \\ & + \int_0^{S_*} \frac{1}{K^2} P(K)dK + \int_{S_*}^{\infty} \frac{1}{K^2} C(K)dK \Bigg]. \end{aligned} \quad (21)$$

Going back to equation (9) we need to calculate the expected values of the above equation to obtain a solution formula for the variance strike price, K_{var}. This yields that

$$\begin{aligned} K_{var} = \frac{2}{T}\Bigg[& rT - \left(\frac{S_0}{S_*}e^{rT} - 1\right) - \log\left(\frac{S_*}{S_0}\right) + \\ & + e^{rT}\int_0^{S_*} \frac{1}{K^2} P(K)dK + e^{rT}\int_{S_*}^{\infty} \frac{1}{K^2} C(K)dK \Bigg] \end{aligned} \quad (22)$$

where S_0 is the initial price of the underlying, S_* is the positive arbitrary cutoff and $C(K)$ as well as $P(K)$ are vanilla European call and put options with strike prices K. One interesting note is that if we were to choose the cutoff as $S_* = S_0 e^{rT}$ we could reduce the equation of the variance strike above into a simpler form of

$$K_{var} = \frac{2e^{rT}}{T}\left[\int_0^{S_0 e^{rT}} \frac{1}{K^2} P(K)dK + \int_{S_0 e^{rT}}^{\infty} \frac{1}{K^2} C(K)dK\right]. \quad (23)$$

3.2.5 Discrete Approximation

As we saw earlier we can replicate a variance swap by a portfolio of vanilla options, and we didn't require any particular model assumption to determine the variance strike price. For a discrete approximation of the replicating scheme the only complicated task resides in how to compute the call and put options, which means that we have to find a reasonable approach to replicate the log contract. For this approximation we have chosen so that the cutoff is equal to the initial stock price ($S_* = S_0$).

Following the frameworks of Demeterfi et al. (1999, [13]), we will check if our initial approach can match the reality which only requires the usage of a discrete set of options. To properly replicate the aforementioned log contract in terms of standard European options we remind ourselves with the definition of the variance strike by equation (13), which can be rewritten as

$$K_{var} = \frac{2}{T} E\left[\int_0^T \frac{dS_t}{S_t} - \frac{S_T - S_*}{S_*} - \log\left(\frac{S_*}{S_0}\right) + \frac{S_T - S_*}{S_*} - \log\left(\frac{S_T}{S_*}\right) \right].$$

Taking expectations of the equations above we will have a variance strike price of

$$K_{var} = \frac{2}{T}\left[rT - \left(\frac{S_0}{S_*}e^{rT} - 1\right) - \log\left(\frac{S_*}{S_0}\right)\right] + e^{rT}\Pi_{CP} \tag{24}$$

Here we then have a term which can be instantly calculated through the choices in cutoff, time of maturity, rate etc. The function Π_{CP} will denote the present value of the portfolio with vanilla options, and has a payoff at expiration given by

$$f(S_T) = \frac{2}{T}\left[\frac{S_T - S_*}{S_*} - \log\frac{S_T}{S_*}\right]. \tag{25}$$

In practice we are only able to use a discrete set of option to replicate this, and we need to determine how to appropriately set this up. Lets assume that we are able to trade call options with strikes

$$K_0 = S_* < K_{1c} < K_{2c} < ... \tag{26}$$

as well as put options with strikes

$$K_0 = S_* > K_{1p} > K_{2p} > ... \tag{27}$$

By a piece-wise linear approximation we can determine how many options of each particular strike we will put into the portfolio. As an example, for a call option with strike price K_0 we would have

$$w_c(K_0) = \frac{f(K_{1c}) - f(K_0)}{K_{1c} - K_0}. \tag{28}$$

The following segment will then be a combination of strikes K_0 and K_{1c}:

$$w_c(K_1) = \frac{f(K_{2c}) - f(K_{1c})}{K_{2c} - K_{1c}} - w_c(K_0). \tag{29}$$

Continuing in a similar fashion one can build the entire payoff curve, one step at a time. The general formula for the weights for call and put options is thus

$$w_c(K_{n,c}) = \frac{f(K_{n+1,c}) - f(K_{n,c})}{K_{n+1,c} - K_{n,c}} - \sum_{i=0}^{n-1} w_c(K_{i,c}) \tag{30}$$

$$w_p(K_{n,p}) = \frac{f(K_{n+1,p}) - f(K_{n,p})}{K_{n,p} - K_{n+1,p}} - \sum_{i=0}^{n-1} w_p(K_{i,p}). \tag{31}$$

As seen above the principle for calculating put options is done in a similar fashion as the call options. Once we have calculated the weights from the framework above we can then obtain Π_{CP} from

$$\Pi_{CP} = \sum_i w(K_{ip})P(S, K_{ip}) + \sum_i w(K_{ic})C(S, K_{ic}). \tag{32}$$

Below we illustrate a numerical example where we approximate the price of a variance swap using this approach.

Type of Option	Strike	Volatility	Weight	B-S Price	Contribution
Put	105	24	36	0.0272	0.9792
Put	110	23	33	0.0650	2.145
Put	115	22	30	0.1102	3.6366
Put	120	21	27	0.2476	6.6852
Put	125	20	26	0.6392	16.6192
Put	130	19	23	1.2324	28.3452
Put	135	18	22	2.2498	49.4956
Put	140	17	10	3.8766	38.766
Call	140	17	9	5.5921	50.3289
Call	145	16	19	3.0585	58.1115
Call	150	15	17	1.3806	23.4702
Call	155	14	17	0.4764	8.0988
Call	160	13	15	0.1125	1.6875
Call	165	12	15	0.0154	0.231
Call	170	11	13	0.0010	0.013
				Total cost	**288.6129**

Table 1: Replication of a variance swap

We assume that the initial stock price level S_0 is 140, the riskless interest rate r is 5%, the stock pays no dividends (dividend yield = 0), and the maturity of the swap is 3 months ($T = 0.25$). We suppose that we can buy options with strikes ranging from 100 to 175, uniformly spaced apart by a range of 5. The option prices are calculated using standard Black-Scholes pricing methods. We also assume that the at-the-money implied volatility is 17% and that it is increased by 1 point for every 5 points of decrease in the strike price. The weights are calculated by the formulas (30) and (31), where the contribution is given by the calculated weights for each level times the individual options. The total sum of each option gives the total cost of the portfolio.

The value of options and weights are given in Table 1, and the total cost of replicating future variance yields that $\Pi_{CP} = 288.6129$, which then using equation (24) gives us that $K_{var} = (17.095)^2$. Since the assumed implied volatility was at 17% this gave us a fairly good approximation.

3.2.6 Limitations in Accuracy and Performance

Since this approximation does not fully capture the variance strike price, it is a good question to ask about the accuracy and performance of doing this in practice. How well will the replicating strategy work with limited data and will there be any practical problems?

The first obvious imperfection with this strategy is that we can't capture the variance strike exactly because of the limitations of using all strike prices of call and put options between 0 and infinity. As we are able to trade only a very limited range of strikes, we might obtain deceiving result that inaccurately replicates the log contract. Secondly the stock price may exhibit jumps which makes the log contract insufficient and will thus not capture the realized variance correctly.

In terms of accuracy the procedure of approximating a log contract will, as pointed out by Demeterfi el al. (1999, [13]), always over-estimate its value causing the value gained from this method to be higher than the true theoretical value. The discrete usage of different strike prices of our call and put options will over-estimate the variance strike, and including more options in our portfolio will sink down the variance strike towards its theoretical value. This means that we will have a convergence as ΔK (the spacing between strikes) goes to 0 towards the theoretical value, and so the procedure will always exceed or match the theoretical value.

3.3 Volatility Swaps

A volatility swap is a forward contract on future realized volatility. Although seemingly very similar to the variance swap, the volatility swap is more commonly traded in practice but also more theoretically complicated. The Financial Times in 2006 posted an article quoting a derivatives trader saying [20]:

"Variance is easier to hedge. Volatility can be a nightmare."

An assets volatility is a good way in measuring riskiness and uncertainty, and thus such a swap will provide a direct exposure to volatility making it an attractive choice. The volatility swap payoff lies close to the variance swap, and is defined as

$$\left(\sqrt{V_d(0,n,T)} - K_{vol}\right) \times N \tag{33}$$

where $\sqrt{V_d(0,n,T)}$ is the square root of the realized variance, i.e. equation (1). n is again the number of sampling dates and K_{vol} is the volatility strike. N refers to the notional amount of the swap.

3.3.1 Example of a Volatility Swap

To get a better feeling of the volatility swap works in practice, we again take a look at a very simple example. Going back with the same variables as in the variance swap case, we put the notional amount as N = 1,000,000 and we take a long position in this swap with a volatility strike K_{vol} of 15%. At the end of the contract we find that the realized variance V is calculated to be 20%. Using again our payoff function we can see that we end with a final profit of $(0.20 - 0.15) \times 1,000,000 = \$50,000$.

3.3.2 Pricing

As in the case of a variance swap, the pricing of a volatility swap starts out quite similarly. We will again define the fair continuous volatility strike as such as the present value of the contract at time zero is equal to zero. This corresponds to solving the equation

$$E_0^Q\left[e^{-rT}(\sqrt{V_c(0,T)} - K_{vol})\right] = 0. \tag{34}$$

Thus, as we recall from equation (9) with the variance strike we can express the volatility strike price as

$$K_{vol} = E_0^Q\left[\sqrt{V_c\,(0,T)}\right] = E_0^Q\left[\sqrt{\frac{1}{T}\int_0^T \sigma^2 dt}\right]. \tag{35}$$

In order to get a solution formula we again need to solve the volatility strike K_{vol}. Since we are dealing with the square root of the variance equation we can't use the same replicating strategy as we had in the variance swap section. A naive approach to solving this problem would be to use the same principle as in the variance swap and just use the square root right away, i.e.

$$K_{vol} = \sqrt{K_{var}}. \tag{36}$$

However doing this directly will most likely result in an error since we don't consider the convexity notion (or sometimes also known as convexity error) of the square root function. Thus because we are dealing with the square root of the variance, we cannot use a model-independent replication approach and thus we have to look for something different. Without having any specified model we will thus not get any proper formula for K_{vol}, but what we can do however is to create an upper bound with the help of the results from earlier.

To deal with the convexity of the square root we apply Jensen's inequality, which says that for a concave square root function we have that $E[\sqrt{x}] \leq \sqrt{E[x]}$. Applying this to the strike prices for both variance and volatility swaps, we obtain the following upper bound:

$$K_{vol} = E_0^Q[\sqrt{V_c(0,T)}] \leq \sqrt{E_0^Q[V_c(0,T)]} = \sqrt{K_{var}}. \tag{37}$$

Hence we can say that the volatility strike is bounded upwards by the square root of the variance strike price. This difference is usually called the convexity correction. Thus the choice on which model to be used when calculating the strike prices, the magnitude of the convexity will be highly dependent upon this choice. One approximation to derive the convexity error is to use a second order Taylor expansion made by Brockhaus and Long (2000, [6]). The result of the calculations is shown below, where the proof is left for the reader in Appendix B:

$$\sqrt{K_{var}} - K_{vol} \approx \frac{Var[V_c(0,T)]}{8\sqrt{K_{var}}^3}. \tag{38}$$

Worth noting is that the convexity correction term is based on the continuous realized variance, rather than the discrete one. A discrete version, i.e. a finite number of sampling dates n, will also perform a reasonable good approximation in a model-independent setup (using the replicating approach), as long as n is big enough.

As pointed out by Broadie and Jain (2008, [27]) the convexity correction formula will not work very well in a Heston stochastic volatility model or Mertons jump-diffusion, due to the fact that the Taylor expansion will have higher order terms that are not negligible. Thus in those models this approximation does not provide a good estimate for the fair volatility strike price. The error term is these cases will consist of 3rd and 4th order terms as well, making it more complex and perhaps not very suitable for receiving a proper approximation.

3.3.3 Laplace Transforms

An alternative approach in receiving the price of a volatility swap is to make use of Laplace transforms. Evaluating a square root function as a Laplace transformation and using the realized variance we can solve the volatility strike price. Introduced by Broadie and Jain (2008, [27]), they proposed an analytical exact solution using Laplace transforms. They noticed that the square root function can be expressed (from the work of Schürger in 2002, [38]) as:

$$\sqrt{x} = \frac{1}{2\sqrt{\pi}} \int_0^\infty \frac{1 - e^{-sx}}{s^{\frac{3}{2}}} ds. \tag{39}$$

Taking expectations on both sides and using Fubini's theorem we get

$$E[\sqrt{x}] = \frac{1}{2\sqrt{\pi}} \int_0^\infty \frac{1 - E[e^{-sx}]}{s^{\frac{3}{2}}} ds. \tag{40}$$

Changing the x above to the realized variance $V_c(0,T)$ we thus obtain a solution formula for the volatility strike price. Thus we have that the strike is given by

$$K_{vol} = E^Q[\sqrt{V_c(0,T)}] = \frac{1}{2\sqrt{\pi}} \int_0^\infty \frac{1 - E[e^{-sV_c(0,T)}]}{s^{\frac{3}{2}}} ds. \tag{41}$$

Thus we need to use numerical integration techniques in order to solve the above integral which yields the volatility strike price.

4 Jump Diffusion Model

Up to this point we have considered an approach which only utilizes the classic Black-Scholes framework, something that is known to have certain unrealistic assumptions. In this section we will focus on the fact that stock returns do not always behave in the nice continuous way dictated by the GBM, which implicate that they may exhibit jumps. The extension from the B-S model would be to implement a model with the presence of jumps.

To check if a jump diffusion model is even necessary we look at some historical data. Below we show daily closing prices of the S&P500 over the course of 10 years.

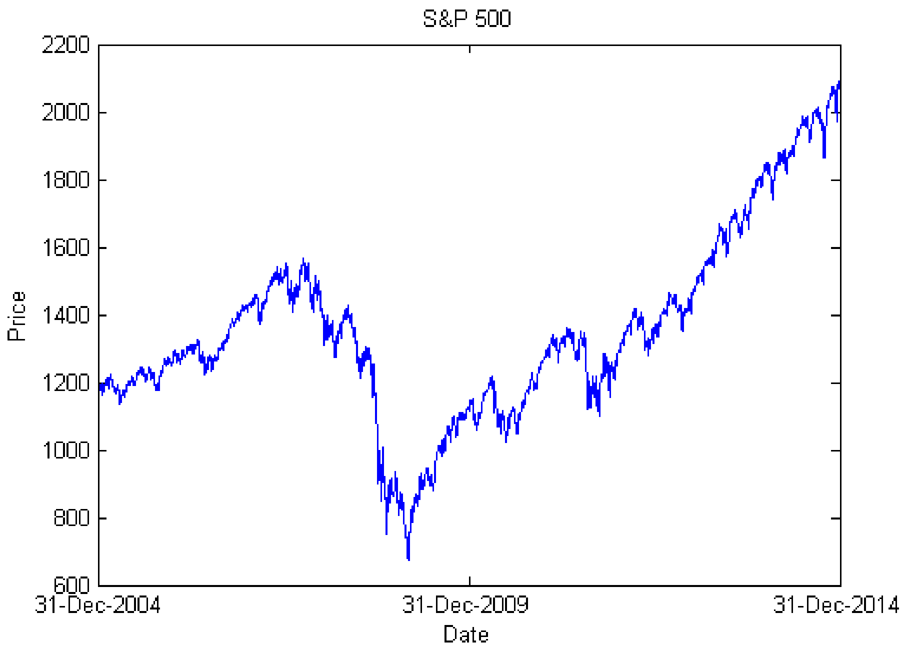

Figure 3: Historical development of the S&P500 from 31/12/2004 to 31/12/2014 [18]

As we can see in the market drastic changes occurs from time to time, something that would encourage a jump model. Most notable of this phenomenon are in crisis periods, where we can see rapid changes in the price. Although not fully convenient, based on this dataset we can say that the prices in S&P500 tend to experience certain jumps. Whether stock prices in general really follows a jump process has to be statistically tested and evaluated, and will not be carried out in this thesis. For such a statistic test we can recommend Stamp and Thorsen (2011, [40]).

Assuming a jump process, there exists a multitude of different models that uses processes with sudden discrete shifts. It was first proposed that stock prices follow a jump process by John Carrington Cox and Stephen Ross in 1976 [12]. They presented a model with a pure jump process such that it works in discrete time. It was later expanded by Robert Merton to a combination of jumps and small continuous movements, and this new process is generally known as the Merton Jump Diffusion model (MJD model) [35].

4.1 Jump Dynamics

In a risk-neutral world the dynamics of a jump diffusion model will usually be given by

$$dS_t = (r - \lambda m)S_{t^-}dt + \sigma S_{t^-}dW_t + S_{t^-}dJ_t. \tag{42}$$

In a jump model we have a few new parameters in contrast to the original setup. J is the added jump process, which will later be defined in equation (43). We have that m is defined as the average jump size, and λ is the intensity of the Poisson process. The subtraction of λm in the risk neutral drift is done to compensate for the drift related to the jump component. As the amount of jumps will be random we need a stochastic counting process and the most common one is a Poisson process, denoted $N(t)$. If we only used this process to model the jumps each jump would have unit length so in order to compensate for that we scale it with a random variable. This new process is called a compounded Poisson process and will be a sum of the random variables defined as

$$J_t = \sum_{i=1}^{N(t)} J_i. \tag{43}$$

Each term in the sum represents a jump at time $0 < t_i < T$, with $-1 < J_i$ as independent identically distributed random variables. Worth noting is that the jump direction and jump size will be based on J, which can be either fixed or random. Although as pointed out in Gatheral (2006, [21]), letting the jump sizes be at a constant value is quite unrealistic. Therefore it is better in practice to assign a distribution to the jump sizes, and a appropriate and popular fit have shown to be a log-normal distribution.

For a more indepth reasoning behind the decision of the distribution we recommend Matsuda (2004, [32]). Thus we will for convenience during this section assume that J follows a lognormal distribution with mean-jump parameter a and standard deviation b such that

$$J \sim LN(a, b^2) \Leftrightarrow \log(J) \sim N(a, b^2). \tag{44}$$

In a Merton jump-diffusion model the floating leg can be shown to be given as follows

$$V_c\left(0,T\right) = \frac{1}{T}\int_0^T \sigma^2 dt + \frac{1}{T}\sum_{i=1}^{N(T)} \log\left(J_i\right)^2. \tag{45}$$

Applying the same arguments as in earlier sections we can again take the expected value of the continuous realized variance, and solve the variance strike price. This yields that assuming a MJD model the continuous strike price is given by

$$K_{var} = E[V_c\left(0,T\right)] = \sigma^2 + \lambda(a^2 + b^2). \tag{46}$$

One of the great benefits for using the MJD model is that we allow for sudden large movements in the spot price, something that has nice implications for option pricing. The extension from B-S to MJD allows us to add features which tries to capture negative skewness and excess kurtosis of the log stock price, an area the standard B-S fails to account for.

4.2 The Effect of Jumps

We devote this section to compare the previous approach we had where we priced a variance swap with as few assumptions as possible, and applying a replication scheme. We saw that within the Black-Scholes framework we were able to receive a closed solution formula for the variance strike price using a replicating approach. Here we extend the Black-Scholes model, and proceed to look into the same pricing scheme when the dynamics are allowed to exhibit jumps. Thus we will derive the same types of contract as in the previous chapter, with the added jump feature.

As we are using Itô's lemma in our derivation without jumps, we will in order to compare it effectively with our jump model naturally use Itô's lemma again for a compounded Poisson process. The proof for the calculations is left for the reader in Appendix A. It follows from the lemma that for any sufficiently smooth function f we have

$$f\left(J_T\right) - f\left(J_0\right) = \sum_{i\geq 1, t_i \leq T}\left[f\left(J_{t_i^-} + J_i S_{t_i^-}\right) - f\left(J_{t_i^-}\right)\right]. \tag{47}$$

It is important to note that the Poisson jump part of a process and the normal diffusion part under Itô's lemma are independent, such that Itô's lemma for a process which is the sum of a drift-diffusion process and a jump process is just the sum of the Itô's lemma for the individual parts.

Returning to our specific question we derive the strike price the same way as we did earlier, with the added change that the stock now follows the equation

$$S_t = S_0 + \int_0^t \mu S_s ds + \int_0^t \sigma S_s dW + \sum_{i=1}^{N(t)} J_i S_{t_i^-}. \tag{48}$$

Our previous approach included both a evaluating a log contract and a stock position. Thus our next step will be to examine the properties of these two identities. Using Itô's lemma on a log contract with the new stock dynamics we arrive at

$$\begin{aligned} &\log\left(\frac{S_T}{S_0}\right) = \\ &\int_0^T \left(\mu - \frac{\sigma^2}{2}\right) dt + \int_0^T \sigma dW_t + \sum_{i=1}^{N(T)} \left(\log(S_{t_i^-} + J_i S_{t_i^-}) - \log(S_{t_i^-})\right). \end{aligned} \tag{49}$$

Similarly for the stock position we get

$$\int_0^T \frac{dS_t}{S_t} = \int_0^T \mu dt + \int_0^T \sigma dW_t + \sum_{i=1}^{N(T)} \left(\frac{S_{t_i^-} + J_i S_{t_i^-}}{S_{t_i^-}} - \frac{S_{t_i^-}}{S_{t_i^-}}\right). \tag{50}$$

Combining these equations to arrive at the floating variance leg now adds a sum as well as the original functions

$$\int_0^T \frac{\sigma^2}{2} dt = \int_0^T \frac{dS_t}{S_t} - \log\frac{S_T}{S_0} - \sum_{i=1}^{N(T)} \left(J_i - \log\left(1 + J_i\right)\right). \tag{51}$$

This gives an idea about the error received when a non-jump model is falsely assumed or vice versa. To get a better overview of the effect from each jump we consider expanding a log contract as a Taylor expansion such that the initial term cancel. This gives an error term for each jump with the following characteristics

$$\left(J_i - \log\left(1 + J_i\right)\right) = J_i - J_i + \frac{1}{2}\left(J_i\right)^2 + ... \tag{52}$$

This leaves us with a dominating term for the error received from each jump as the square of the jump. This shows that the error on the strike price if a jump free model was assumed. The formula for the variance strike when a jump model is assumed is thus given by the equation

$$
\begin{aligned}
K_{var}^{jump} =& \frac{2}{T} E\left[\int_0^T \frac{dS_t}{S_{t^-}} - \log\left(\frac{S_T}{S_0}\right)\right] - \frac{2}{T} E\left[\sum_{i=1}^{N(T)} \left(J_i - \log\left(1 + J_i\right)\right)\right] \\
=& \frac{2}{T} E\left[\int_0^T \frac{dS_t}{S_{t^-}} - \log\left(\frac{S_T}{S_0}\right)\right] - 2\lambda E\left[J_i - \log\left(1 + J_i\right)\right].
\end{aligned} \tag{53}
$$

This formula is very similar to the formula for the strike price in the case when no jumps were assumed although it is important to remember that the characteristics of the underlying asset price is different. Although if correctly priced European call and put options exist in the market then the calculations for hedging the swap with an infinite number of calls and puts are done as in the model without jumps.

Exploring the possibility with a market with infinitely many correctly priced call and put options we can approximate the error received by comparing a jump model towards its counterpart with no jumps. As the formula for the strikes are only slightly different we notice that the extra term is

$$2\lambda E\left[J_i - \log\left(1 + J_i\right)\right] = \lambda E\left[J_i^2 + ...\right]. \tag{54}$$

This shows that the approximated error is close to the jump intensity times the expected square of the jump size, which gives a linear increase in error if the intensity changes but a quadratic increase if the jump size increase.

Thus we can see that proceeding with the replication approach when falsely assuming jumps or vice versa can cause quite ineffective results. Demeterfi et al. (1999, [13]) also points out that neglecting jumps in a synthetic replication scheme can give substantial errors.

5 Stochastic Volatility Models

To measure the dynamics of asset prices or other financial aspects one often uses stochastic processes to properly explain the random movements. Using stochastic processes is a big part of financial mathematics, and has been known to play a crucial role in modelling option pricing.

When speaking about random processes and modelling of random movements, one of the most famous processes used in the Black-Scholes model is the geometric Brownian motion (GBM). Empirically it has been shown however that the GBM has failed to suffice as a valuable model for pricing and hedging different securities. Looking for better empirical motivation, one idea has been to look towards Lévy processes (for an example see Carr, Wu (2004, [9])).

We are however in this chapter going to focus on volatility, and attaching a random process on this component. Remember that in the B-S model we assume constant volatility, which is something that is generally not true in the market. When it comes to volatility there are quite a few things that empirically speaks against the Black-Scholes framework. An observation by Mandelbrot (1963, [31]) said that "large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes". This effect is usually known as volatility clustering, and doesn't exist in the B-S model world. Another fallout is the "leverage effect", which says that decreases in the underlying asset price often results in increases in volatility. In practice it can also be shown that volatility takes different values depending on moneyness (how much out-of-the-money or in-of-the-money the option is at the current time) as well as term structure. This phenomenon is usually called "volatility smile", and is a widely debated and discussed area in option pricing. To relax the B-S assumption about constant volatility in search for better empirical support, stochastic volatility models becomes a natural choice.

After the remarkable introduction of the B-S model in the 70s, there have been as pointed out above many debates and arguments about the accuracy and validity of the solution formulas and the assumptions made in the model. In search for models to fit the market better, many researches have turned to stochastic volatility models. These types of models, which now allows variance to vary in time and be driven by its own process, has been a ground-breaking result in the area of option pricing. Stochastic volatility models can be seen as an extension of Black-Scholes, and tries to capture some of the flaws in the original model. In this thesis we will focus on the kind of SV models which takes the particular form:

$$dS_t = rS_t dt + \sqrt{v_t} S_t \left(\rho dW_t^1 + \sqrt{1-\rho^2} dW_t^2 \right) \tag{55}$$

$$dv_t = \kappa \left(\theta - v_t \right) dt + \sigma_v v_t^\gamma dW_t^1. \tag{56}$$

The first equation (55) gives us the dynamics of the underlying asset: where S_t denotes the asset price at time t, r is the riskless interest rate and $\sqrt{v_t}$ is the volatility at time t. The two W's, W_t^1 and W_t^2, are two standard Brownian motions (under the risk neutral measure Q) which allows the diffusion processes to be dependent. The correlation between the two is characterized by the factor ρ. The dynamics of the first equation above will be the same for the different types of models we will be looking at in this chapter. Worth mentioning is that for simplicity we assume that the asset pays no dividends and that the interest rate is constant.

The second equation (56) specifies the behaviour of the variance v_t as a mean-reverting process. The structure for the variance dynamics can be altered in several ways, for example different types of γ (a free parameter $\gamma > 0$) provides different models. The ones we will be looking at in this paper will be $\gamma = 1$ which corresponds to the GARCH model and $\gamma = 1/2$ the famous Heston model. A slight alteration in the dynamics of the variance along with $\gamma = 3/2$ gives the 3/2 model.

As for the other parameters in the second equation we define θ as the long-run mean variance, κ as the rate of speed in which v_t tends to θ, and σ_v as the volatility of the volatility, a parameter which will determine the volatility of the variance process. The type of equation that describes the dynamics of the variance (56) is generally known as a time-dependent CIR-process (Cox-Ingersroll-Ross process).

5.1 The Heston Model

As mentioned our previous approach in pricing volatility derivatives have been to work in a semi model-independent world where no great emphasis or assumptions have been put on the underlying model. In this section we look at the Heston model, a stochastic volatility model which assumes that the variance of the asset follows a random process. This model is one of the most commonly used when assuming stochastic volatility and was introduced by Steven Heston in 1993. [23]

Working with the equations we presented in the beginning of this chapter, we now have that $\gamma = 1/2$. The Heston model is thus under the risk neutral measure determined by the stochastic process:

$$dS_t = rS_t dt + \sqrt{v_t} S_t \left(\rho dW_t^1 + \sqrt{1-\rho^2} dW_t^2 \right) \tag{57}$$

$$dv_t = \kappa\left(\theta - v_t\right)dt + \sigma_v\sqrt{v_t}dW_t^1. \tag{58}$$

The parameters are defined by equation (55) and (56), as well as $\gamma = 1/2$. To ensure that the variance process stays positive we restrict that the parameters obey the inequality that $2\kappa\theta > \sigma$, a condition known as the Feller condition (1951, [19]). As usual we denote the index level by S_0 as well as the initial variance process by v_0. To obtain the variance process v_t, an unobservable process, one has to estimate it from some sort of data (for example a time series of S_t).

To price volatility derivatives in this model we already have several parameters to our disposal, which makes us able to calculate the conditional expectation of the variance process since it is an ordinary integral equation. The calculations are shown below:

$$\begin{aligned} E[v_t|v_0] &= E\left[v_0 + \int_0^t \kappa(\theta - v_s)ds + \int_0^t \sigma_v\sqrt{v_s}dW_s^1 \middle| v_0\right] \\ &= E\left[v_0 + \int_0^t \kappa(\theta - v_s)ds \middle| v_0\right] \\ &= v_0 + \int_0^t \kappa(\theta - E[v_s|v_0])ds. \end{aligned} \tag{59}$$

This integral equation is equivalent to the ordinary differential equation given with corresponding initial condition

$$\begin{aligned} \frac{dE[v_t|v_0]}{dt} &= \kappa(\theta - E[v_t|v_0]) \\ E[v_0|v_0] &= v_0. \end{aligned} \tag{60}$$

Solving this ODE is a simple task and will, by standard calculations, yield a solution given by

$$E[v_t|v_0] = \theta + (v_0 - \theta)e^{-\kappa t}. \tag{61}$$

When we delve further into the stochastic volatility models this relation will be of significant importance and have great implications in pricing volatility derivatives and especially the variance swap.

5.2 The GARCH Model

In a GARCH (Generalized Autoregressive Conditional Heteroskedasticity) model we assume that the stochastic part of the variance process varies with the variance, compared to the square root of the variance above in the Heston case. Thus for a standard GARCH(1,1) model we define the following dynamics:

$$dS_t = rS_tdt + \sqrt{v_t}S_t\left(\rho dW_t^1 + \sqrt{1-\rho^2}dW_t^2\right) \tag{62}$$

$$dv_t = \kappa\left(\theta - v_t\right)dt + \sigma_v v_t dW_t^1. \tag{63}$$

The various parameters here are defined the same as above, where now $\gamma = 1$. Comparing the GARCH model to the Heston model we notice that the only mathematical difference lies in the dynamics of the variance, but we also have in this model that the variance and spot price are uncorrelated. This means that for this model we have that $\rho = 0$. The GARCH(1,1) model, formulated by Engle and Mezrich (1995, [16]), were given in a discrete setting which had the form

$$\log\left(\frac{S_n}{S_{n-1}}\right) = r_n = \mu + u_n, \; u_n = v_n\varepsilon_n, \; \varepsilon_n \sim N(0,1) \tag{64}$$

$$v_{n+1} = (1-\alpha-\beta)V + \alpha u_n^2 + \beta v_n \tag{65}$$

Here V is the long-run variance, u_n is the drift-adjusted stock return at time n, α the weight assigned to the squared drift-adjusted return (u_n^2) and β the weight assigned to the last period variance (v_n). This discrete model can be transformed into the continuous time model, where we have the following parameter relationships:

$$\theta = \frac{V}{dt} \tag{66}$$

$$\kappa = \frac{1-\alpha-\beta}{dt} \tag{67}$$

$$\sigma_v = \alpha\sqrt{\frac{\xi-1}{dt}} \tag{68}$$

In the above equation ξ stands for the Pearson kurtosis (fourth moment), calculated from historical data on u_n. As we stated earlier the GARCH(1,1) model implies an uncorrelation between the spot price and the volatility, something that is certainly not always the case for all models. To approach the matter one could adopt a Nonlinear GARCH model (NGARCH), which introduces an extra term defined as

$$v_{n+1} = (1-\alpha-\beta)V + \alpha(u_n - c)^2 + \beta v_n. \tag{69}$$

Here we have added an extra parameter, c, which also needs to be estimated and will facilitate the correlation with the underlying process. We will however not put our focus on this relation here, and the standard GARCH(1,1) will accurately suffice for now. More about the NGARCH process can be found in the original paper by Engle and Ng (1993, [17]).

A considerable difference between the Heston and the GARCH model is that in this paper we will now utilize historical data in order to estimate future levels of volatility. This can be seen in contrast to the Heston model in which we use option data and market implied estimates of future volatility as calibration tools.

Solving the expected value of the variance we use the same approach as we had in the Heston model, and after some quick calculations we see that we achieve the same result as in the Heston case. With the parameters at our disposal, and considering that the stochastic component vanishes as seen in equation (59), all calculations have already been done. Thus the solution of the conditional expected value in a GARCH(1,1) model is given by

$$E[v_t|v_0] = \theta + (v_0 - \theta)e^{-\kappa t}. \tag{70}$$

Thus we see that we achieve the same conditional expectation for both those model, and the only difference lies in the estimation method of the parameters.

5.3 The 3/2 Model

Working in the 3/2 model we can see close similarities with the Heston and GARCH model, but the dynamics of the variance process alter slightly. One of the major questions when working with a stochastic volatility model is the choice of γ, and which value of the parameter would fit reality the best. Bakshi et al. (2006, [2]) estimated in their study a value of approximately 1.28 for γ, which suggests that the 3/2 model would be better suited than the Heston model which has $\gamma = 0.5$.

Following the same setup as earlier, we define the model by the following dynamics:

$$dS_t = rS_t dt + \sqrt{v_t} S_t \left(\rho dW_t^1 + \sqrt{1-\rho^2} dW_t^2\right) \tag{71}$$

$$dv_t = \kappa v_t (\theta - v_t) dt + \sigma_v v_t^{\frac{3}{2}} dW_t^1. \tag{72}$$

Again the parameters are defined the same as earlier together with $\gamma = 3/2$. The difference here can be seen as having the same dynamics as in the Heston model multiplied by an extra factor of v_t, i.e.

$$dv_t = v_t[\kappa (\theta - v_t) dt + \sigma_v \sqrt{v_t} dW_t^1]. \text{ [15]} \tag{73}$$

One thing to notice in this model is the interpretation and motivation of some of the parameters. The speed of mean reversion (previously κ) will now be given by the product $\kappa \times v_t$ which is now a stochastic quantity. This will imply that the variance will revert more quickly to the mean when it is high. This product then makes the model more suitable to handle fast volatility increases or decreases, a point at which Heston fails. σ_v is again the volatility of volatility, and will for higher values result in heavier tails on the sides. [29]

Since the dynamics of the 3/2 models alters in the dynamics of the variance compared to Heston and GARCH, the calculation of the conditional expectation becomes slightly different and the quadratic term complicates the process. We will however apply the relationship adopted by Drimus (2009, [14]) which gives a link between the 3/2 model and the Heston model. This approach says that the dynamics of $1/v_t$ follows a Heston process with some specific parameters. Applying Itô's lemma to $1/v_t$ when v_t follows the dynamics of the 3/2 model as in (72) we get

$$d\left(\frac{1}{v_t}\right) = \kappa\theta\left(\frac{\kappa+\sigma_v^2}{\kappa\theta} - \frac{1}{v_t}\right)dt - \frac{\sigma_v}{\sqrt{v_t}}dW_t. \tag{74}$$

Thus we notice that the reciprocal of the 3/2 variance process is a Heston process of parameters $\left(\kappa\theta, \frac{\kappa+\sigma_v^2}{\kappa\theta}, -\sigma_v\right)$. Having this relationship to our disposal we can then calculate the conditional expected value for the 3/2 model. For simplicity lets denote

$$\tilde{\kappa} = \kappa\theta \tag{75}$$

$$\tilde{\theta} = \frac{\kappa+\sigma_v^2}{\kappa\theta} \tag{76}$$

$$\tilde{\sigma_v} = -\sigma_v. \tag{77}$$

Then our solution for equation (74) is calculated the same way as in (61) with a slight change to the initial condition. Given this solution for $E[\frac{1}{v_t}|v_0]$ we conclude by inverting back to achieve the final solution which ends up being

$$E\left[v_t | v_0\right] = \frac{1}{\tilde{\theta} + \left(\frac{1}{v_0} - \tilde{\theta}\right)e^{-\tilde{\kappa}t}}. \tag{78}$$

5.4 Variance Swaps

To derive the price for a variance swap across the different stochastic volatility models we will use the same initial thoughts as in the variance swap section. In a stochastic volatility model (SV model), we start by defining the continuous realized variance by

$$V_c\left(0,T\right) = \frac{1}{T}\int_0^T v_s ds. \tag{79}$$

The fair variance strike price, again denoted by K_{var}, defines the value in which the contract's net present value is 0. In terms of equations we have, under the risk neutral measure Q:

$$E_0^Q\left[e^{-rT}(V_c(0,T) - K_{var})\right] = 0. \tag{80}$$

With this we are now able to calculate the continuous realized variance given the dynamics we have set up, because we can explicitly calculate the expected value for the different models. Thus given a SV model, we can solve this above equation for the variance strike price. In the presence of a GARCH or Heston model, we get the following price:

$$\begin{aligned} K_{var} &= E[V_c(0,T)|v_0] = E\left[\frac{1}{T}\int_0^T v_t dt \middle| v_0\right] \\ &= \frac{1}{T}\int_0^T E[v_t|v_0]dt = \frac{1}{T}\int_0^T \left[\theta + (v_0-\theta)e^{-\kappa t}\right]dt \\ &= \frac{1}{T}\Big(\theta T + \frac{v_0-\theta}{\kappa}(1-e^{-\kappa T})\Big). \end{aligned} \tag{81}$$

Here we notice that the fourth equality comes from the calculation of the conditional expectation, which had been done in equation (59). Thus working within the stochastic volatility models we are able to find an explicit formula for the continuous variance strike price.

In the frame of the 3/2 model, since the conditional expectation of the variance looks a bit different, we get the following price:

$$\begin{aligned} K_{var} &= E[V_c(0,T)|v_0] = E\left[\frac{1}{T}\int_0^T v_t dt \middle| v_0\right] \\ &= \frac{1}{T}\int_0^T E[v_t|v_0]dt \\ &= \frac{1}{T}\int_0^T \frac{1}{\tilde{\theta} + \left(\frac{1}{v_0}-\tilde{\theta}\right)e^{-\tilde{\kappa}t}}dt \\ &= \frac{1}{\tilde{\kappa}\tilde{\theta}T}\ln\left(v_0\tilde{\theta} + \left(1-v_0\tilde{\theta}\right)e^{-\tilde{\kappa}T}\right) + \frac{1}{\tilde{\theta}}. \end{aligned} \tag{82}$$

Thus given that we can estimate the parameters we are able to obtain the price of a variance swap with a certain time to maturity T. These calibrations and pricing procedures will be done in chapter 6 and 7.

5.5 Volatility Swaps

For the volatility strike however there is no easy way to obtain the strike price explicitly, and the only approximations given has been the upper bound that we saw in the volatility swap section. As we noted earlier, we have to our disposal the convexity correction formula in equation (38). However as pointed out by Broadie and Jain (2008, [27]) the correction formula might not be as accurate as one would like in the case of SV models, because higher order terms in the Taylor expansion are not negligible. Thus we have fairly limited information of how to properly transition into the volatility swaps. One thing to solve that would be to allow more terms in the Taylor expansion for a SV model, but that is a fairly long stretch.

Thus in order to compute the fair volatility strike we will instead in this section derive a partial differential equation which numerically can be solved to approximate the volatility strike price. The approach is built on the fact of a no-arbitrage relationship between volatility and variance swaps. We will also derive a analytical solution for the volatility swap which includes the use of Laplace transforms and numerical integration, an approach set up by Broadie and Jain (2008, [27]).

5.5.1 PDE Approach

Variance swaps are in their own right very liquid instruments and can be used as a valid trading tool. They can however also be regarded as underlying assets in order to price other types of volatility instruments including volatility swaps and variance options. In this section we derive a PDE to price volatility derivatives, and in particular give an approximated price of the volatility strike price.

Initially we define the price process of the floating leg, denoted by

$$X_t^T = E_t^Q\left[\frac{1}{T}\int_0^T v_s ds\right]. \tag{83}$$

This process is sometimes used as an argument in other price processes that have a close relation to the floating leg of a variance swap. Since the price process of the floating leg is not completely stochastic because it has, at time t, a deterministic part we will define a state variable $I_t = \int_0^t v_s ds$ to be able to separate the stochastic and the deterministic part of the process. This variable has a SDE representation of $dI_t = v_t dt$. Thus the price process of the floating leg has the representation

$$X_t^T = \frac{I_t}{T} + E_t^Q\left[\frac{1}{T}\int_t^T v_s ds\right]. \tag{84}$$

Then we have for any derivative depending on the floating leg, and in particular the variance swap, a price process $G\left(t, X_t^T\right) = F\left(t, v_t, I_t\right)$. Applying Itô's lemma on F we get the PDE

$$dF = \frac{\partial F}{\partial t}dt + \frac{\partial F}{\partial v}dv + \frac{\partial F}{\partial I}dI + \frac{1}{2}\frac{\partial^2 F}{\partial v^2}dv^2. \tag{85}$$

Using the above equation (85) we can derive the various PDEs for our different SV models.

Working with our price process F we thus need to substitute the dynamics of each model into equation (85). Starting out in the Heston model, the result can be simplified by collecting terms, arriving at

$$dF = \left[\frac{\partial F}{\partial t} + \frac{\partial F}{\partial v} \kappa \left(\theta - v_t\right) + \frac{\partial F}{\partial I} v_t + \frac{1}{2} \frac{\partial^2 F}{\partial v^2} v_t \sigma^2 \right] dt + \frac{\partial F}{\partial v} \sigma \sqrt{v_t} dW_t^1. \quad (86)$$

Since F is a forward price process, its drift under the risk neutral measure must be zero. Hence solving the PDE for a derivative on the floating leg we can get an approximation to the price of F:

$$\frac{\partial F}{\partial t} + \frac{\partial F}{\partial v} \kappa \left(\theta - v_t\right) + \frac{\partial F}{\partial I} v_t + \frac{1}{2} \frac{\partial^2 F}{\partial v^2} v_t \sigma^2 = 0. \quad (87)$$

This above PDE is then solved to compute the fair volatility strike. The boundary condition needed in order to solve this is

$$F(T, V_T, I_T) = \sqrt{\frac{I_T}{T}} \quad (88)$$

Thus the framework in which we obtained the PDE in equation (87) can also be implemented for the other SV models. The results for a GARCH model yields the following PDE as well as boundary condition:

$$\frac{\partial F}{\partial t} + \frac{\partial F}{\partial v} \kappa \left(\theta - v_t\right) + \frac{\partial F}{\partial I} v_t + \frac{1}{2} \frac{\partial^2 F}{\partial v^2} v_t^2 \sigma^2 = 0. \quad (89)$$

$$F(T, V_T, I_T) = \sqrt{\frac{I_T}{T}} \quad (90)$$

Substituting the variance dynamics of the 3/2 model into the price process we get a slightly different PDE, but the mechanics will be done the same as above. We again have that since F is a price process, the drift under the risk neutral measure has to be zero. Thus we arrive at the PDE

$$\frac{\partial F}{\partial t} + \frac{\partial F}{\partial v} \kappa v_t \left(\theta - v_t\right) + \frac{\partial F}{\partial I} v_t + \frac{1}{2} \frac{\partial^2 F}{\partial v^2} v_t^3 \sigma^2 = 0. \quad (91)$$

$$F(T, V_T, I_T) = \sqrt{\frac{I_T}{T}} \quad (92)$$

5.5.2 Laplace Transforms

Using the same argument as we did earlier in this, we can numerically solve the volatility strike price with the use of Laplace Transforms. The continuous volatility strike can again be evaluated by the formula defined as

$$K_{vol} = \frac{1}{2\sqrt{\pi}} \int_0^\infty \frac{1 - E[e^{-sV_c(0,T)}]}{s^{\frac{3}{2}}} ds. \quad (93)$$

Thus in this setting we have to evaluate the expected value in the case of a Stochastic Volatility model. As an example we look at the Heston model which gives, from Lian (2010, [30]) the following formula:

$$E^Q[e^{-sV_c(0,T)}] = \exp[A(T,s) - B(T,s)v_0] \tag{94}$$

where

$$A(T,s) = \frac{2\kappa\theta}{\sigma_v^2}\log\left(\frac{2\gamma(s)e^{\frac{(\gamma(s)+\kappa)T}{2}}}{(\gamma(s)+\kappa)(e^{\gamma(s)T}-1)+2\gamma(s)}\right) \tag{95}$$

$$B(T,s) = \frac{2s(e^{\gamma(s)T}-1}{T[(\gamma(s)+\kappa)(e^{\gamma(s)T}-1)+2\gamma(s)]} \tag{96}$$

$$\gamma(s) = \sqrt{\kappa^2 + 2\frac{\sigma_v^2 s}{T}}. \tag{97}$$

After estimating the parameters we again rely on numerical integration techniques in order to solve the integration part in order to obtain the volatility strike price.

5.6 Comparison of the Models

As we can see there are some small variations and differences between the models. Starting off with the variance swap we see that the Heston and GARCH model gives the same explicit formula, where as the 3/2 model becomes slightly different. For the volatility swaps we received three different PDEs which can be solved in order to obtain a volatility strike price. Looking more into the PDEs we can notice that the variance terms in the second derivative of the variance alters. For this term we have v_t in the Heston model, v_t^2 in the GARCH model and v_t^3 for the 3/2 model. We can also see that we get an extra v_t term in the first derivative of the variance for the 3/2 model, due to its extra term in the variance dynamics.

So which model is the optimal one to use? Well we can't say that yet, and need to estimate the parameters in the models in order to evaluate our results. This question also depends on what we are looking for particularly, whether it is accuracy or computational speed, or how the model calculates data. All these attributes are all valuable when looking over large datasets as too complex models may induce heavy computational power.

The most popular out of the three models that we have presented in this section is undoubtedly the Heston model. It is usually the first model that comes to mind when speaking about stochastic volatility models. After several attempts in extending the Black-Scholes framework, Heston was one of the first in creating a model which had the analytical tractability to give semi-closed form expressions. While GARCH mainly have had its success in econometrics it has over the years found its way into the area of volatility derivatives, being in close relation to the famous Heston model. It is also often estimated using a price-based approach, compared to the others which are option-based, which will yield different set of parameters when doing the estimations. By Drimus (2009, [14]) the 3/2 model have been shown to give a better empirical support in certain areas, for example supporting upward sloping implied volatility of variance smiles.

We will later on investigate how well these models perform empirically with estimated parameters. This will be done in chapter 7.

5.7 Stochastic Volatility Models with Jumps

The developments for stochastic volatility models have been revolutionary in the area of finance in the sense that it gives a more realistic view of how stock prices behave, thus giving a more comprehensive option valuation. As it stands with most models, there is almost always room for improvements. Generally SV models are good at capturing the volatility smile on longer maturities, but fail to fully support the skew at short maturities as pointed out by Gatheral (2006, [21]). Thus we extend to allow for jumps in the stochastic volatility model, which basically means adding a jump component to our already existing model.

A model which allows jumps can be divided into two sections, the simplest case being the Stochastic Volatility with spot Jump model (SVJ model) with jumps in the price process. An extension to the SVJ model is also considering jumps in the variance process, which we denote as a Stochastic Volatility with spot Jump and volatility Jump model (SVJJ). We will thus in this section be looking at both the SVJ and SVJJ model, and give an introduction into both of these extensions.

5.7.1 Stochastic Volatility with Jumps in the Underlying (SVJ)

The implications for introducing jumps to a SV model will basically look the same, and we will thus in this section focus our attention on the most famous model, the Heston model. If we first allow for jumps in the underlying stock process (a so called SVJ model) the risk neutral dynamics then becomes:

$$dS_t = (r - \lambda m)S_{t^-}dt + \sqrt{v_t}S_{t^-}\left(\rho dW_t^1 + \sqrt{1-\rho^2}dW_t^2\right) + S_{t^-}dJ_t \tag{98}$$

$$dv_t = \kappa\left(\theta - v_t\right)dt + \sigma_v\sqrt{v_t}dW_t^1. \tag{99}$$

The parameters given above are the same as in the Heston model specified in (57) and (58), as well as the applied Merton Jump-Diffusion (MJD) model which is defined in (42). We assume here that the jump process N_t and the

Brownian motion are independent.

Calculating the variance strike price in this jump model we apply the same procedure as in the regular Heston model, where now we have to into consideration the added jump process. Using these results we can then calculate the price of a variance swap in a SVJ model as

$$K_{var} = \frac{1}{T}\Big(\theta T + \frac{v_0 - \theta}{\kappa}(1 - e^{-\kappa T})\Big) + \lambda(a^2 + b^2) \text{ [27]}. \tag{100}$$

As we can see the result is fairly straight forward compared to earlier, and does not add much complexity. The extra jump process is as we can see only added on top of what we already calculated in a Heston model without jumps. Thus as long as we know the dynamics for the jumps and the parameters, i.e. λ, a and b, we can receive the price for a variance swap without much trouble the same way we did in the original Heston model. The downside to this model is that the jump process is uncorrelated with the variance process, implicating that that volatility stays the same after a jump occurs in the underlying.

5.7.2 Stochastic Volatility with Jumps in Stock Price and Volatility (SVJJ)

Although adding a jump process to the stochastic volatility models was considered a major improvement, there will always be the quest to finding the perfect model. Up to this point we have considered a model which only contains jumps in the price process, but not allowing for jumps in the variance process. If we think about this assumption it is quite unrealistic that the instantaneous volatility would not jump if the stock price jumps. Thus it would be appropriate to apply a SVJJ model, such that we have jumps in both processes. This goes in line with the quote we had earlier about volatility clustering, that "large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes". This model, defined in 1999 by Andrew Matytsin [33], will have the dynamics given by:

$$dS_t = (r - \lambda m)S_{t^-}dt + \sqrt{v_t}S_{t^-}\left(\rho dW_t^1 + \sqrt{1-\rho^2}dW_t^2\right) + S_{t^-}dJ_t^s \tag{101}$$

$$dv_t = \kappa\left(\theta - v_t\right)dt + \sigma_v\sqrt{v_t}dW_t^1 + dJ_t^v. \tag{102}$$

Added here is a jump process in the dynamics of the variance, namely J_t^v. We assume that the jump size in this process are exponential with mean η_J and the correlation between the two jump processes are defined as follows:

$$J^v \sim \exp\left(\frac{1}{\eta_J}\right),\ J^s|J^v \sim LN(a + \rho_J J^v, b^2). \tag{103}$$

Concluding this section the question to ask now becomes which model is preferable, the original Heston, the single jump model (SVJ) or the double jump model (SVJJ). One might assume that including more things and making a more complex model would fit better, but that is arguably not always the case. A comparison as a whole for the benefits of adding jumps into the model is a trade off. Even though we might mitigate some problems that the original model cannot account for (such as short maturity skews), we have on the other hand increased complexity and in some cases loss of tractability.

Mathematically it becomes more complex when adding jumps, which of course affects the computational speed. As pointed out in Gatheral (2006, [21]), the SVJJ model has more parameters but were also harder to fit observed option prices. Working with more parameters might be good in some cases, but the main issue with SVJJ is that the added performance is overmatched by the computational technicality and power needed. The SVJ model has more parameters than the original Heston model, but as we pointed out the added parameters are technically not that difficult to implement, thus giving a better fit to observed data compared to the original model.

5.8 Variance Options

Another interesting derivative in this area is options which are not based on an underlying asset, but instead written on variance or volatility. We will in section derive an approach which introduces variance options. In order to price variance options, we use a no-arbitrage argument which is similar to the PDE approach we saw in the case of volatility swaps.

Using this strategy we create a specific portfolio which lets us find a partial differential equation which will be similar to the ones we saw for the volatility swap. As usual we denote the price of a standard European call option as:

$$C_t = E_t^Q[e^{-r(T-t)}\max(V_c(0,T) - K, 0)] \times N. \tag{104}$$

We will derive a partial differential equation to price a variance call option, and the payoff of the call can be replicated by continuous trading of a variance swap. This replicating approach will yield the price of the variance call option. Thus we need to form a portfolio including both a call option and α units of variance swaps. The created portfolio will at time 0 have a value of

$$\Pi_0 = \alpha(X_0^T - K_{var}) + C_0. \tag{105}$$

Worth noting is that the portfolio value will be the same as the variance call option value because there is no cost to a unit of variance swap at the start of the contract. The price process for the variance call, denoted C_t^T, can be expressed as

$$C_t^T = G(t, v_t, I_t) \tag{106}$$

Applying Itô's lemma to our price process we get

$$dG = \frac{\partial G}{\partial t}dt + \frac{\partial G}{\partial v}dv + \frac{\partial G}{\partial I}dI + \frac{1}{2}\frac{\partial^2 G}{\partial v^2}dv^2. \tag{107}$$

This equation can be simplified using the parameters from our SV models, and we will perform the calculations in the case of a Heston model. The procedures for creating the PDE for variance options are technically done the same for our different models, and will thus only be carried out in the Heston case. We will only apply this approach in terms of either a GARCH, Heston or 3/2 model, and will in this paper not consider the extensions into the jump frameworks (SVJ and SVJJ). Putting the variance dynamics from the Heston model into G we obtain

$$dG = \left[\frac{\partial G}{\partial t} + \frac{\partial G}{\partial v}\kappa\left(\theta - v_t\right) + \frac{\partial G}{\partial I}v_t + \frac{1}{2}\frac{\partial^2 G}{\partial v^2}v_t\sigma^2\right]dt + \frac{\partial G}{\partial v}\sigma\sqrt{v_t}dW_t^1. \tag{108}$$

From equation (105) the change in portfolio value for a small time step dt is

$$d\Pi_t = \alpha dF + dG. \tag{109}$$

Now we notice that we have done both dF (in the volatility swap section) and dG (in this section), which we can use in order to simplify. Thus we substitute dF from equation (86) and dG from equation (108) into equation (109) which yields:

$$\begin{aligned} d\Pi_t = {} & \alpha\Big(\frac{\partial F}{\partial v}\sigma_v\sqrt{v_t}dW_t^1\Big) \\ & + \Big[\frac{\partial G}{\partial t} + \frac{\partial G}{\partial v}\kappa(\theta - v_t) + \frac{\partial G}{\partial I}v_t + \frac{1}{2}\frac{\partial^2 G}{\partial v^2}v_t\sigma_v^2\Big]dt \\ & + \frac{\partial G}{\partial v}\sigma_v\sqrt{v_t}dW_t^1. \end{aligned} \tag{110}$$

We can now recognize that a suitable choice of α can remove the stochastic component. Thus the random segment vanishes if we choose

$$\alpha = -\frac{\partial G}{\partial v}\Big/\frac{\partial F}{\partial v}. \tag{111}$$

We can then simplify (110) to

$$d\Pi_t = \Big[\frac{\partial G}{\partial t} + \frac{\partial G}{\partial v}\kappa(\theta - v_t) + \frac{\partial G}{\partial I}v_t + \frac{1}{2}\frac{\partial^2 G}{\partial v^2}v_t\sigma_v^2\Big]dt. \tag{112}$$

Noting that the portfolio Π_t is riskless, we want it to equal the risk free rate of return (to avoid arbitrage opportunities), which yields

$$\Big[\frac{\partial G}{\partial t} + \frac{\partial G}{\partial v}\kappa(\theta - v_t) + \frac{\partial G}{\partial I}v_t + \frac{1}{2}\frac{\partial^2 G}{\partial v^2}v_t\sigma_v^2\Big]dt = rGdt. \tag{113}$$

This can thus be rewritten as

$$\frac{\partial G}{\partial t} + \frac{\partial G}{\partial v}\kappa(\theta - v_t) + \frac{\partial G}{\partial I}v_t + \frac{1}{2}\frac{\partial^2 G}{\partial v^2}v_t\sigma_v^2 - rG = 0. \tag{114}$$

For the cases of a GARCH or 3/2 model we get slightly different PDEs, which will look similarly as in the section of volatility swaps. The corresponding PDEs will be

$$\frac{\partial G}{\partial t} + \frac{\partial G}{\partial v}\kappa(\theta - v_t) + \frac{\partial G}{\partial I}v_t + \frac{1}{2}\frac{\partial^2 G}{\partial v^2}v_t^2\sigma_v^2 - rG = 0. \tag{115}$$

$$\frac{\partial G}{\partial t} + \frac{\partial G}{\partial v}\kappa v_t(\theta - v_t) + \frac{\partial G}{\partial I}v_t + \frac{1}{2}\frac{\partial^2 G}{\partial v^2}v_t^3\sigma_v^2 - rG = 0. \tag{116}$$

Solving the partial differential equation (114) in a specific region with proper boundary conditions gives the price of a variance call option. These details will not be explored in this paper, and more information are given by Broadie and Jain (2008, [26]).

6 Parameter Calibrations

Given the different pricing methods and strategies we have introduced so far, we need to estimate the various parameters in order to use them. Thus this chapter will be devoted on calibrating and thereby estimating the parameters from our different models. These estimations are based on two different calibration techniques, which is either option-based or price-based. Most of the models we are looking at will be option-based (which will be the B-S, MJD, Heston and the 3/2 model), and they will be estimated from option data from S&P500. The GARCH model is the only model which is price-based, and will be estimated from historical data. For this we will be looking at the S&P500 index over a historical time period.

6.1 European Options

In order to estimate the parameters from B-S, MJD, Heston or the 3/2 model we have gathered European style vanilla options from the S&P500. All options have been collected from the Historical Option Data database, obtained via www.historicaloptiondata.com. The raw dataset consists of 1 273 460 European options (both put and calls) on the S&P500. The range of the data collected is between 2014-01-01 and 2014-12-31, i.e. one year of daily data, which corresponds to 252 trading days of data.

6.1.1 Filtering Process

As we have quite a large dataset it is a good idea before calibration to perform a filtration to eliminate outliers and thus stabilize parameter estimation. This will be done to ensure that the estimations are valuable and to obtain efficient and reliable calibrations. There exists a variety of different filtration techniques, but we will in this paper apply the filtering scheme done by Bakshi et al. (1997, [3]). Both put and call options can be used but it suffices to only concentrate on call options. Thus we start off by eliminating all put options from our dataset. We then proceed to remove all non-traded options, which is done due to illiquidity. We will also remove the options with a time to maturity less than 7 days or greater than 182 days.

Thus in this study we start off with 1 273 460 European S&P500 options. We begin with removing all put options, which consists of 636 730 options. Next we remove the non-traded options due to illiquidity, removing a total of 524 069 options. In accordance with Bakshi et al. (1997, [3]) we also remove options that have a time to maturity which is less than 7 days. This is because these options are considered illiquid, even though it is worth noting that these types of options have increased since 1997. Since we are focusing on swaps in this thesis, which are usually traded in short maturities, we will also remove any options which has a higher maturity than $T > 0.5$ (182 calendar days). These options eliminated from our dataset are also in accordance with Sepp (2011, [39]).

There should be noted that it is possible to remove even more options, which would induce even better estimations. These conditions are however quite rare and will not remove a large portion of options, and the final estimations would therefore not drastically change. This includes removing options with negative bid-ask spread (or zero bid-ask spread), or that violates a specific no-arbitrage condition. We will in this paper not impose on all possible filtering schemes, as this is not a main concern of the thesis. More about the different removal techniques and implications are done by Bakshi et al. (1997, [3]). The filtration process is summarized in the table below:

Filter description	Removed options
Removal of European put options	636 730
Removal of non-traded options	524 069
Removal of options with maturity < 7 or > 182 days	24 147
Total removal	1 184 946

Table 2: Filtering Process

After the filtrations we are left with 88 514 European call options for the time period of 2014 (2014-01-01 to 2014-12-31). Given these options we are now able to properly estimate our parameters for our different models. Therefore we have about 351 call options available on average per day, at a total of 252 trading days. For the last day of our time period 2014-12-31, which is a selected day we will use to estimate our models, we have 396 call options available.

6.2 Option-Based Calibration

As it is in most calibration schemes, there will be several different ways to estimate a option-based diffusion model. Some of the examples in this area include a Nonlinear GLS, Nonlinear OLS, Simulated Method of Moments (SMM), Efficient Method of Moments (EMM) etc. An overview of the different available models are provided by Chernov and Ghysels (2000, [11]).

Alternatively there has been empirical evidence showing that an approach based on minimization of a nonlinear loss function using either deterministic or stochastic algorithms provide accurate results. We will thus in this thesis adopt the calibration techniques proposed by Mikhailov and Nögel (2003, [34]) based on this scheme.

This method is based on minimization of squared errors, and will be defined as the difference between the model imposed and the market observed prices. We will in thesis use an algorithm which falls within the class of deterministic algorithms, and in MATLAB goes by the name ”lsqnonlin”. For more information about the properties and the various structure of this minimization method we recommend Stamp and Thorsen (2011, [40]).

6.3 B-S Estimation

Being in a B-S world we are only concerned about 1 parameter, namely the diffusion term σ. The minimization procedure is thus quite simple and converges fast (about 5-10 iterations) towards its global minimum. The calibration results for the last day of our time period as well as the average are presented in the table below:

	σ
Average	0.1035
2014-12-31	0.1325

Table 3: Black-Scholes calibration results

As we can see the implied volatility parameter averages over the year of about 10%. During the last day the parameter is at 13.25%, indicating that the year may displayed some fluctuations. We can see that the volatility was quite a lot higher (about 3%) at the end of the year compared to the average value.

6.4 MJD Estimation

Working in a Merton Jump Diffusion model we now include the added jump process, which adds 3 extra parameters to fit. Thus we have to estimate σ, λ, a and b. Naturally this calculation requires more iterations compared to the above Black-Scholes in order to find a suitable fit, considering we have more parameters to estimate. The calculations in this case are quite dependent on starting values, and upper and lower bounds for the parameter are implemented to ensure proper values of our different parameters. The results of the calibration are shown below:

	λ	a	b	σ
Average	2.4075	-0.0479	0.0474	0.0682
2014-12-31	2.5825	-0.0727	0.0418	0.0779

Table 4: Merton Jump Diffusion calibration results

Here we see that the parameters for the last day is basically the same as the average over the whole year, with a slightly higher λ. We can also note that the jump intensity λ is quite high, if we compare it to studies by Stamp and Thorsen (2011, [40]) or Tankov and Voltchkova (2009, [41]). Since all these papers are looking at different time periods and are using different estimation methods, the various change in values are thus exclusive to each study and our parameters appear plausible.

6.5 Heston Estimation

For the calibrations in a Heston SV model, we need to fit 5 parameters. For this model we need to estimate κ, θ, v_0, σ and ρ. As we gain more parameters than before we naturally require more computational time compared to the earlier models. In order to receive reasonable parameters we add the following constraints to the minimization problem.

$$v_0, \kappa, \theta, \sigma > 0 \tag{117}$$

$$-1 \leq \rho \leq 1 \tag{118}$$

$$2\kappa\theta > \sigma^2 \tag{119}$$

These constraints are all in logic with what one would expect. The first constraint indicates that the parameters all are positive, which is very natural. The correlation term ρ can either yield a positive or negative correlation, and this restriction ensures that. The third constraint, which is known as the Feller condition, will ensure that the variance process is positive. Another important aspect of calibrating the model is the choice of pricing formula for European call options, as different approaches might yield different results. In this paper we implement the approach of Moodley (2005, [36]), which calculates the European call options using Fast Fourier Transforms (FFT). With more parameters to estimate as well as the pricing procedure the computational increases quite significantly in comparison to the earlier models. The table below presents the calibration results:

	κ	θ	σ_v	ρ	v_0
Average	9.4673	0.0181	0.5859	-0.6596	0.0096
2014-12-31	16.9965	0.0236	0.8956	-0.8995	0.0237

Table 5: Heston calibration results

Comparing the parameters to other studies we notice that the obtained results vary quite widely depending on the chosen time period, calculation method and the option data. From our estimations we have a fairly high speed of mean reversion, κ, and we can also see that we can assume negative correlation between the variance process and the underlying, as ρ estimates a negative value. In general our results appear plausible and give somewhat similar values compared to previous studies such as Bakshi et al. (1997, [3]) or Stamp and Thorsen (2011, [40]).

6.6 The 3/2 Estimation

As the 3/2 model lies in close relation to the Heston model, we again have to estimate 5 parameters. This method also requires quite a lot of computational power, as both the amount of parameters have increased and the complexity of the pricing schemes are heavier compared to the non stochastic volatility models of B-S and MJD. As we have stated earlier in chapter 5 we can rework the 3/2 model into a Heston model, where in this case the dynamics of $1/v_t$ follows a Heston process.

Thus we apply the same procedure as the Heston calibrations, and we only need to change the so called Feller condition. As pointed out by Drimus (2009, [14]), the transformations performed sets up a new condition which is defined as

$$2\kappa\theta \cdot \frac{\kappa + \sigma^2}{\kappa\theta} \geq \sigma^2 \tag{120}$$

or

$$\kappa \geq -\frac{\sigma^2}{2}. \tag{121}$$

Thus we apply this new non-explosion condition along with the original Heston conditions

$$v_0, \kappa, \theta, \sigma > 0 \tag{122}$$

$$-1 \leq \rho \leq 1. \tag{123}$$

Given these conditions we are now able to estimate a Heston process which after that can be transformed into the 3/2 model. The results of the calibrations are shown below:

	$\tilde{\kappa}$	$\tilde{\theta}$	$\tilde{\sigma_v}$	ρ	v_0
Average	0.1121	46.0703	-0.5768	-0.6237	0.0101
2014-12-31	0.1715	32.3396	-1.0452	-0.7365	0.0233

Table 6: 3/2 model calibration results

Looking at the parameters in comparison to the Heston model, we see that the comparable differences lies in the 3 first parameters as expected. Transformations for these 3 have been done accordingly to (75), (76) and (77), yielding suitable parameters for the pricing in the 3/2 model.

6.7 Price-Based Calibration

As we mentioned earlier the price-based estimation scheme will only be applied to the GARCH model. It is possible to use option-based methods like the ones we proposed earlier, but it might not be as effective. These estimations will be based on historical data, and we will look at the evolution of the spot price on the S&P500 index. This particular method is discussed by Heston and Nandi (2000, [24]), where they noticed that a discretized GARCH(1,1) outperforms different versions of the B-S model.

The calculations done will be performed using standard econometric methods. There are several methods in this area which are applicable to use, and no special emphasis has been put on the performance of the different available choices. Some of the possible approaches include Feasible Generalized Least Squared (FGLS), Generalized Method of Moments (GMM) or Maximum Likelihood (ML). In this thesis we will proceed to estimate the parameters in a discrete GARCH setting by the means of ML (Maximum Likelihood).

6.8 GARCH Estimation

In this section we present the estimation results for our GARCH(1,1) model. We will estimate 3 parameters (ω, β and α) in a discrete setting, which we will use in order to obtain the continuous estimations. To estimate the model we have used daily log-return observations on the S&P500 index over the course of 4 years (2010-2013). The amount of years are chosen to ensure asymptotic properties of the estimator.

The gathered data has been implemented into MATLAB, where we have fitted a GARCH(1,1) model to obtain the parameters. After the estimation we shift the sample window by 1 observation, creating a forecast over 2014. This yields 252 datapoints, and the average estimated parameters over the year as well as the last set estimated on 2014-12-31 are shown below:

	ω	β	α
Average	4.1206e-06	0.8356	0.1239
2014-12-31	2.39023e-06	0.8291	0.1458

Table 7: Discrete GARCH(1,1) calibration results

As we can see the model is stationary, as $\alpha + \beta < 1$ and all parameters are strictly positive. Comparing the results to other academic GARCH estimation papers such as Haug et al. (2002, [22]) or Stamp and Thorsen (2011, [40]), we can see that our parameters are within reasonable limits. The difference to those papers are of course expected since we are looking at different year intervals, Haug et al. are looking at the years between 1996 and 2001 while Stamp and Thorsen have gathered data between 2006 and 2010.

For now we have estimated the parameters in a discrete GARCH(1,1) setting, and what leaves us is converting it into the continuous case. Converting into the diffusion model we use the relationships presented in equations (66), (67) and (68). Along with those we have that $dt = 1/252$ to reflect daily observations. The results of this transformation yields:

	V	ξ	κ	θ	σ_v	v_0
Average	1.0173e-04	9.4240	10.206	0.0256	5.7086	0.01
2014-12-31	9.5228e-05	11.4186	6.3252	0.0239	7.4707	0.0092

Table 8: Continuous GARCH(1,1) calibration results

Comparing those results to other studies by Stamp and Thorsen (2011, [40]) and Haug et al. (2002, [22]) we receive fairly similar estimations. Some parameters such as κ varies quite a bit along these studies, but are within suitable ranges because of the different time periods used.

Thus having estimated the parameters we are now ready to price volatility derivatives in the framework of a GARCH model, and compare it to the option-based estimations of the other models.

7 Empirical Evidence

As for all models, it would be difficult to talk about their performance if we don't have any valid parameters and were able to properly implement them. Thus we will for this chapter apply our estimated parameters and use them along with our specific formulas from our different models we have arrived at so far, and see how they perform in the case of volatility derivatives. The models that we will be looking at are the ones that we have estimated parameters for, namely the Black-Scholes (B-S), Merton Jump Diffusion (MJD), Heston, GARCH as well as the 3/2 model. We will mainly be focused on evaluating the performance of variance and volatility swaps with maturity lengths of 1, 3 and 6 months. The choice of maturity dates T are done based on that these contracts are some of the most commonly traded in the market.

When looking for empirical evidence, one normally evaluates model performance against the ability to fit or replicate observed market prices. Secondly in order to consider a model attractive one concerns the ability to fit model features onto other products, which in our case are swaps written on variance or volatility. We will thus devote this section to use our estimated parameters in the previous chapter to price swaps, and evaluate their performance against each other. We will for the variance swap use our closed form solutions which we have derived in this paper. As there exist no such formulas in the case of volatility swaps, we will instead rely on Monte-Carlo simulations as well as numerical integration techniques to obtain the strike price for the volatility swap.

7.1 Variance Swaps

Most variance swaps in the market are written on realized variance and calculated from log-returns and a daily discrete sampling process ($AF = 252$). As we have derived several formulas for the continuous strike price of a variance swap across different models, we will in this section use our calibrated results gathered from the previous chapter and implement this into our solution formulas for the variance strike price.

Worth reminding is that the variance strike is often quoted in units of volatility squared, i.e. $20\%^2 = 4\%$ or sometimes also denoted as $20\%^2 = 400$ variance points. We will start by looking at a specific day, which in our case will be the last day of our time period 2014-12-31, to give the reader a taste for the range of swap rates. After this we also present the whole sample period (the year of 2014) along with some analysis and conclusions. The continuous swap rates on 2014-12-31 are given in the table below:

Maturity	B-S	MJD	Heston	GARCH	3/2
$T = 1/12$	13.25	15.56	15.38	12.28	15.27
$T = 3/12$	13.25	15.56	15.37	14.13	15.30
$T = 6/12$	13.25	15.56	15.36	15.02	15.34

Table 9: Variance swap prices on 2014-12-31

As we can see from the table above the only model to induce some variety depending on maturity is the GARCH model. As we are only looking at 1 day, the estimations for the Heston or 3/2 model becomes almost the same, and the Heston model are even going down as the maturity time increases. These prices are of course a bit skewed since we are only looking at one day, and some estimations may give a false image of the certain trends for the different models. Although we can give some general remarks that the prices are higher in MJD, Heston and 3/2 at smaller maturities. The GARCH models catches up when increasing the maturity, and all prices for $T = 6/12$ are all around 15% volatility across all models except for the B-S.

Next up we look over the whole sample period (2014-01-01 to 2014-12-31), to get a better grasp of the swap prices. Looking over just one day might give a false indication as how these models perform, as we are looking over one day which is only about 400 options (this day in particular could have been an extreme). Therefore we will proceed to look at the variance swap prices using the average estimated parameters over the year 2014, which is presented in the table below:

Maturity	B-S	MJD	Heston	GARCH	3/2
$T = 1/12$	10.35	14.62	11.05	11.16	10.06
$T = 3/12$	10.35	14.62	12.18	12.85	10.08
$T = 6/12$	10.35	14.62	12.77	13.94	10.12

Table 10: Variance swap prices over the whole sample period

Here we see that for a 1 month swap the volatility levels are all almost the same except in the MJD case. In the 3/2 model we have very small changes over the different maturities, compared to the Heston and GARCH model. Again we see the largest fluctuations are in the GARCH model, where we go from around 11% up to 14% depending on the length of the swap. We can also see that the swap rates are quite high in the MJD model compared to the other models.

Comparing the average prices over the year 2014 against the last sample day 2014-12-31, we can see some general trends. Overall the volatility is a few percentage points lower in the average prices compared to the last day, as indicated by most models. If we start by looking at $T = 1/12$, the last day indicates a price around 15% for most models (MJD, Heston and 3/2) while looking at the average prices the volatility level is ranging between $10\% - 11\%$. The same can in general be said for the other maturities as well, although the swap prices in the average time period does seem to raise as the maturity increases, and the Heston and GARCH model moves a bit closer to the MJD model.

7.2 Volatility Swaps

Although the volatility swaps are easier to understand in practice and more intuitive, the pricing framework are not as straightforward as the variance swap. The square root function of the volatility swap adds an element which mathematically becomes a lot more complex. Since there are no pricing formulas for the volatility swap, we rely on numerical methods in order to price this swap.

We will calculate the pricing process in the Black-Scholes model as well as the Merton Jump Diffusion model based on the analytical formula we presented earlier, which relies on Laplace transformations and numerical integration. For the stochastic volatility models we use Monte-Carlo simulations in order to approximate the volatility strike price. The Monte-Carlo simulations used will be based on 100 000 simulations with 20 steps per day. The volatility swap prices for the last day of our time period are given in the table below:

Maturity	B-S	MJD	Heston	GARCH	3/2
$T = 1/12$	13.24	11.83	10.4452	7.1593	10.0368
$T = 3/12$	13.24	13.36	10.5386	7.7858	10.0605
$T = 6/12$	13.24	14.20	10.6668	8.2674	10.0875

Table 11: Volatility swap prices on 2014-12-31

Comparing the variance swap findings to the table above we now have generally lower strikes for most models over different maturities. Not surprisingly the Black-Scholes prices have not changed much, and stays around 13% for both the variance and volatility swap. The MJD model however increases quite heavily depending on the maturity, and contracts over 1 month for the MJD are fairly close to the SV models. We also see that for this day the volatility level of the GARCH model are lower than the other models, which is in line with the results of Stamp and Thorsen (2011, [40]). In general all models produce a lower strike price compared to their variance counterpart which is suspected due to the convexity adjustment discussed earlier.

Again we also look at the average behaviour over the time period to get a better overview. The average swap prices for a volatility swap across our different models are given in the table below:

Maturity	B-S	MJD	Heston	GARCH	3/2
$T = 1/12$	10.34	9.60	7.47	8.3060	7.0978
$T = 3/12$	10.34	10.68	8.26	9.3832	7.1174
$T = 6/12$	10.34	11.30	8.76	9.9857	7.1402

Table 12: Volatility swap prices over the whole sample period

In this content we see that the GARCH model performs higher than the last day, something that is not occurring in the other models. In the other 4 models we have a slight decrease in volatility levels for the average time period, which indicates that the volatility is on average lower than the last day. This is also reminiscent in the variance swap, where the average prices were at a 2-4% lower level.

Looking at the values of the volatility swap, we can see that all models imply the same patterns, except the Heston model. As the B-S and 3/2 model are quite stable through different maturities and the GARCH and MJD are steadily increasing for higher maturities, this is not the case in the Heston model. For the last day, the Heston swap prices are almost locked around 10.5%, compared to a slight increase for the average prices.

Overall we see that the stochastic volatility models offers slightly different swap prices compared to the B-S and MJD models. The GARCH model with its price-based approach fluctuates a lot over both swaps and does not generally price within the same range as the option-based counterparts.

8 Concluding Remarks

In this thesis we have investigated the field of volatility derivatives, a financial area that has become quite popular in the recent years. Mainly focused on the variance and volatility swap, we have introduced the pricing strategies for some of the most popular models available as well as given a "model-independent" replicating scheme to price variance swaps which we saw only requires the use of European call and put options.

To properly evaluate those models we have estimated the various parameters across our models using either a option-based or price-based approach. For the B-S, MJD, Heston and 3/2 model we gathered European call and put options from the S&P500 during 2014 and applied a squared error minimization technique to estimate the various parameters. For the price-based approach in which the GARCH model was calibrated we used log-returns on the S&P500 index over the course of 4 years (2010-2013), together with the maximum likelihood calibration technique for the estimations.

Looking over the literature we found that the variance swap was well covered in terms of analytical formulas and we were able to find closed form solutions for our models. As goes for the volatility swap there were limitations due to the complexity that the square root function opposes, and no general solution formulas could be obtained. Thus we had to rely on numerical integration techniques as well as Monte-Carlo simulations in order to evaluate the pricing scheme for a volatility swap across all models.

Performance wise we could note that the GARCH model stood out and did in general not perform exceptionally well. The Heston and 3/2 model produced mostly the same patterns depending on the time stamp, and gave a decent performance. The B-S model was typically outside of the range of the other models, and was not entirely unexpected fairly underwhelming. The added extension with jumps for the MJD model performed a bit better compared to the original Black-Scholes model.

In conclusion it appears that the option-based calibration techniques has a superior performance. The price-based method which we used for GARCH showed some flaws in the pricing of swaps, and in general the option-based counterpart that has a forward-looking ability in the options seem to prevail over relying on historical data.

9 Extensions and Further Research

Our approach in this thesis have been to apply some of the most popular models available in the area, as well as putting the ”standard” yet fairly simple assumptions on the asset. We assumed for instance that the stock pays no dividends, which is in the market usually not the case. Thus extending to allow for cash dividends in the pricing scheme would be a natural progress. A study by Klassen (2009, [28]) derives the price for a variance swap in case of cash dividends, and are using the replication scheme to find the variance strike price using this extension.

Most of our formulas presented in this thesis are done in a continuous matter, something that is typically not applicable when looking at market data. Although the error between a discrete setting compared to its continuous counterpart might be low, one should remember that even a 1% error could implicate major losses when entering contracts of higher values. Thus a natural extension would be to analyze the different settings against each other in the framework of volatility derivatives. Both Lian (2010, [30]) as well as Broadie and Jain (2008, [27]) describes the effect of this phenomenon and applies it to variance and volatility swaps.

Another improvement would be to use stochastic interest rates, since in the market those rates usually fluctuates over time. Instead of having a regular constant rate, Hörfelt and Torné (2010, [25]) removes the typical constant assumption and presents the value of variance swap in case of stochastic interest rates. They look into the case where the interest is driven by its own Itô process, which obviously makes for a more advanced pricing structure which may result in reflecting the market better.

Other aspects would be to improve the assumptions about the models, or the components in them. As we have mentioned earlier, one could exchange the geometric Brownian motion for a more general Lévy process. Carr and Lee (2009, [7]) expands the framework and applies a time-changed Lévy process in order to price variance swaps.

One extension that we touched upon in this thesis was combining the MJD model and the stochastic volatility models. Using these two together are quite natural since both the jump aspect and the stochastic volatility implementation are both very reasonable assumptions and have been seen to be empirically viable. Evaluating the SVJ and SVJJ model would make pricing more realistic and possible better. Stamp and Thorsen (2011, [40]) applies pricing tools for both the variance and volatility swap in the case of the SVJ and SVJJ as well as some of the common stochastic volatility models.

There are of course even more extensions that can be done to further enhance the pricing process, as well as other relevant aspects. One of those areas include looking at different parameter calibration methods, because even with more advanced models the necessity for reasonable estimation-techniques are equally relevant in the pricing of volatility derivatives.

10 Appendix

10.1 A - Compound Poisson Process

We use in the paper Itô's lemma for a compounded Poisson process and the proof is rather simple. Assumptions include that the process

$$S_t = \sum_{i=1}^{N(t)} X_i \tag{124}$$

is adapted and that the function f is continuous. To illustrate this fact assume that there is only one jump at time t_j then

$$\begin{aligned} & f(S_t) - f(S_0) = \\ & \big(f(S_t) - f(S_{t_j+\epsilon})\big) + \big(f(S_{t_j+\epsilon}) - f(S_{t_j-\epsilon})\big) + \big(f(S_{t_j+\epsilon}) - f(S_0)\big) . \end{aligned} \tag{125}$$

The assumed continuity causes most of the terms to vanish, so that we end up with

$$f(S_t) - f(S_0) = f(S_{t_j+\epsilon}) - f(S_{t_j-\epsilon}). \tag{126}$$

Letting ϵ go towards zero give us Itô's lemma for one jump, which can then be extended to more jumps by induction with all the paths between the jump vanishing due to continuity.

10.2 B - Brockhaus-Long Convexity Approximation

A proof concerning the convexity error regarding the error between variance and volatility swaps. We start out by defining a square root function F as:

$$F(x) = \sqrt{x}. \tag{127}$$

Calculating the derivatives we know that

$$F'(x) = \frac{1}{2\sqrt{x}} \tag{128}$$

$$F''(x) = -\frac{1}{4\sqrt{x}^3}. \tag{129}$$

Using a Taylor expansion for F around x_0 we get

$$\begin{aligned} F(x) &\approx F(x_0) + F'(x_0)(x - x_0) + \frac{1}{2}F''(x)(x - x_0)^2 \\ &\approx x_0^{1/2} + \frac{x - x_0}{2\sqrt{x_0}} - \frac{1}{8}\frac{(x - x_0)^2}{\sqrt{x_0}^3} \\ &\approx \frac{x + x_0}{2\sqrt{x_0}} - \frac{(x - x_0)^2}{8\sqrt{x_0}^3}. \end{aligned} \tag{130}$$

Choosing $x = V$ and $x_0 = E[V]$ we get

$$\sqrt{V} \approx \frac{V + E[V]}{2\sqrt{E[V]}} - \frac{(V - E[V])^2}{8\sqrt{E[V]}^3}. \tag{131}$$

Finally we take expectations on both sides which yields

$$E[\sqrt{V}] \approx \frac{E[V] + E[V]}{2\sqrt{E[V]}} - \frac{E[(V - E[V])^2]}{8\sqrt{E[V]}^3} \tag{132}$$

or simplified

$$\sqrt{E[V]} - E[\sqrt{V}] \approx \frac{Var(V)}{8\sqrt{E[V]}^3}. \tag{133}$$

Transforming back to our notations in the text we immediately get that $E[\sqrt{V}] = K_{vol}$ as well as $\sqrt{E[V]} = \sqrt{K_{var}}$, which is their respective definitions.

References

[1] Applebaum, David *Lévy Processes - From Probability to Finance and Quantum Groups* American Mathematical Society, 2004

[2] Bakshi, G., Ju, N., Yang, H. *Estimation of Continuous Time Models With an Application to Equity Volatility* Journal of Financial Economics, 2006

[3] Bakshi, G., Cao, C., Chen, Z. *Empirical Performance of Alternative Option Pricing Models* The Journal of Finance, 1997

[4] Black, Fischer & Scholes, Myron *The Pricing of Options and Corporate Liabilities* Journal of Political Economy, 1973

[5] Breeden, D. & Litzenberger, R. *Prices of State-Contingent Claims Implicit in Option Prices* Journal of Business, October, 1978

[6] Brockhaus, Oliver & Long, Douglas *Volatility swaps made simple* Risk Magazine, 2000

[7] Carr, Peter & Lee, Roger *Pricing Variance Swaps on Time-Changed Lévy Processes* 2009

[8] Carr, Peter & Lee, Roger *Volatility Derivatives* Annual Review of Financ. Econ., 2009

[9] Carr, Peter & Wu, Liuren *Time-Changed Lévy Processes and Option Pricing* Journal of Financial Economics, 2004

[10] Chicago Board Options Exchange *The Volatility Index (VIX)* http://www.cboe.com/micro/vix/historical.aspx

[11] Chernov, M. & Ghysels, E *Estimation of Stochastic Volatility Models for the Purpose of Option Pricing* Computational Finance, 2000

[12] Cox, John C. & Ross, Stephen A. *The Valuation of Options for Alternative Stochastic Processes* Journal of Financial Economics, 1976

[13] Demeterfi, Kresimir et al. *More Than You Ever Wanted To Know About Volatility Swaps* March, 1999

[14] Drimus, G. *Option on Realized Variance by Transform Methods: A non-affine Stochastic Volatility Model* University of Copenhagen, 2009

[15] Drimus, G. *Volatility-of-Volatility Perspectives: Variance Derivatives and Other Equity Exotics* University of Copenhagen, 2011

[16] Engle, Robert F & Mezrich, J. *Grappling With Garch* Risk Magazine, 1995

[17] ENGLE, ROBERT F. & NG, VICTOR K. *Measuring and Testing the Impact of News on Volatility* Journal of Finance, 1993

[18] FINANCE, YAHOO *Daily Closing Prices of S&P500* https://finance.yahoo.com/

[19] FELLER, W. *Two Singular Diffusion Problems* Annals of Mathematics, 1951

[20] GANGAHAR, ANUJ *Volatility Becomes an Asset Class* Financial Times, May 2006

[21] GATHERAL, JIM *The Volatility Surface - A Practitioner's Guide* John Wiley & Sons, Inc., 2006

[22] HAUG, ESPEN G. ET AL. *GARCH and Volatility Swaps* Wilmott Magazine, 2002

[23] HESTON, S. *A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options* The review of Financial Studies, 1993

[24] HESTON, S. & NANDI, S. *A Closed-Form GARCH Option Valuation Model* The review of Financial Studies, 2000

[25] HÖRFELT, P. & TORNÉ, O. *The Value of a Variance Swap - A Question of Interest* Risk, June 2010

[26] JAIN, ASHISH & BROADIE, MARK *Pricing and Hedging Volatility Derivatives* January, 2008

[27] JAIN, ASHISH & BROADIE, MARK *The Effect of Jumps and Discrete Sampling on Volatility and Variance Swaps* International Journal of Theoretical and Applied Finance, 2008

[28] KLASSEN, T. *Pricing Variance Swaps with Cash Dividends* Wilmott Journal, 2009

[29] KOLEVA, DESSISLAVA *Option Pricing under Heston and 3/2 Stochastic Volatility Models: an Approximation to the Fast Fourier Transform* Aarhus University, 2012

[30] LIAN, GUANGHUA *Pricing Volatility Derivatives with Stochastic Volatility* University of Wollongong, 2010

[31] MANDELBROT, BENOIT *The Variation of Certain Speculative Prices* The Journal of Business, 1963

[32] MATSUDA, KAZUHISA *Introduction to Merton Jump Diffusion Model* The City University of New York, 2004

[33] Matytsin, Andrew *Modelling Volatility and Volatility Derivatives* J.P. Morgan, 1999

[34] Mikhailov, S. & Nögel, U. *Heston's Stochastic Volatility Model Implementation, Calibration and Some Extensions* Wilmott Magazine, 2003

[35] Merton, Robert C. *Option Pricing When Underlying Stock Returns Are Discontinuous* Journal of Financial Economics, 1976

[36] Moodley, Nimalin *The Heston Model: A Practical Approach with Matlab Code* University of the Witwatersrand, 2005

[37] Schultes, Renée *Volatility kills market for volatility swaps* The Wall Street Journal Europe, 2008

[38] Schürger, Klaus *Laplace Transforms and Suprema of Stochastic Processes* University of Bonn, 2002

[39] Sepp, Artur *Pricing Options on Realized Variance in the Heston model with Jumps in Returns and Volatility - Part II: An Approximate Distribution of Discrete Variance* Journal of Computational Finance, 2011

[40] Stamp, Emil S. F. & Thorsen, Thomas F. *Pricing of Variance and Volatility Swaps* Aarhus University, 2011

[41] Tankov, P. & Voltchkova, E. *Jump-Diffusion Models: A Practitioner's Guide* Banque et Marches, 2009

[42] Øksendal, Bernt *Stochastic Differential Equations, An Introduction with Applications* Springer-Verlag Berlin Heidelberg, 2010