Study of Modifications of The Option Pricing Theory: A Practical Approach[1]

An Honours Thesis at the

University of New Mexico

Author:
Juan Pablo Madrigal Cianci

Supervisor:
Dr. Jens Lorenz

A thesis submitted in fulfilment of the requirements
for the honours degree of B.Sc in Mathematics
at the
Department of Mathematics and Statistics.

April 2015

[1]Presented at the Southwestern Undergraduate Mathematics Research Conference (SUnMaRC) at El Paso, TX. February, 2015.

Thanks to my mother, the greatest and kindest person I know. I would not be here without her efforts and unconditional love. All my accomplishments are hers. To my family and my father, for all the support, guidance, and advice. To my friends, for all the good times.

I would like to thank my advisor, prof. Jens Lorenz, for all the invaluable help he has given me regarding this project.

I would also like to thank prof. Nitsche, for all the effort she has put on the MCTP program, which has benefit so many students, including myself.

Finally, I would like to thank all the wonderful faculty from the University of New Mexico, Universidad Nacional de Colombia, and the Universidad de Antioquia that I've come across these years.

Mathematical reasoning may be regarded rather schematically as the exercise of a combination of two facilities, which we may call intuition and ingenuity.

—Alan Turing

Abstract

This research project is a collection of several ideas, condensed mainly into three broader concepts. These ideas cover what can be considered as perhaps the milestones of the option pricing theory.

The first idea results from exploring the consequences of modifying an assumption of one of the most important results in financial mathematics. I am talking about modifying the assumption of constant volatility in the Black-Scholes model. I explored different ideas and different parameterizations of the volatility, σ, and then I proceeded to compare the results of the Black-Scholes model using these parameterizations and the original Black-Scholes model in order to asses the performance of each method. A similar work was examined by Guernsey (2014) who modeled the volatility as a discontinuous function. However, I believe that my project provides a more comprehensive approach, using different parametrization and methods. After presenting a model, the new parametrization of the volatility is then used on the Black Scholes partial differential equation and is then solved using a Crank-Nicholson scheme. The price returned by the difference scheme is then compared to the market price and the price predicted by the constant volatility model.

The other two parts are studies and practical approaches on other option pricing methods, called the Heston Model and the GARCH option pricing model. Both these models are called stochastic volatility models (SVM).

For the Heston model, I start by describing and analyzing said model. Then, following the steps of Moodley (2005) and Rouah(2013), I proceed to discuss, implement and test several calibrations for this model. Finally, I simulate it and compare results with the market price and the price predicted by Black and Scholes.

Finally, for the last part, I introduce an option pricing method using a General Autoregressive Conditional Heteroskedasticity (GARCH) model. I implement the model as proposed by Duan (1993) and Heston & Nandi (2000). Finally, I implement an improvement proposed by Chorro et al, on which we consider the innovations for the GARCH model to follow a generalized hyperbolic distribution. As with the other two parts, I will compare this model with the market price and with the standard Black Scholes Model.

I consider that this is a rather interesting approach. Not only it proposes alternatives to the option pricing theory, but it also implements alternatives and their respective improvements. Also, as opposed to the general literature, I shall use a variety of option

chains to test my models. The option chains include companies such as Google, Facebook and Yahoo, as opposed to the widely studied S&P 500. This is important because it is interesting to test these option pricing models in other companies individually, rather than in an economic indicator.

Contents

Part I

Background

Chapter 1

Introduction and Motivation

1.1 Brief History

The financial world is, by nature, full of risk. Financial markets have been, for the most part, rather unpredictable. The industry went through a paradigm change since the mid 80's when firms at Wall Street started hiring mathematicians, physicists, and engineers to model and provide complex financial instruments as well as trying to minimize the risk associated to financial decision making.

Many things have changed since this became a common and accepted practice in the industry; "quants" (The scientists and engineers working for the financial industry) have adopted a wide variety of models from physics and math and applied them to quantify operations. This lead to a rapid expansion in this particular field of knowledge, as well as creation of graduate programs in this area in order to supply more candidates to the industry. The demand for this type of professionals seems to keep increasing faster than its supply; according to different sources, such as specialized forums, work networks, industry related press, etc, many of the highest regarded financial mathematics/financial engineering programs have been able to place above 80 % of their graduates upon three months of competition[1].

However, not everything is perfect. There have been 3 market crashes in the last 30 years mainly because the financial industry still relies a big deal on human made decisions, which are not always perfect. As an example, the last crisis (2007-2008) gave negative returns of around 40% , which translates into more than 6 standard deviations

[1] Sites include but are not limited to Quantnet, Business insider, QuantOasis, Mergers and Inquisitions, NYU courant institute of Math, Carnegie Mellon Tepper School of Business, Columbia School of operation research and industrial engineering, Columbia Mathematics of finance, etc.

of the expected change, making an (theoretically) almost impossible event (probability of less than 10^{-10}) happen, costing millions of dollars and million of jobs worldwide. It is because of events like this that we are always in need of better models to predict prices and therefore minimize risk and have better measures, so we can be able to make more accurate decisions.

This is where this paper comes to play. In this work, I shall present some modifications and discuss some alternatives to the commonly used methods in the industry. I will present the mathematical foundation underlying the methods as well as a comparison of the results of such simulations and an indication of under what circumstances they should be used.

The usual practice of financial mathematics, and the aspects which gives it its paramount importance in today's financial industry, is the fact that we develop and test models on liquid assets (asset that exist today), so we can model and predict future assets and derivatives.

1.2 Important Definitions

Through out this text there are some terms that might not be quite familiar for a reader that is not interested in the financial market or that has no financial background whatsoever. Even though this text is about the mathematics that shapes the financial model, there are some *technical words* that have to be used. To learn more about financial terms, I've found that *http://www.investopedia.com* provides some very clear and insightful definitions.

Derivative (Finance): a derivative is a contract that derives its value from the performance of an underlying entity. This underlying entity can be an asset, index, or interest rate, and is often called the "underlying". Derivatives can be thought as "bets" on the state of the underlying, and can be relatively simple, like betting on the price of the stock in the future, or can get rather complicated, like betting that the price of stock X will go up while the price of stock Y will go down and, at the same time, it will rain in NY; this is not a joke! According to Bloomberg, weather derivatives are a 19 Billion dollar market[2] This instruments are rather "tricky"; the use and abuse of derivatives involving credit default swaps was one of the detonators of the financial crisis of 2008[3]

[2] http://www.bloomberg.com/apps/news?pid=newsarchive&sid=aK9BWCFSRPnM
[3] *Inside Job*, 2010. Documentary

Option: A financial derivative that represents a contract sold by one party (option writer) to another party (option holder). The contract offers the buyer the right, but not the obligation, to buy (call) or sell (put) a security or other financial asset at an agreed-upon price (the strike price) during a certain period of time or on a specific date (exercise date).

Option Parameters: Options have different parameters. The most important are the time to expiry or maturity, which is the time on which the option contract obtains it's maturity (or time limit depending on the type of option). Asset price, denoted S is the price of the stock or other asset (bond, mortgage, commodity, etc) on which we are calculating the option. Finally, the other important parameter is the strike price, denoted by K. The price at which a specific derivative contract can be exercised. When an option contract is signed, both the buyer and the seller know all of these parameters.

Risk Free Measure: Also called equivalent martingale measure is a probability measure such that each share price is exactly equal to the discounted expectation of the share price under this measure.

Volatility: Volatility measures the changes in time of the price of a financial instrument. Big changes will yield a high volatility while small changes will produce a smaller volatility. Volatility is high during economic crisis and bubbles and it is relatively low otherwise. It is, in a way, a measure of risk. In general volatility can't be observed, only calculated.

Implied Volatility: It is the volatility that we would need to plug into the Black-Scholes equation in order to get the real market price. This can be calculated using a root finding scheme and the Black-Scholes equation.

Chapter 2

Theory: Derivation of The Black Scholes Model

In this chapter I shall explain how the Black and Scholes model was obtained. There are different approaches that, ultimately, will yield the same result. However, as I consider it the most clear, I will follow the same approach as Wilmott (1995). In order to obtain the Black Scholes PDE, we need to make use of Ito's lemma:

Important Results From Ito's Calculus: Consider the following SDE

$$\frac{dS}{S} = \sigma dX + \mu dt, \tag{2.1}$$

where dX is a Wiener process. This is the term that contains the "randomness" of the model. Here, σ and μ are constants and S is our variable of interest, which we will denote as the underlying price. The wiener process has the following form:

$$dX \sim N(0, dt). \tag{2.2}$$

That is, it is normally distributed with mean 0 and standard deviation dt. Ito's calculus provides the following results[1]:

$$dX \sim O(\sqrt{dt}), \tag{2.3}$$

$$dX^2 \to dt \;\; as \;\; dt \to 0. \tag{2.4}$$

It also states that $dtdX$ and dt^2 tend to 0 faster than dt and dX^2.

[1] Wilmott (1995) provides a more in depth explanation of these results.

Squaring equation (2.1) gives:

$$dS^2 = \sigma^2 S^2 dX^2 + 2\sigma\mu S^2 dt dX + \mu^2 S^2 dt^2. \tag{2.5}$$

Then, using the above results (equations (2.3) and (2.4)), we get that

$$dS^2 \to \sigma^2 S^2 dt, \;\; as \;\; dt \to 0, \tag{2.6}$$

which is a result that we shall use shortly.

The Black-Scholes model is based on some assumptions, namely:

- The asset price, S, follows a log-normal random walk with the same form as equation (2.1). In this case μ is the drift coefficient, and σ is the volatility coefficient of the process.
- The risk free interest rate r of the underlying asset and its volatility, σ are known functions of time during the life of the option. [2]
- there are no transaction costs associated to hedging a portfolio.
- The underlying asset pays no dividend during the life of the option.
- there's an arbitrage free world.
- trading is continuous over the underlying asset, that is, it is possible to trade any amount of the underlying.
- short selling is permitted and the underlying assets are divisible.

Having stated the conditions of the Black-Scholes world, we can then derive the model. Let's suppose that we have an option (or a portfolio, a collection of options, etc) where the price, V, only depends on time to expiration (maturity) and underlying price (generally a stock). That is, an option given by $V = V(S, t)$. Expanding $V(S + dS, t + dt)$ in a Taylor series about (S, t):

$$dV = \frac{\partial V}{\partial S} dS + \frac{\partial V}{\partial t} dt + \frac{1}{2}\frac{\partial^2 V}{\partial S^2} dS^2 + \dots \tag{2.7}$$

[2]This is not particularly true; we can have different rates of return as seen on Wilmott, 1995 and Hull, 2003) Also, the study and parametrization of non-constant volatilities is the scope of the next chapter.

Replacing dS and dS^2 by equations (2.1) and (2.6) gives that:

$$dV = \sigma S \frac{\partial V}{\partial S} dX + \left(\mu S \frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} dS^2 + \frac{\partial V}{\partial t} \right) dt. \tag{2.8}$$

Now, we proceed to construct a portfolio that has an option and Δ units of the underlying asset. The value of the portfolio is then going to be given by

$$\Pi = V - S\Delta. \tag{2.9}$$

In a time step, we can see that the change in the portfolio is going to be given by:

$$d\Pi = dV - \Delta dS. \tag{2.10}$$

If we put the last three equations together we get that:

$$d\Pi = \sigma S \left(\frac{\partial V}{\partial S} - \Delta \right) dX + \left(\mu S \frac{\partial V}{\partial S} + \frac{1}{2} \frac{\partial^2 V}{\partial S^2} + \frac{\partial V}{\partial t} - \mu S \Delta \right) dt. \tag{2.11}$$

To eliminate the random component of this random walk, we make the following assumption, called “Delta-Hedging Strategy”:

$$\Delta = \frac{\partial V}{\partial S}. \tag{2.12}$$

This leaves us with the deterministic price for the portfolio, given by

$$d\Pi = \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) dt. \tag{2.13}$$

Here, we are faced an argument due to the laws of supply and demand that is pretty reasonable and straightforward. In a risk-less world (as stated before) the return on Π, invested at a rate r equals to an amount $r\Pi dt$. For a more detailed argument about this value and its implication in the arbitrage free world see Wilmott, 1995. We then have:

$$r\Pi dt = \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) dt. \tag{2.14}$$

Dividing by dt and subtracting, we finally get the Black Scholes partial differential equation, given by

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} - rV = 0. \tag{2.15}$$

This equation plays an important role in finance and in the content of this paper. As we can see, the way it was derived doesn't depend much on the of σ, therefore, we can introduce a σ function given by $\sigma(S,t)$, making equation (2.15) of the form:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma(S,t)^2 S^2 \frac{\partial^2 V}{\partial S^2} dt + rS\frac{\partial V}{\partial S} - rV = 0. \tag{2.16}$$

It can be shown that the solution for (2.15) is given by

$$V(S,t) = SN(d_1) - Ee^{-r(t-\tau)}N(d_2), \tag{2.17}$$

where $N(*)$ is the standard normal cumulative distribution function and d_1 and d_2 are given by

$$d_1 = \frac{\log(S/E) + (r + \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}, \tag{2.18}$$

and

$$d_2 = \frac{\log(S/E) + (r - \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}. \tag{2.19}$$

The solution of equation (2.16) will clearly take a more complex form than the solution for equation (2.15).

Levy and Paraz (1998) as well as Lorenz and Qiu (2009) found a closed form solution when sigma is modeled as a discontinuous function bounded between two numbers. In general, we are not guaranteed that there is a closed form solution depending on the form the sigma. Therefore, we resort to solve the equation using numerical methods, in particular the fully implicit Scheme and the Crank Nicholson Scheme. The Black Scholes PDE can be transformed into a heat equation (Hull 2003), and it is well known that the fully explicit scheme for this type of equation is unstable unless the time step is chosen to be very small (Sauer, 2011).

Part II

Models, Approaches, and Experiments

Chapter 3

Modeling Volatility: An Empirical Approach

As we have discussed before, the Black-Scholes model in its closed form takes the volatility as a constant. It comes natural to think that the volatility shouldn't be a constant because of the nature of the market, and that options with different expirations and/or strike prices are traded with different volatilities. Moreover, we can see that the volatility surface is not a constant, as it changes for different values of S (asset price), K (strike price), and t (time to maturity), as we can see in the following figure:

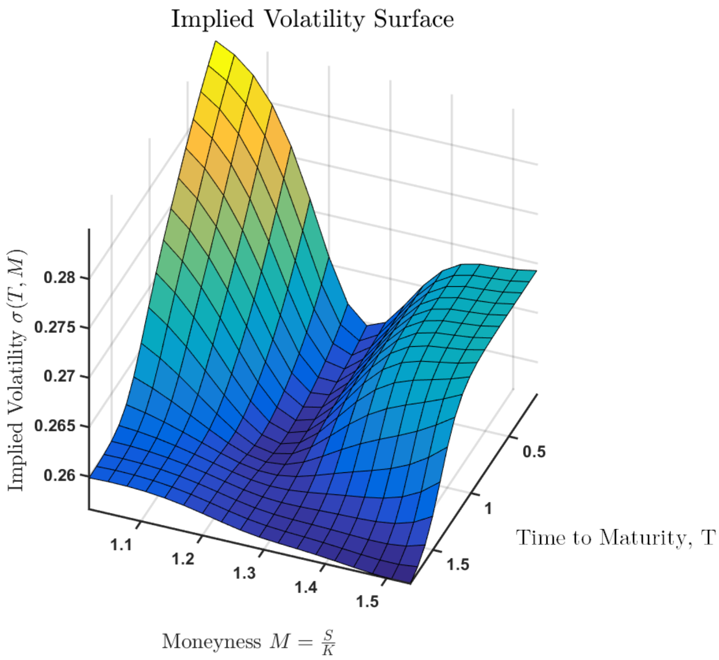

FIGURE 3.1: Implied volatility for Facebook, Inc.

We can see from the figure above that the volatility doesn't behave as a constant surface, but rather as a curved surface. This is called the Volatility Smile and it has been vastly studied[1]. It is because of this phenomenon that we will propose some empirical approaches to model the volatility.

3.1 Modeling and Calibration

To asses the efficiency of these models, we will clearly have to test their predictability against two different parameters. The first one is against the market price of the option, which is fairly easy, considering that the market prices for option chains are readily available on sites like Yahoo, Google, NASDAQ, etc. The Options charts available freely on the financial services websites, and therefore the ones used on this project, have the following form:

Maturity (Years)	**Strike**	**Market Price**
0.0164	72	12.22
0.0164	72.5	9.4
0.0164	73	10.1
0.1315	79.5	4.55
0.1315	80	4.35
0.1315	80.5	4.17
0.7835	20	63.15
0.7835	23	56.75
0.7835	25	59.09
0.7835	30	52.1
1.7972	67.5	23.26
1.7972	70	20.70
1.7972	72.5	17.45
1.7972	75	17.65

TABLE 3.1: Selected option chains for Facebook, Inc.

The second parameter we would like to calibrate against is the implied volatility. This parameter is a little bit trickier, mainly because this is not data that we can read directly from the options chain. In fact, it is the only piece of information on the Black Scholes

[1]Professor Emanuel Derman (a widely known financial mathematician from Columbia University) has a course on this topic.

equation that we can't get directly from the market. However, we shall explain how to calculate it, using a well known trick.

3.1.1 Implied Volatility

Suppose we know that the price for an option at time t and strike price K is given by $C = C(S, t, K, r, \sigma)$. We can get the option price as the price that the option is being traded on the market. fixing r, t, K, and S as constants, we would get that the predicted price is just a function of the volatility; $C_{BS} = B(\sigma)$, where $B(\sigma)$ is the Black Scholes formula (2.17) taken just as a function of the volatility(Holding all the other parameters as constant). Now consider the function $g(\sigma) = B(\sigma) - C_{Market}$. We want to find the zeros of the function, that is, we want to find a σ for which the Black Scholes equation gives us the same price as the one observed at the market. In order to do so, we apply the Newton-Raphson method, or other root finding algorithm.

$$\sigma_{i+1} = \sigma_i + \frac{B(\sigma_i) - C_{market}}{B'(\sigma_i)}. \tag{3.1}$$

Here, $B'(\sigma)$ is known as "Vega" and it quantifies the change in price due to a change in the volatility. The explicit formula for Vega is given by (Wilmott, 1996) :

$$B'(\sigma) = SN'(d_1)\sqrt{T-t}, \tag{3.2}$$

where $N'(x)$ is the standard normal probability density function. The Newton method converges pretty fast (Sauer, 2011): we will only need around 6 iterations to converge to an acceptable number of decimal places (around 4 or 5 decimal places).

3.1.2 Calibration Methods

Now that we know the parameters against which the model is going to be tested, we will start the discussion of the optimization methods. For this chapter, the optimization algorithms used are the non linear minimum least squares and the constrained optimization. MATLAB includes functions to perform both of these procedures. MathWorks provides more information about these topics, but the algorithms used by MATLAB for this optimization are called "trust-region-reflective" and the "Levenberg-Marquardt" for the non linear least squares and interior-point', "trust-region-reflective" and "active set" for the constrained optimization (MathWorks, 2015). A good reference for optimization numerical methods is given by Papadimitriou and Kenneth Steiglitz (1998). To perform these type of algorithms in MATLAB, we need to create an objective function,

$$f = f(x_1, \ldots, x_n),$$

that requires n inputs, an initial guess for the parameters $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, and at least n simulations to make the system of equations complete.

3.2 Volatility As a Markov Chain

Markov chains are random processes that undergo transitions from one state to another in a state space. The state at time t_n only depends on the state at t_{n-1}. This is a required condition for this type of process and it's called "memoryless". A rather clear explanation (pretty much the standard reference in this topic) is given by Ross (1972).

In this section, we are interested in modeling σ as a two dimensional Markov chain. To do so, suppose that for a given maturity, at strike price K_n, the volatility has a value σ_n that can be larger or smaller than the volatility at the $(n-1)^{th}$ strike price step. We propose the following two state Markov chain: The volatility will increase by a with probability P_{11} given that it increased for the past strike. It will increase by the same value, with probability P_{12} given that it decreased in the past (last step). On the other hand, the probability will decrease with probability P_{21} given that it increased in the past and it will decrease with probability P_{22} given that it decreased in the past. A graphical interpretation of this 2-state Markov chain is given by:

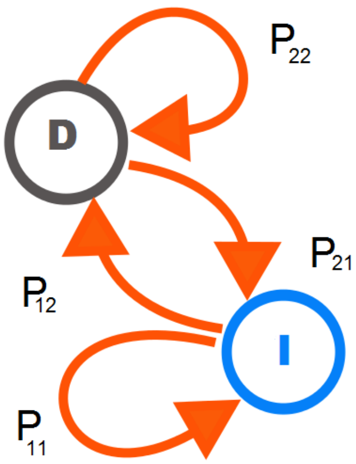

FIGURE 3.2: Graphical description of the Markov Process.

Mathematically, the transition matrix is given by:

$$P = \begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix}. \tag{3.3}$$

So if we know that at a given maturity (that is, at a fixed time), for a strike price k_0 the volatility increased (state 1, x_1) or decreased (state 2, x_2), then, we can find the volatility at the state K_n by

$$x^{(n)} = x^{(n-1)}P = x^{(n-2)}P^2 = \cdots = x^{(0)}P^n, \tag{3.4}$$

where x is a two-dimensional row vector. Then the expected value of the volatility at strike step K_n will be given by

$$(\sigma + a)(x_1^{(n)}) + (\sigma - a)(x_2^{(n)}). \tag{3.5}$$

That is, we are computing the expected value by multiplying increase in volatility, $(\sigma+a)$ times the probability that the volatility increased at state n, $x_1^{(n)}$ and adding it to the contrary case, on which the volatility decreased by a at state $x_2^{(n)}$. The main issue with this approach is that we don't know the probability for the change on each state, so we will have to "guess" an initial value and optimize via non linear least squares or constrained optimization. Performing this approach on the FaceBook, Inc (`FB`) option chart, we get that the optimal parameters are $P_{11} = 0.200$, $P_{21} = 0.3427$, $\sigma = 0.3533$ $a = 0.1000$. The simulation takes 2.44 seconds. After calibrating we simulate the model for different maturities of the stock, which gives us:

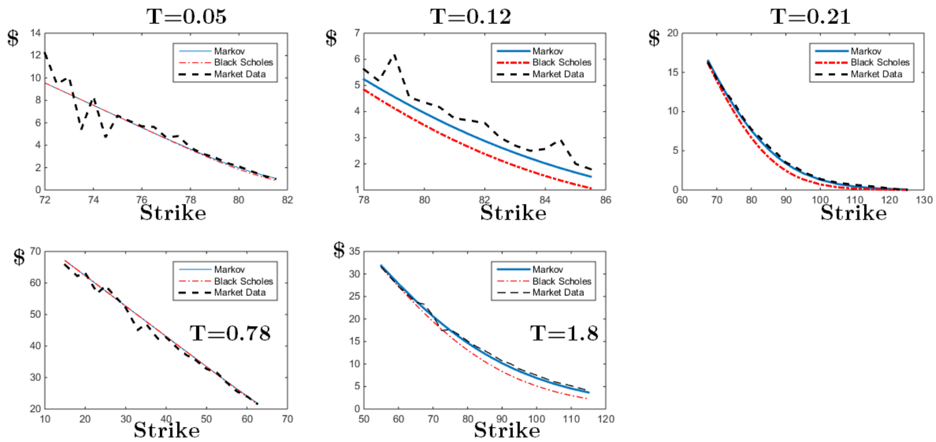

FIGURE 3.3: Comparisons between the market price, the Black-Scholes model and the Black-Scholes model with Markovian volatility for different maturities.

As we can see it gives a much better approximation for all the prices compared to the Black Scholes model. Generally this seems to be a good approach and its study should be taken forward. One drawback of this model, however, is that we need to calibrate $4 \times t$ parameters, where t is the number of maturities. However, if we have enough data and we are not concerned about changes in time on the order of the optimization (around 3 seconds, each) we get a more accurate model. As we can see, there's no such thing as a free lunch.

3.3 Periodic Volatility

Market has cycles. It would come natural to think that the volatility index of an option also does, therefore being an approximately periodic function (of time, of course). In this section, we shall consider the volatility as a periodic function of time. Let's start by analyzing a graph of the volatility index of Google. The plotted time series can be found on Figure 3.4. We can see that the volatility index for Google seems to have a periodicity, which is a fact that we will explore on this work.

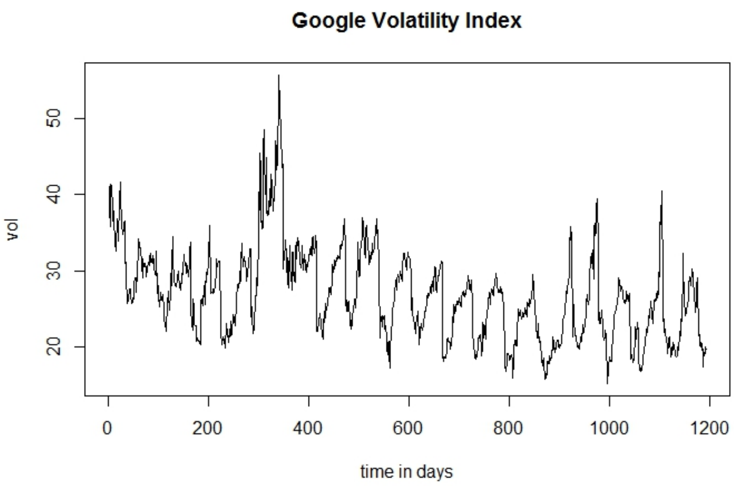

Figure 3.4: GOOG VIX Data from the past 5 years.

In this section we shall focus on modeling this volatility as a function of time.

3.3.1 First Approach: Volatility as a Trigonometric Function

Let's start by considering a model of the form:

$$\sigma(t) = \sigma_0 + A\sin(kt) + B\cos(kt). \tag{3.6}$$

Now, the problem consist on defining σ_0, A, B, and k for the given model. We explore some models and comment on their performance.

3.3.1.1 First Model

From Figure 3.4, we can see that the volatility completes a little bit more than two cycles every two houndred days. As a first approach, we considered modeling the volatility with a periodicity of a quarter year, so I made $k = 8\pi$ in order for the function to repeat itself four times per year. Assuming that the periodicity is given quarterly (every 3 months, 90 days). Let's also assume one of the amplitudes to be 0. Therefore we have: $k = 8\pi, A = 0.01, b = 0$. Then,

$$\sigma_1(t) = 0.18 + 0.01\sin(8\pi t). \tag{3.7}$$

Let's consider this as our first model (hence, the 1 in the subscript of the sigma function). The election of σ_0 is done by choosing the volatility on which the asset was being trade at the moment of this analysis ($\sigma = 0.15$, February, 2015). A is chosen in a way that makes the periodic term always smaller than the initial volatility, σ_0.

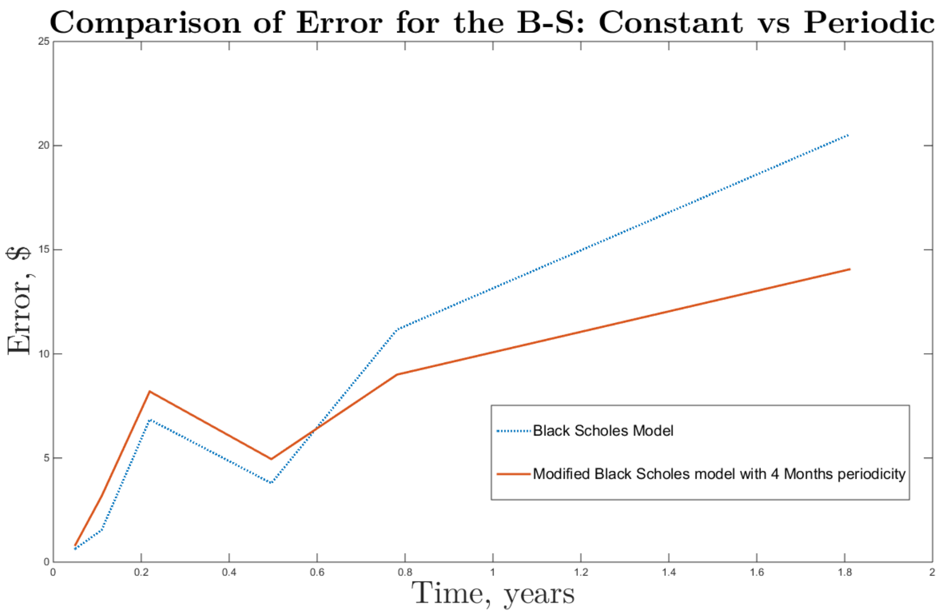

FIGURE 3.5: Periodic Volatility fo k=4π.

As we can see, from Figure 3.5, our first model gives a smaller error for times bigger than 0.6 years (around seven months). Running the same procedure for different A parameters we get that the best parameter compared to the BS price is $A = 0.1$, as we can see on figure 3.6.

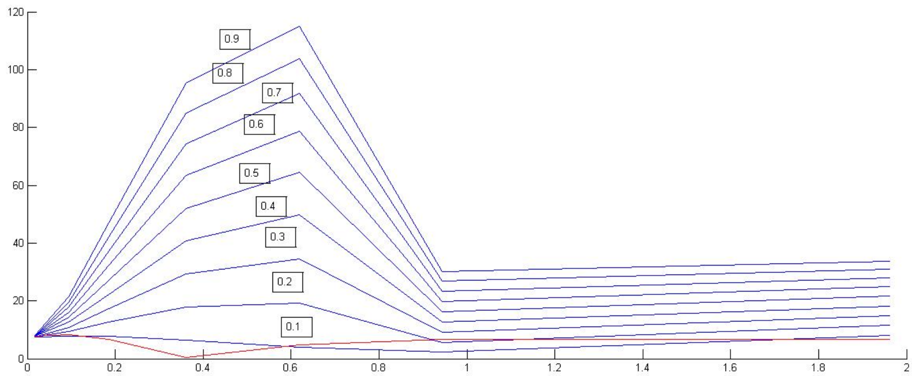

FIGURE 3.6: Mean Error for various A. Red line is the Black Scholes error with constant volatility for comparison.

Note: I did not considered smaller values of A (smaller than 0.1) because then the change in volatility would be very small and wouldn't affect the result on a significant manner. On a more general case for the periodicity of the function, we consider an extended variety of periods:

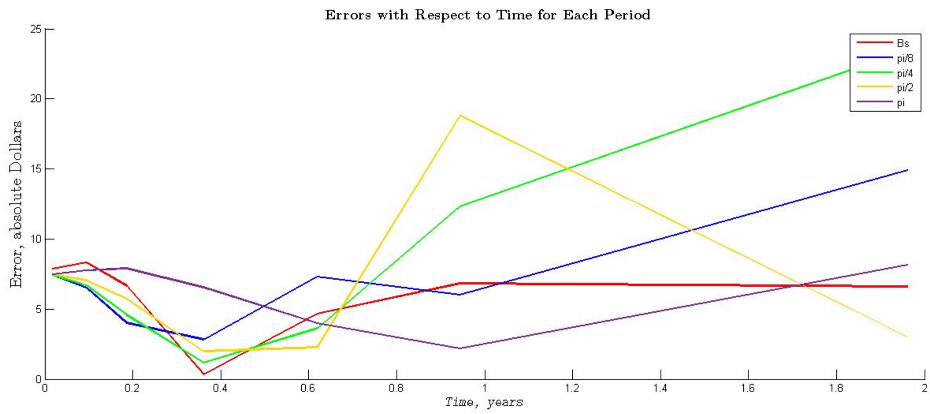

FIGURE 3.7: Mean error for different time periods, $k \leq \pi$.

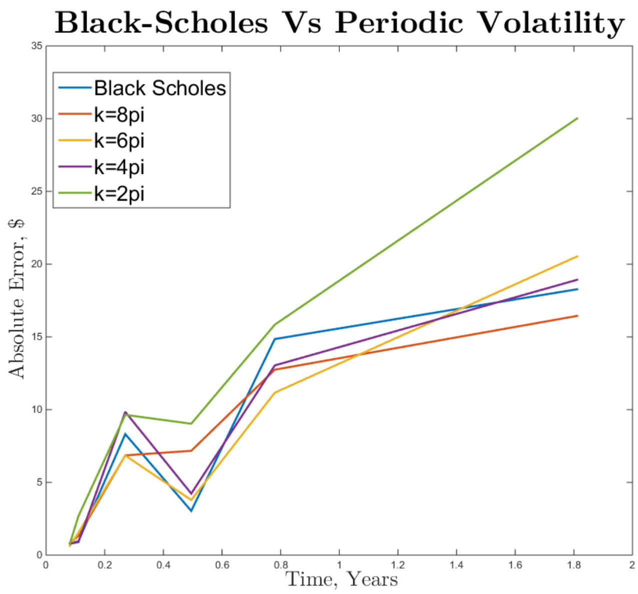

FIGURE 3.8: Mean error for different time periods, $k \geq 2\pi$.

On Figure 3.7 and Figure 3.8, we can see that different values of k give better values for the price depending on the time. For example, for the model with $k = \pi$, we get that the error plot yields:

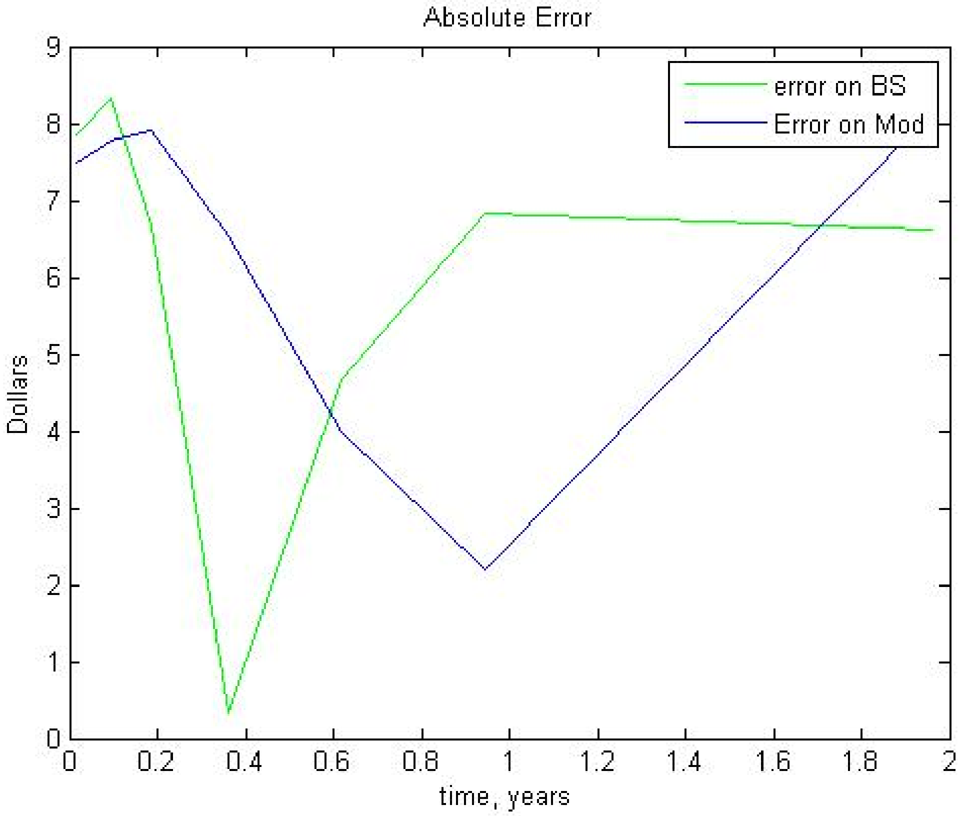

FIGURE 3.9: Error plot when $k = \pi$.

We can see how this last model gives better results than the Black-Scholes equation between 7 months and around 19 months.

3.3.2 Second Approach: Modeling Volatility With a Fourier Series

From the last section, we can see that modeling the volatility as a trigonometric function yields some success for some cases, (for some time intervals). Now, we consider the use of a more complex model, the use of a Fourier series, that is, modeling the volatility as

$$\sigma_2(t) = \sum_{i=1}^{n} a_0 + a_i \cos(iwt) + b_i \sin(iwt). \tag{3.8}$$

It is important to note that for this fitting I will be using the time in days. Fitting different Fourier series, we concluded that the best fit is given by the Fourier series representation with $n = 3$. This is done to avoid over-fitting. The model chosen is given by:

$$\sigma_2(t) = 0.2 + 0.2275\cos(wt) - 0.1969\sin(wt) - 0.1152\cos(2wt) - 0.2144\sin(2wt) \\ -0.1279\cos(3wt) - 0.5.625\sin(3wt), \quad (3.9)$$

where we have that $w = 0.003528$ (again, time is in days!). The fit gives the following goodness of fit table as a result:

SSE	5.571e+04
R-square	0.6087
Adjusted R-square	0.6064
RMSE	6.811

and yields the plot presented on Figure 3.10:

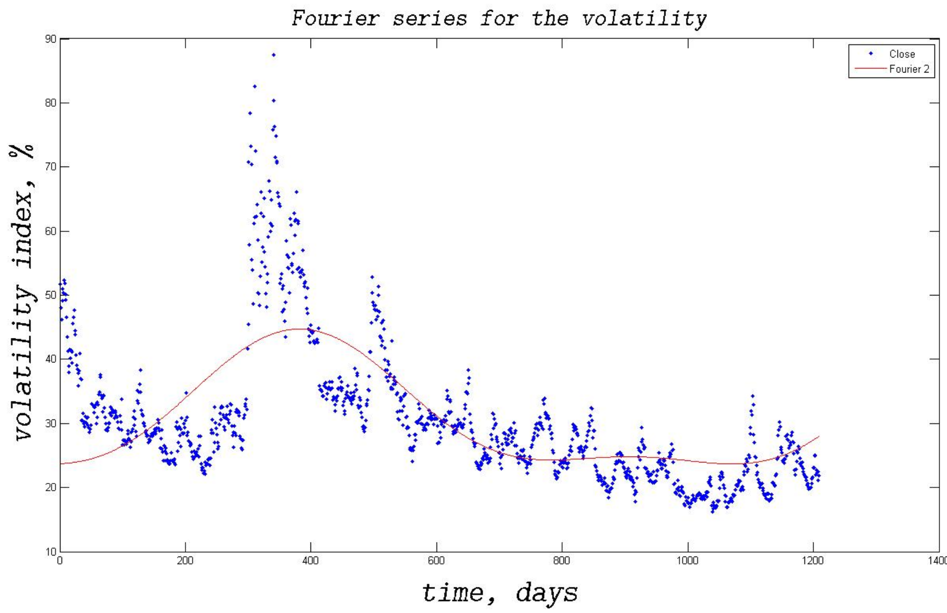

FIGURE 3.10: Fourier Approximation of the Volatility, Google Inc.

Applying this model to the Google option data, we get that this modification gives a much better pricing than the standard Black-Scholes model:

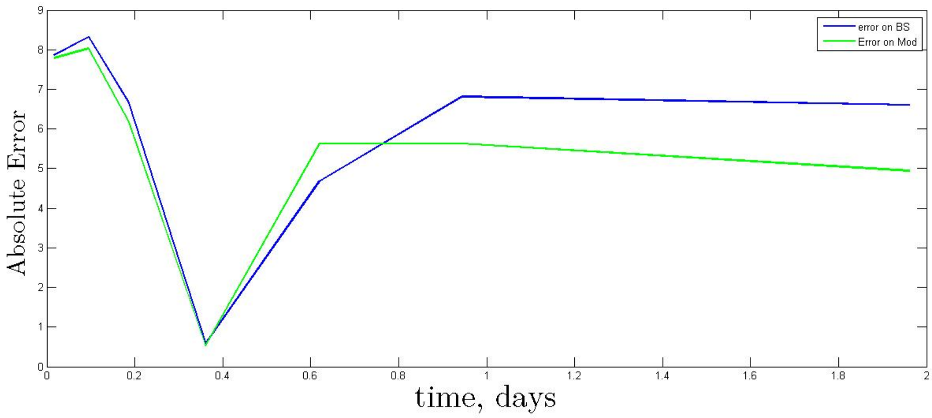

FIGURE 3.11: Error plot for the Fourier series model.

As we can see, this model gives a smaller error compared to the standard model, except from values between 3 and 5 months. We can also see that the error is smaller when longer times are considered, therefore, this provides a better model to predict prices in the more distant future (more than a year).

3.3.3 Third Model: Periodic with a linear trend.

Now, we propose to model the volatility as a periodic function with a trend, that is, modeling the volatility as:

$$\sigma(t) = \mu + \beta t + g(t), \tag{3.10}$$

where $g(t)$ is a periodic function of time, as used before. From section (3.2) we get that one of the best fitting model for the periodic function is given by $\sigma(t) = 0.15 + 0.01sin(8\pi t)$. Now, we just need to fit a linear trend to the volatility. Doing so we get that the fit is given by:

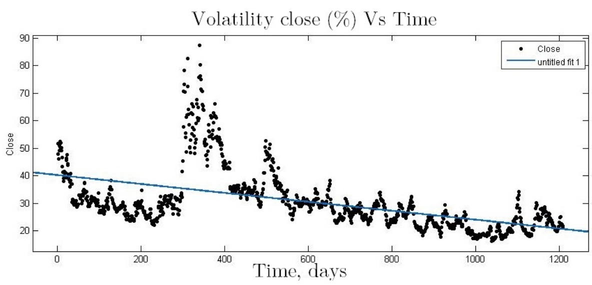

FIGURE 3.12: Linear fit for the data.

and has parameters β= -0.056 and μ =0.3921. Then, we proceed to simulate the model, which gives us as a return:

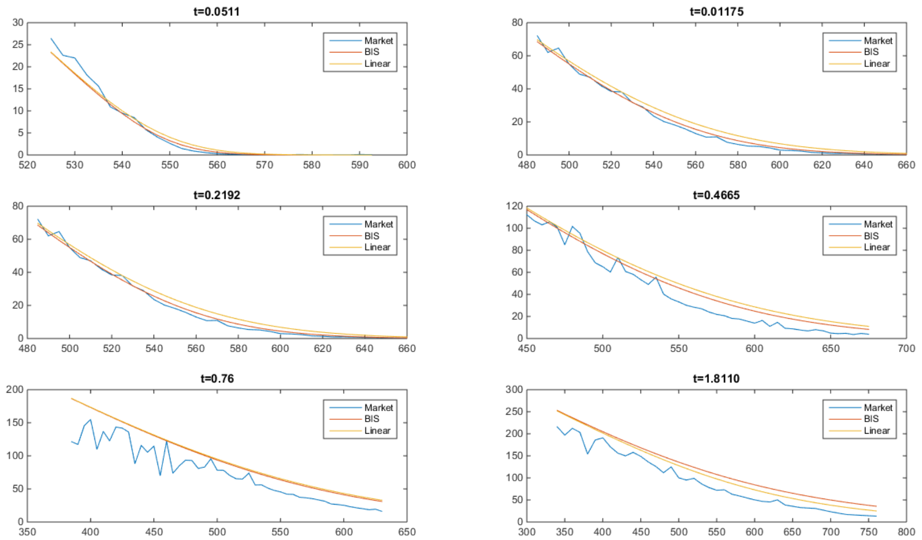

FIGURE 3.13: comparison of prices for the linear model.

In general, this method tends to overprice the data. However, this is not necessarily a bad thing. This might be something desirable because overestimation is somewhat safer; it makes more sense to make financial plans and take financial decisions by overestimating rather than by underestimating.

3.4 Random Models: An Empirical Approach

In this section we will consider the random nature of the financial market and make use of this to model the volatility. Under some circumstances, volatility can become uncertain, specially when modeling complex and exotic derivatives (Avellaneda & Paras, 1998. Hull, 2003). We shall then propose and discuss a model that captures this essence.

3.4.1 Volatility from a distribution

This model assumes a total randomness of the volatility. In practice, this is not completely true, for the volatility will have some dependencies. For example, the implied volatility tends to decrease with time (Guernsey, 2014). However, just for the sake of the model, let's assume that the volatility is a random number generated by a distribution of the form

$$\sigma_r \sim f(\theta_1, \theta_2, \ldots, \theta_n), \tag{3.11}$$

where $(\theta_1, \theta_2, \ldots, \theta_n) = \vec{\theta}$ is the vector of parameters of this distribution; mean, variance, shape parameter, etc. Having proposed the model, we are just left with one question; what distribution (i.e., Gaussian, Poisson, Weibull, etc) should we use to model the volatility?

3.4.1.1 Obtaining the data

We used the option data for the Google Inc. option chain (`GOOG`) taken on march 31, 2015, containing five different maturities for different strike prices, making up for a total of 292 options. Then, we proceeded to compute the implied volatility of this data using the Newton-Raphson method. Once we found the 292 implied volatilities for the data, we proceeded to find the distribution with the best fit. Using MATLAB and the Bayesian information Criterion (BIC), we got that the implied volatilities are distributed as shown in the graph bellow. To find this, we used a MATLAB code that iterated over all the predetermined distributions and assigned a BIC (Bayesian information criterion) to them. We can see that the distribution is given by:

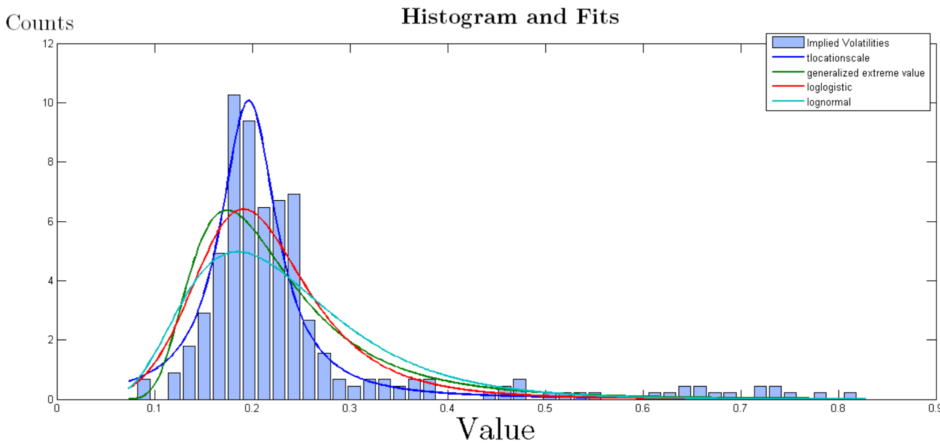

FIGURE 3.14: Distribution of the implied volatility of GOOG. Distributions are listed according to their BIC.

And the volatility surface is given by:

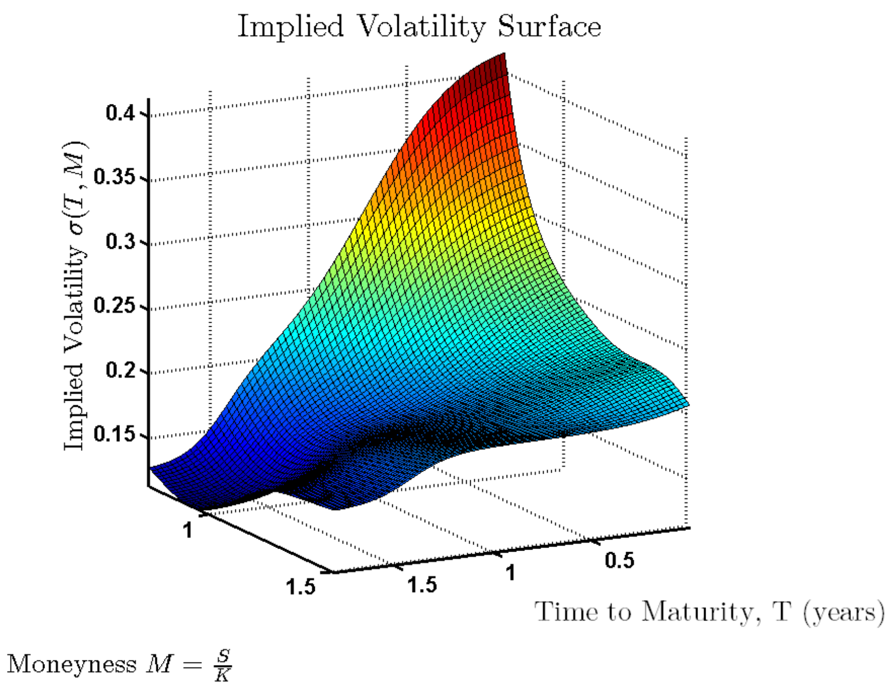

FIGURE 3.15: Volatility Surface.

From the plot we can see that the distribution that yields the best fit is the Student's-t location scale with mean μ, scale parameter ν and standard deviation σ, given by $V \sim t(\mu, \sigma, \nu)$, where V represents the volatility as a random variable. The PDF of said distribution is given by

$$t(\mu, \sigma, \nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sigma\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)} \left[\frac{\nu + \left(\frac{x-\mu}{\sigma}\right)^2}{\nu}\right]^{\left(\frac{\nu+1}{2}\right)}. \tag{3.12}$$

The t location-scale distribution is useful for modeling data distributions with heavier tails (more prone to outliers) than the normal distribution (MathWorks, 2015). The estimated parameters for this distribution are

Parameter	Value
μ	0.2345
σ	0.0347
ν	0.00854
BIC	-1044.9031

3.4.1.2 Confining the volatility to σ^- and σ^+

Finally, let's make a small refinement to this model. The following line of thought follows an approach somewhat similar to that of Avellaneda and Paras (1998), Lorenz and Qiu(2009), and Guernsey (2014), in the sense that we have an unknown volatility that takes values between a minimum volatility σ^-, and a maximum volatility σ^+. However, instead of modeling it as a function of $\frac{\partial^2 V}{\partial S^2}$, we model it a compete random number that takes a value between σ^-, and σ^+ for each time, t and each strike price, K.

To perform the simulation, we use the implied volatility calculated before. Then, we proceed to remove the values that are outside σ^- and σ^+. In order to choose our limits, we make use of the `VIX` index[2]. Doing so, we get that the minimum value of the volatility for last year was 0.115 and the maximum was 0.2911. We shall use this values as our volatility interval.

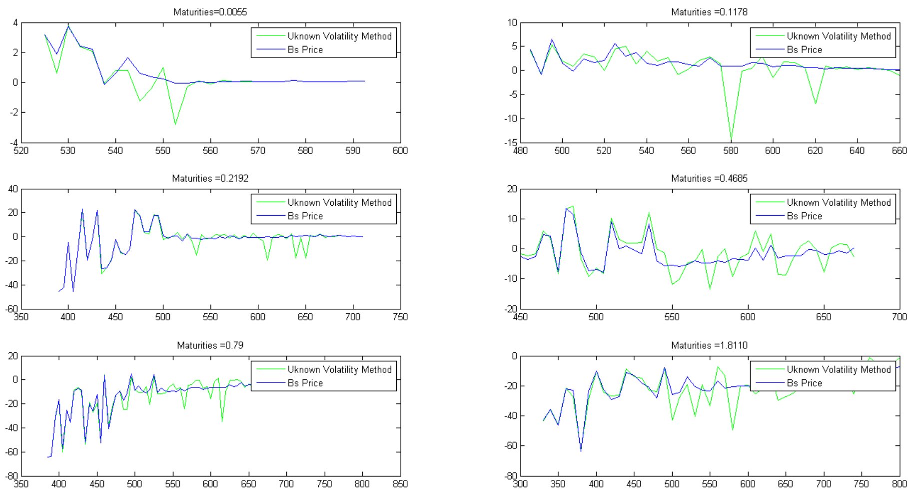

FIGURE 3.16: Error plots for the Unknown Volatility model and the Black Scholes Model.

We can see that, for certain strike prices, we find that the errors in the model are sometimes bigger and sometimes smaller than the errors from the Black-Scholes Model. However, considering that this is a method to price financial derivatives whose volatility is unknown, we could say that this model is, in overall terms, satisfactory. This model seems to behave better for in the money options, that is, options whose strike price K is smaller than the underlying price, S ($548 on this particular case).

[2]Supplies the market expectations of near-term volatility conveyed by S&P 500 stock index option prices. http://www.cboe.com/micro/charts/vix.aspx.

3.5 Using a confidence interval

Finally, I propose the use of a confidence interval for the volatility. We start by sampling some implied volatilities from a training data set. Assuming that the sampled volatilities follow a certain distribution, we proceed to create a $(1-\alpha)$ confidence interval, a procedure fairly standard in statistics. We would have three volatilities, $\bar{\sigma}$, $\bar{\sigma}_L$ $\bar{\sigma}_U$, that would give us three predicted prices, $\hat{C}$, $\hat{C_U}$, and $\hat{C_L}$. For this case, I shall also use the Google, Inc. option chain. As usual, I will compare with the market price and with the price predicted by the Black-Scholes equation. Taking a random sample of implied volatilities we get that they behave as a log-normal distribution, which agrees with well known empirical results:

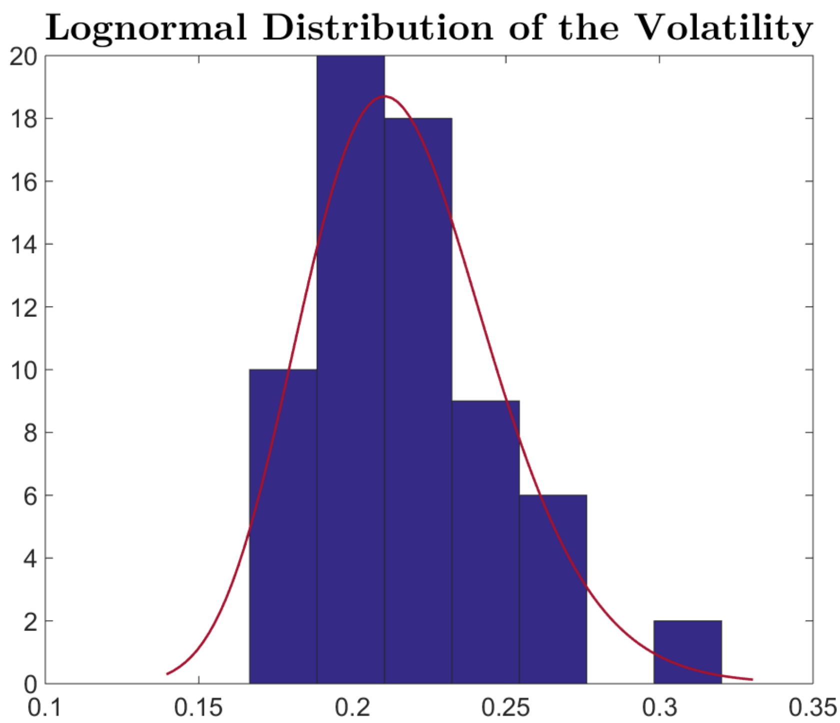

FIGURE 3.17: log normal distribution of the volatility.

Using a 95% confidence interval, we get that the mean, the lower limit and the upper limits are given by:

Lognormal distribution			
Parameter	mean	Lowe Limit	Upper Limit
$log(\mu)$	-1.5379	-1.57336,	-1.50243
μ	0.2148	0.2073	0.2226
$\sigma = 0.0318$			
$p - value = 0.5763$. Can't reject H_0: Distribution is Log-Normal			

TABLE 3.2: Summaries for the distribution.

As we can see from the visual parameter and the p-value, said distribution seems to fit the data very well. Therefore, I shall construct a 95% price interval using both volatilities. The result, tested on the Google Inc option data is given as follows:

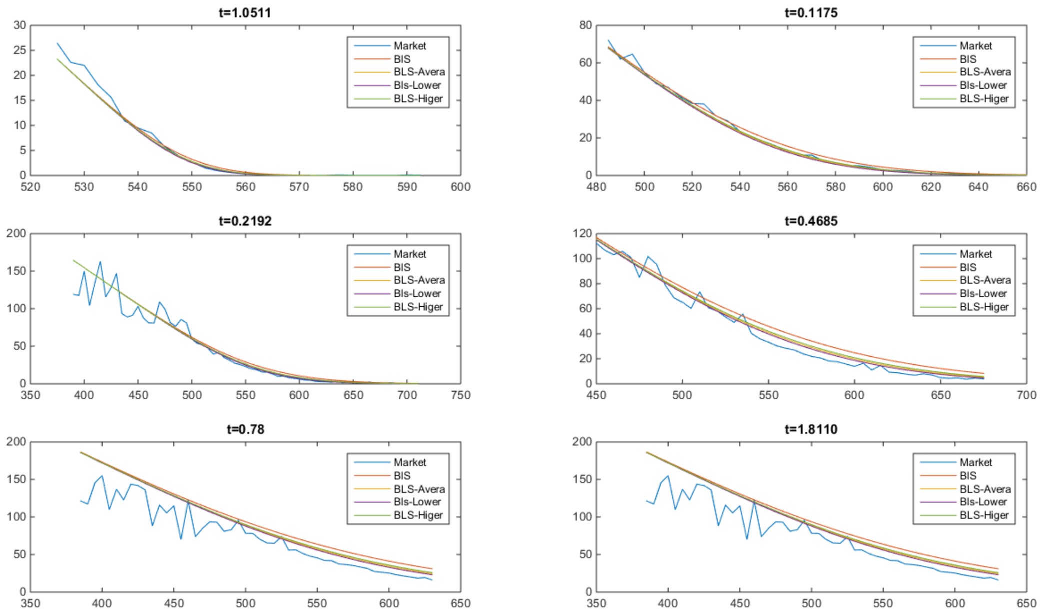

FIGURE 3.18: Comparison for the confidence interval method.

As we can see, this method provides more accurate results compared to the original Black-Scholes Model. One of the main advantages of this model is the fact that it doesn't require any complicated computational models, it rather based on a simple idea from statistics, and it seems to work well.

3.6 Summary

We have seen then how some model behave better than other under certain circumstances. it is interesting to propose and model the volatility because, even though it is the most important parameter of the Black Scholes model, it is not something that is observable from the market; it something that can only be estimated. Having a more accurate estimation of said parameter is of crucial importance to minimize risk and optimize portfolios, which is translated into a more secure and overall better practice of the option pricing.

These models are subjected to further evaluation. However, they yield quite interesting results so far. Due to the time constraints and the extension of this type of work we couldn't explore other methods, such as the parametrization over both the strike price and the time. Such parametrization are rather scarce. Gatheral has studied this topic and has presented some methods in his papers (2004, 2011, 2012).

Note:

In the following sections I will discuss two practical approaches to the option pricing theory. The first one is called the Heston option pricing model, which models the volatility as a stochastic process. The other approach comes from the time series analysis, where I will model the log-returns of a stock price (Yahoo, Inc.) as a GARCH model. I will use the Black-Scholes equation as a benchmark for these methods.

Chapter 4

A Practical Approach to the Heston Model

4.1 Introduction

Proposed by Heston (1993) while he was a professor at Yale, this model offers, under certain conditions, some advantages (more accurate prices) compared to the Black Scholes equation, which is a particular case of this model (Heston, 1993. Moodley, 2006). However, as we will discuss later, this model requires certain parametrization, which is not trivial -or unique for that matter-. Also, one of the drawbacks of this model is that it seems to be outperformed by the Black Scholes model for short maturities. Some of the derivations of this model require some notions of stochastic calculus and measure theory. Good references for these topics are Mikosh (1999) for the earlier, and Stein & Shackrachi (2005) for the latter.

The Heston Stochastic Volatility Model is specified by the following equation:

$$\frac{dS(t)}{S(t)} = \mu dt + \sqrt{V(t)} dW_1, \tag{4.1}$$

$$dV(t) = \kappa(\theta - V(t))dt + \sigma\sqrt{(V(t))} dW_2. \tag{4.2}$$

In order to take the leverage effect into account, the following condition shall be met:

$$dW_1 dW_2 = \rho dt. \tag{4.3}$$

That is, the two Wiener processes are correlated.

As we can see, the model depends on some parameters; μ, the drift coefficient, κ which is the mean reversion rate[1], θ, the long run variance, V_0 the initial variance, the volatility of the volatility, σ, and finally, the correlation between both Wiener processes, ρ. This last coefficient describes the correlation between the logarithm of the price and the volatility of the stock. According to Heston (1993), σ, the volatility of the volatility, affects the kurtosis of the probability density distribution. If $\rho > 0$, then the volatility will increase as the stock price increases. On the other hand, if $\rho < 0$, then the volatility will increase while the stock price decreases. This model is subject to some constrains; κ, V_0, and θ should be bigger than 0. Also, To make the variance always positive, we are subjected to:

$$2\kappa\theta > \sigma^2. \tag{4.4}$$

The only parameter that can be less than zero is ρ, which takes values between $[-1, 1]$. These constraints will be of fundamental interest at the time to calibrate the model.

I will follow a risk neutral approach to derive the formula, as shown by Moodley(2006). However, there are other derivations that yield to the same result. For example, Heston(1993) derives it a tad differently, using stochastic volatility arguments and a "Black-Scholes Like" PDE. In an equivalent martingale measure, $\mathbb{Q}$, we get that the option price is given by

$$C = E_t^{\mathbb{Q}}[e^{r(T-t)}H(T)], \tag{4.5}$$

where $H(T)$ is the payoff function at T and r is the risk free interest rate (this value can be determined from the market data. In practice, they use the rate of return of the US treasury bond on 30 years). Using Girasov's theorem, which describes how the dynamics of an stochastic process changes when the measure is changed to an equivalent martingale measure (Kallianpur and Karandikar, 2000), we get that

$$d\tilde{W_t^1} = dW_t^1 + \nu_t dt, \tag{4.6}$$

$$d\tilde{W_t^2} = dW_t^2 + \Lambda(S, V, t)tdt. \tag{4.7}$$

Using the Radon-Nykodim theorem:

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left[-\frac{1}{2}\int_0^t (\nu^2 + \Lambda(S, V, t)^2)ds - \int_0^t \nu_s dW_s^1 - \int_0^t \Lambda(S, V, t)dW_t^2\right], \tag{4.8}$$

where $\nu = \frac{\mu - r}{\sqrt{V_t}}$, $\mathbb{P}$ is the physical world measure[2], and $\tilde{W_t^i}$ (for $i = 1, 2$) are Brownian motions under the equivalent martingale measure. Under this measure, equations (4.1),

[1] that is, assumption that a stock's price will tend to move to the average price over time.

[2] The physical world measure, $\mathbb{P}$ uses real world probabilities, while the equivalent martingale measure, $\mathbb{Q}$, uses a risk free approach. Hatfield (2012) provides a more in depth discussion of this issue.

(4.2) and (4.3) become:

$$dS_t = rS_t dt + \sqrt{V_t} S_t d\tilde{W}_t^1, \tag{4.9}$$

$$dV_t = \kappa^*(\theta^* - V_t)dt + \sigma\sqrt{V_t} d\tilde{W}_t^2, \tag{4.10}$$

and

$$d\tilde{W}_t^1 d\tilde{W}_t^2 = \rho dt, \tag{4.11}$$

respectively. The following changes were made: $\kappa^* = \kappa + \lambda$, and $\theta^* = \kappa\theta/(\kappa + \lambda)$, which results in an elimination of λ under the equivalent martingale measure. The election of λ is given by Heston in his original paper and it is related to Breeden's (1979) consumption based model.

Heston presents that, by analogy to the Black-Scholes model, the solution of the pricing equation is of the form:

$$C(S_t, V_t, t) = S_t P_1 - K e^{r(T-t)P_2}, \tag{4.12}$$

where, for $j = 1, 2$, we have that:

$$P_j(x, V, T, K) = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty Re\left(\frac{e^{\phi \ln K} f_j(x, V_T, T, \phi)}{i\phi}\right) d\phi, \tag{4.13}$$

$$x = \ln(S_t), \tag{4.14}$$

$$f_j(x, V_t, T, \phi) = \exp\left[C(T-t, \phi) + D(T-t, \phi)V_t + i\phi x\right], \tag{4.15}$$

$$C(T-t, \phi) = r\phi i r + \frac{a}{\sigma^2}\left[(b_j - \rho\sigma\phi i + d)\tau - 2\ln\left(\frac{1 - ge^{dr}}{1 - g}\right)\right], \tag{4.16}$$

$$D(T-t, \phi) = \frac{b_j - \rho\sigma\phi i + d}{\sigma^2}\left(\frac{1 - e^{dr}}{1 - ge^{dr}}\right), \tag{4.17}$$

$$g = \frac{b_j - \rho\sigma\phi i + d}{b_j - \rho\sigma\phi i - d}, \tag{4.18}$$

$$d = \sqrt{(\rho\sigma\phi i - b_j)^2 - \sigma^2(2y_j\phi i - \phi^2)}, \tag{4.19}$$

where, $u_1 = 1/2 \quad u_2 = -1/2, \quad a = \kappa\theta, \quad b_1 = \kappa + \lambda - \rho\sigma, \quad b_2 = \kappa + \lambda.$
The above integral seems monstrous, but in fact, it is really easy to implement in MATLAB. As we can see, the equation (4.13) represents the integral to be calculated, and the equations (4.14) to (4.19) describe the parameters of the aforementioned integral. Therefore we can solve it by defining each part of the integral and by making the upper limit of the integral a rather large number.

4.2 Implementation

Now, let's show how we can change from the theory to the practice, that is, let's examine the behavior of the Heston model and compare its results for different parameters.

4.2.1 Integration Schemes

Following a similar approach to that of Nimalin (2006) and Mikhailov & Nogel (2003), we get that, so far, the model has been reduced to two problems: the first one, finding the appropriate parameters for the function, which I shall do shortly, and the second one: Integrating equation (4.13). There are different formulations in order to do the latter, for example, Carr (1998) formulated a solution using the fast Fourier transformation (FFT) algorithm, which we shall use shortly. The Heston model uses a characteristic function of the form:

$$f(\phi_j) = Re\left[\frac{e^{-i\phi_j \ln K} f_k\left(\phi_j; x, \nu\right)}{i\phi_j}\right], \tag{4.20}$$

where $j = 1, 2$.

In this section I will discuss the use of a Gauss-Legendre and the FFT method proposed by Carr(1998), for which the characteristic function of the model is given by:

$$f(\phi_j) = Re\left[e^{-i\phi_j \ln K} \psi(\phi_j)\right], \tag{4.21}$$

where we have that:

$$\psi(\phi_j) = \frac{e^{-rT} f_2(\phi_j - (\alpha+1)i; x, \nu)}{\alpha^2 + \alpha - \phi_j^2 + i(2\alpha+1)\phi_j}. \tag{4.22}$$

From the above equation we get that f_2 is the characteristic function of the Heston Model, and α is a necessary parameter, as stated by Gil and Pelaez (1958).

4.2.1.1 Gaussian Quadratures

The Gaussian quadratures are an approximation of the definite integral of a function, usually stated as a weighted sum of function values at specified points within the domain of integration. Generally, Gaussian quadratures take the following form:

$$\int_b^a f(x)dx \approx \sum_{k=1}^{N} w_k f(x_k), \tag{4.23}$$

where each w_k is a weight at the k^{th} subinterval. For each method, the weight will be a polynomial given by the method itself. For example, if we use the Gauss Legendre Quadrature, we obtain that the weights are the roots of the Legendre polynomials given by:

$$w_i = \frac{2}{(1-x^2)[P_n'(x_i)]^2}, \tag{4.24}$$

where P_n Represents the Legendre polynomial. I won't go into much detail of this method mainly because it is widely known and studied. However, the reader that would like to learn more about how to perform integrals under these schemes might refer to Sauer (2012) or Burden (2010).

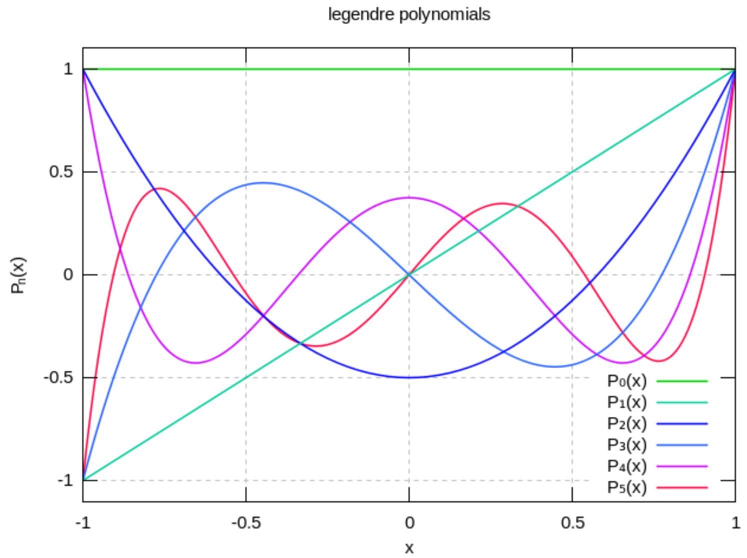

FIGURE 4.1: Five first Legendre Polynomials.

Note: There are other widely known integration schemes that can be used to price the function. Said methods include midpoint. trapezoid, Simpson's and Simpson's 3/8 rule. However, I shall limit my approach to a Gauss-Legendre quadrature.

4.2.1.2 FFT

This method was first introduced by Carr and Madan in 1999 so they could increase the computation speed for the option pricing. Basically, what this algorithm does, is mapping the vector $\mathbf{x} = (x_1, \ldots x_n)$ into $\hat{\mathbf{x}} = (\hat{x}_1, \ldots \hat{x}_n)$ with the function:

$$\hat{x}_k = \sum_{j=1}^{N} x_j e^{\left(-i\frac{2\pi}{N}(j-1)(k-1)\right)}. \tag{4.25}$$

Normally, this algorithm would behave as $O(h^2)$, however, since each sum is computed simultaneously, it takes $N \log_2(N)$ steps (Burrus, 2008, Sawer, 2012). The most commonly used algorithm to calculate the FFT is a *divide an conquer* type of algorithm called "the Cooley-Tukey algorithm"(Barnett, 2008). According to Carr and Madan (1999), the price for a call option at the log of the strike, $\ln K$, is given by

$$C(K) = \frac{e^{-\alpha K}}{\pi} \int_0^\infty Re\left[e^{i\nu K}\psi(\nu)\right] d\nu. \tag{4.26}$$

To implement this method, we divide ν and $\ln K$ into N equally spaced subintervals of length λ

$$x_j = e^{i(b-\ln S_t)\nu_j}\psi(\nu_j)w_j, \tag{4.27}$$

where $b = N\lambda/2$ and $\psi(\nu_j)$ is the equation (4.7) evaluated at the j^{th} subinterval. Consequently, let

$$\psi(\nu_j) = \frac{e^{-rt} f_2(\nu_j - (\alpha+1)i)}{\alpha^2 + \alpha - \nu_j^2 + iv(2\alpha+1)}, \tag{4.28}$$

where $f_2(\nu_j - (\alpha+1)i)$ is the Heston characteristic function evaluated at $(\alpha+1)i$. By doing this, we will be able to discretize (4.26) in a form similar to (4.27). Clearly, we would also need to discretize the integration domain. After said parametrization is made, we could use different integration schemes, but I will limit this paper to the use of the trapezoidal rule. Using said scheme over $[0, b]$ for ν using N equidistant points, we get that

$$\nu_j = (j-1)\eta, \qquad for\ j = 1, \ldots N. \tag{4.29}$$

Clearly, η will be the increment between ν_j and ν_{j+1}. Using the trapezoidal rule, the option price (4.26) becomes

$$C(k) \approx \frac{e^{-\alpha k}}{\pi} Re\left[\frac{1}{2}e^{-i\nu_1 k}\psi(\nu_1) + e^{-i\nu_2 k}\psi(\nu_2) + \cdots + \frac{1}{2}e^{-i\nu_N k}\psi(\nu_N)\right] \tag{4.30}$$

$$= \frac{\eta e^{-\alpha k_u}}{\pi} \sum_{j=1}^{N} Re\left[e^{-iv_j k}\psi(\nu_j)\right] w_j, \tag{4.31}$$

where the weights are represented by w_j and its value are given by $w_j = 1$ for $j = 2, \ldots, N-1$ and by $w_j = 1/2$ for $j = 1,\ j = N$. Thanks to this discretization we can evaluate the characteristic function.

4.2.2 Calibration Methods and Empirical Tricks

In this section I will follow an approach similar to that of Nogel (2003), Moodley (2005), and Rouah (2013) to calibrate the model in different ways. I shall use an objective function in order to do this. MATLAB handles these type of functions using its non linear least squares (`lsqnonlin`) and constrained optimization (`fmincon`) algorithms. I shall also discuss some "tricks" to speed up the calibration. To use said algorithms, I need to create an objective function that is concatenated within the algorithms to price the option via the Heston model. The MATLAB function will then find the values (κ, θ, γ, V_0, and ρ) that minimize the difference between the predicted prices from the objective function and the market prices. An example of said function is given by

```
% Objective Function for Non-Linear least squares
function er=HestonError(input)
load google.txt
%creates variables
k=google(:,2);
t=google(:,1);
marketprice=google(:,3);
%market proice is a matrix,
%prealocates
modelprice=zeros(length(google),1);
er=zeros(length(google),1);
%creates the error matrix
%input matrix = [kappa theta sigma rho v0]

for i=1:length(google)
modelprice(i)=HestonCall(input(1),input(2),input(3)...
input(4),input(5),r,t(i),S,k(i)); %Uses the Heston Gauss Legendre
er(i)=(marketprice(i)-modelprice(i))^2; %parameter to be minimized
end
```

I will use the Google option data in order to use test this model. After calibrating, we get that the parameters are given by:

Using Gauss-Legendre			
Method:	`lsqnonlin`	Method:	`fmincon`
Parameter	Value	Parameter	Value
V_0	0.0829	V_0	0.05601
θ	0.1912	θ	0.1247
σ	1.9998	σ	2.1241
ρ	-0.7277	ρ	-0.7521
κ	1.4788	κ	1.578
Time	382.51 Sec	Time	471.98 Sec

Table 4.1: Parameters for the Heston Model

As we can see, both methods (using a Gauss Legendre integration scheme) take quite some time to calibrate the model. However, we can see that the `lsqnonlin` function from MATLAB seems to be a little bit faster.

As mentioned before, there are other methods to estimate said parameters. Moodley proposes the use of an adaptive simulated annealing algorithm, however, he shows that the time to find the parameters is in the order of hours, so I believe we are better off with these methods. There are other methods on the literature, that include the implementation of a trick to compute the calibration faster. Said trick is proposed by Kilin (2007), that argues that the characteristic function, (4.21) doesn't change much with respect to the strike, as it does with respect to the time. Therefore, he proposes then that the characteristic function should only depend on the maturity, and therefore, we can calculate f_1 and f_2 during the optimization and preallocate them (cache them) for every different maturity. This improves the computation time tremendously! Testing and comparing the models on the Google data, we get that:

Objective Function	κ	θ	σ	V_0	ρ	time	MSE
Gauss-Legendre without trick	1.4788	0.1912	1.9998	0.0892	-0.7277	382.51	4.59e-03
Gauss-Legendre with trick	1.4780	0.1913	1.9998	0.0892	-0.7277	7.2243	4.59e-03
FFT	2.4318	0.1441	2.0000	0.0925	-0.8283	1.2204	4.78e-03

Table 4.2: Parameters for the Heston Model for different calibrations

We can see that the fastest one is the one that uses the FFT, and thus, this is the method we shall use. Each of the implementations simulates the Heston model with a given volatility. For the experiments done above, the implied volatilities are:

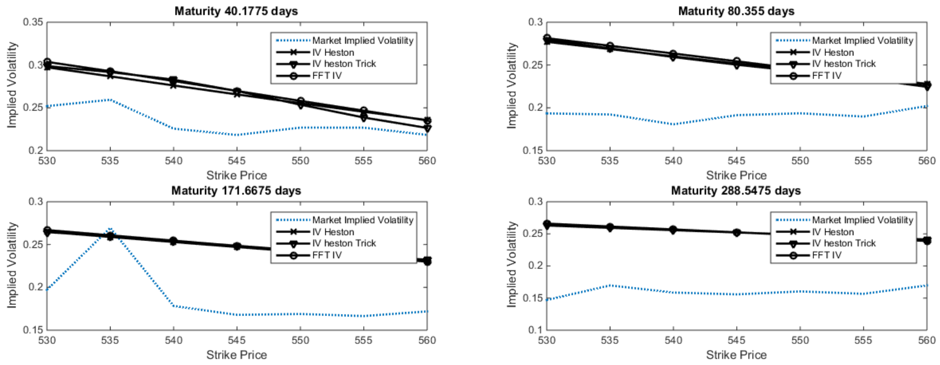

FIGURE 4.2: Volatilities Produced by the Methods

As we can see, all the simulations share an implied volatility similar to each other, however, they are somewhat different (although not by much) than the volatility implied by the market.

It is important to mention that there are other "tricks" when it comes to calibrating the model. For example, it has been shown that the calibration is not very sensible to the parameter κ (Rouah, 2013), and therefore, it would be "ok" to use a conservative and educated guess.

4.2.3 Simulation

Simulating the Heston model for the Google, Inc. data using least squares estimation and the FFT to compute the integral, we compare the predicted price by this model with the predicted price by the Black-Scholes model. The error of the model with respect to the prices as well as the error of the Black-Scholes model are shown bellow. The parameters are the same as the ones in Table 4.2, last row.

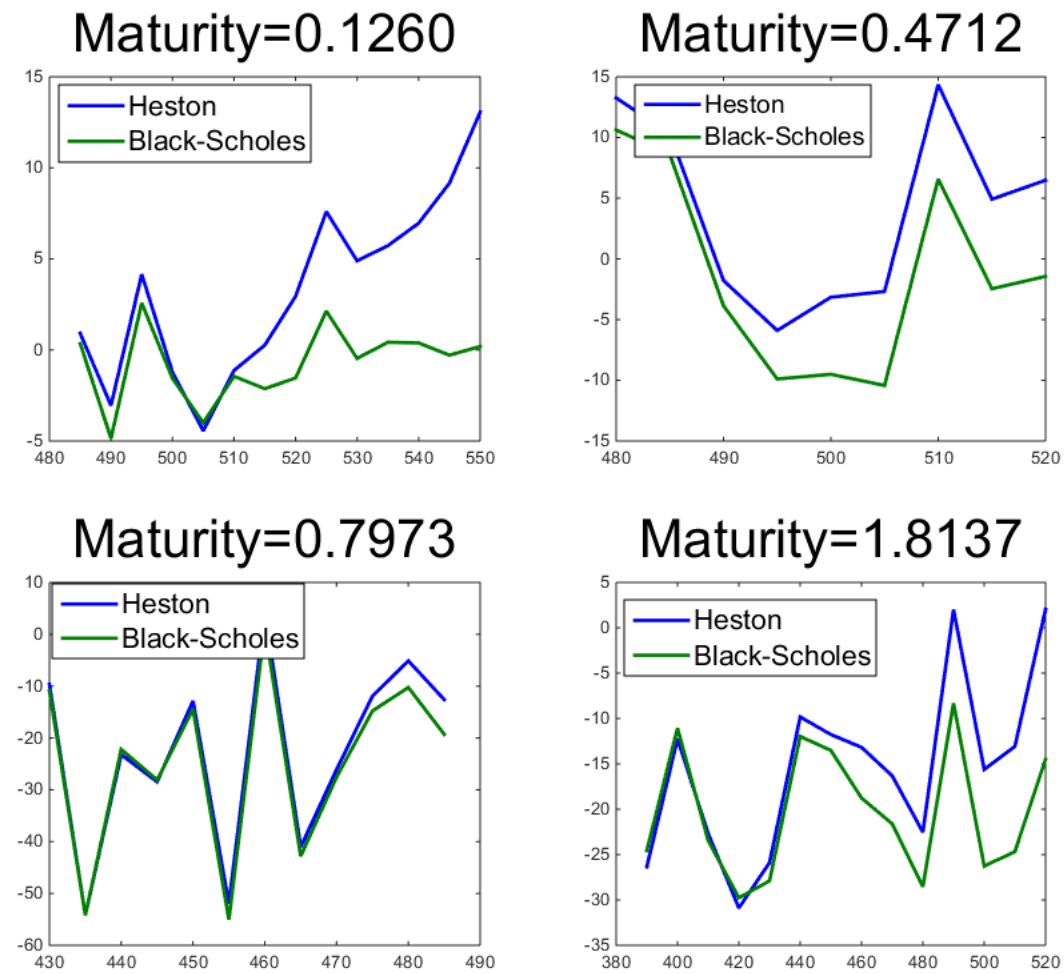

FIGURE 4.3: Difference in prices with respect to the market, Heston and Black-Scholes models.

4.3 Discussion

As we can see, the Heston model gives, in general, more accurate results than the Black Scholes formulation. However, we can see that for smaller maturities this method has a bigger error. It is important then at the time of pricing option to follow this rule; if we are pricing options with smaller maturities, we might be better off using the Black-Scholes model or a modification of it. However, if we want to price options with higher maturities, we are better off using the Heston Model (Rouah, 2015). We can also see that using the FFT reduces the calibration time tremendously.

The Heston model is quite good, assuming that we are not pricing options that are close to maturity. Other than that, it seems to behave better than the Black-Scholes model. Lastly, one of the drawbacks of this model is that it needs to be calibrated with frequency (around twice a week), because the market prices that were used to calibrate the model will obviously change over time.

Chapter 5

A Study of GARCH Models

Finally, we consider the implementation of a method that comes from the time series analysis: The General Autoregressive Conditional Heteroskedasticity model (GARCH). These methods consider the returns of a stock to behave as a GARCH process and are able to predict the prices based on this and the expected value under an EMM. Let's start the discussion defining a GARCH model.

5.1 GARCH Models

First proposed by Engle (1982) and Bolorslev(1985), the GARCH models are a more complete generalization of the ARCH(q) model. The ARCH models are good for modeling scenarios on which variance is not constant. Basically, both models(ARCH and GARCH), deal with time series on which there exists a clustering of the volatility, the difference is that the GARCH model takes this "clustering" as a stochastic process. As we can see, said clustering is given by Figure 5.1.

There is strong empirical evidence that the financial markets behave like a GARCH process (Heston & Nandi, 2000). Furthermore, it makes a lot of sense to think about them this way because the inherent randomness of the market. A good reference for ARCH and GARCH models (and for time series analysis for that matter) is Shumway and Stoffel (2011). In general, GARCH(p,q) models have the form:

$$y_t = \ln\left(\frac{x_t}{x_{t-1}}\right), \tag{5.1}$$

$$y_t = \sigma_t z_t, \tag{5.2}$$

$$z_t \sim^{iid} N(0,1), \tag{5.3}$$

$$\sigma_t^2 = \alpha_0 + \sum_{j=1}^{m} \alpha_j^2 y_{t-j} + \sum_{j=1}^{r} \beta_j^2 y_{t-j}. \tag{5.4}$$

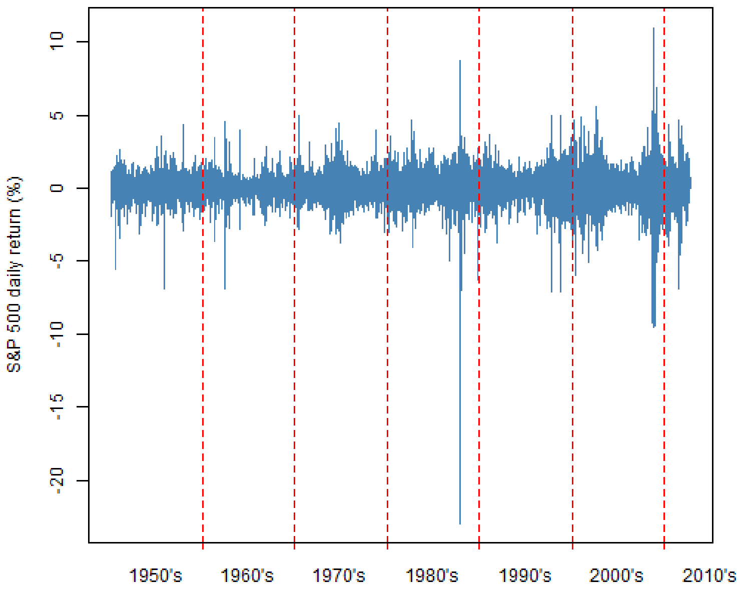

FIGURE 5.1: Volatility clustering for the S&P 500. We can see that there are regions with high returns and low returns.

In general, the most used model in the financial mathematics literature seem to be GARCH (1,1). However, there is no "rule" that says that we can't choose another GARCH (p,q) model, it's just that GARCH(1,1) is the simplest version, and it's also the easiest to implement. For the GARCH (1,1) model we get that:

$$\sigma_t^2 = \alpha_0 + \alpha_1^2 y_{t-1} + \beta^2 y_{t-1}. \tag{5.5}$$

Duan (1990) proposes that the returns of an asset (generally a stock) follow a distribution of the form:

$$\ln\left(\frac{S_t}{S_{t-1}}\right) = r - \frac{1}{2}h - t + \epsilon_t, \tag{5.6}$$

$$\epsilon_t \sim N(0, h_t), \tag{5.7}$$

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i(\epsilon_{t-i} - \lambda\sqrt{h_{t-i}})^2 + \sum_{i=1}^{p} \beta_i h_{t-i}. \tag{5.8}$$

This was the first approach to the option pricing using GARCH models. He gives that the stock price is given by:

$$S_T = S_t \exp\left[(T-t)r - \frac{1}{2}\sum_{s=t+1}^{\tau} h_s + \sum_{s=t+1}^{\tau} \epsilon_s\right]. \tag{5.9}$$

And, the price of the call option is then given by:

$$C = e^{-(T-t)r} E^{\mathbb{Q}}[\max(S_T - K), 0], \tag{5.10}$$

Where T is the maturity and K is the strike price of the option.

5.2 The Heston Nandi Model

Heston and Nandi proposed a model similar to Duan's. However, theirs has a closed form solution. From their original paper (Heston & Nandi, 2000), we can see that this model is based on two assumptions; the first one is that the asset price, $S(t)$, uses the following process on a time step, δ:

$$\log(S(t)) = \log(S(t-\Delta)) + r + \lambda h(t) + \sqrt{h(t)}z(t), \tag{5.11}$$

$$h(t) = \omega + \sum_{i=1}^{p} \beta_i h(t-i\Delta) + \sum_{i=1}^{q} \alpha_i(z(t-i\Delta) - \gamma_i\sqrt{h(t-i\Delta)})^2, \tag{5.12}$$

where r is the interest rate for the given time step and $z(t)$ is a random number that follows the standard normal distribution. γ and λ are parameters to be estimated in a similar form as the parameters in the Heston model (chapter 4). In this case, $h(t)$ is the conditional variance. We can see that this equation is fairly similar to equation (5.8).

The second assumption is that the option price follows a "Black-Scholes-like" pricing formula (see Heston and Nandi, 2000 for the complete proof and derivation). That is, the price of the option has the following form:

$$C = e^{-r(T-t)} E_t^{\mathbb{Q}}[\max(S(T) - K, 0)] \tag{5.13}$$

$$= S(t) \left(\frac{1}{2} + \frac{e^{-r(T-t)}}{\pi} \int_0^\infty Re \left[\frac{K - i\phi f^*(i\phi + 1)}{i\phi} \right] d\phi \right)$$
$$-Ke^{-r(T-t)} \left(\frac{1}{2} + \frac{1}{\pi} \int_0^\infty Re \left[\frac{K - i\phi f^*(i\phi)}{i\phi} \right] d\phi \right), \tag{5.14}$$

where $E_t^{\mathbb{Q}}$ denotes the expected value under the risk neutral (equivalent martingale) measure and $f^*(\phi)$ is the moment generating function of the logarithm of S(T) (equation (5.11)). According to Heston And Nandi, the generating function takes the following form:

$$f(\phi) = S(t)^\phi \exp[A(t; T, \phi)$$
$$\sum_{i=1}^{p} B_i(t; T, \phi) h(t + 2\Delta - i\Delta) + \sum_{i=1}^{q-1} C_i(t; T, \phi)(z(t + \Delta - i\Delta)$$
$$-\gamma_i \sqrt{h(t + \Delta - i\Delta)^2})], \tag{5.15}$$

where A and B can be calculated using the following recursions:

$$A(t; T, \phi) = A(t + \Delta; T, \phi) + \phi r + B_1(t + \Delta; T, \phi)\omega -$$
$$\frac{1}{2} \ln(1 - 2\alpha_1 \beta_1(t + \Delta; T, \phi)), \tag{5.16}$$

$$B_1(t; T, \phi) = \phi(\lambda + \gamma_1) - \frac{\gamma_1^2}{2} + \beta_1 B_1(t + \Delta; T, \phi) + \tag{5.17}$$
$$\frac{(1/2)(\phi - \gamma_1)^2}{1 - 2\alpha_1 \beta_1(t + \Delta; T, \phi)}.$$

For the particular case of a GARCH (1,1) model ($p = q = 1$), we can work the equation backwards by introducing the following end conditions, as stated by Heston and Nandi:

$$A(T; T, \phi) = 0,$$

$$B_1(T; T, \phi) = 0.$$

These equations seem to be very cumbersome and scary, however, they are relatively easy to evaluate at MATLAB. Using the `quad(), quadl()` or `integral()` integrated functions (Attaway, 2013), as well as a nested function to compute the recursive relation for A and B, they can be evaluated relatively easy.

5.3 implementation

Now, I will use the Heston-Nandi method to price some options. I choose this method over the Duan's mostly because it is more used in the industry. For this section, I shall use stock data from Yahoo, Inc. Since this option pricing model requires the stock price to behave as a GARCH process, I start by fitting a GARCH(1,1) to the log returns. Doing so, I get that:

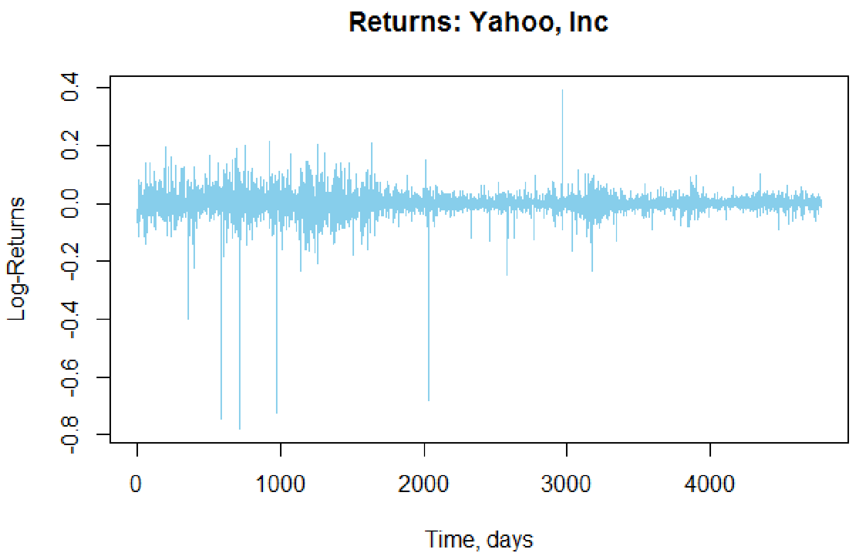

FIGURE 5.2: Log Returns for Yahoo Inc from 2004 to 2015. We can see how volatility is clustered.

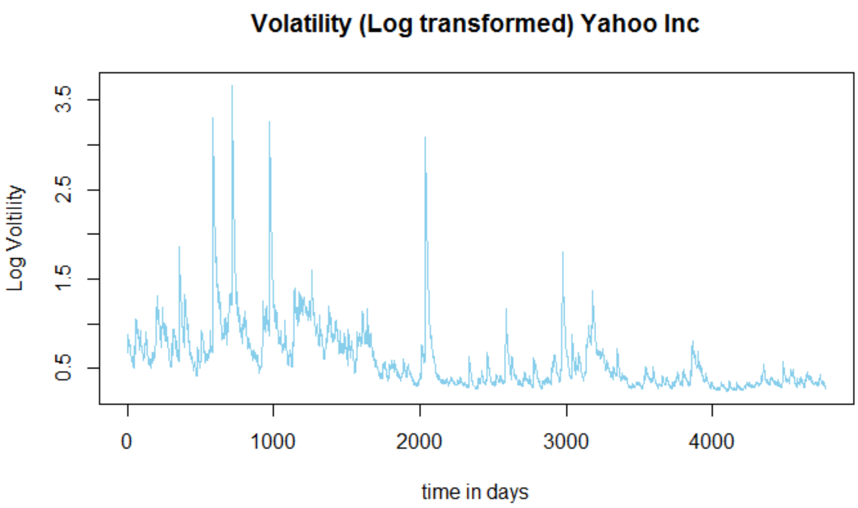

FIGURE 5.3: Volatility for Yahoo, Inc.

We can see that the volatility is clustered; there are periods of time with more volatility than others. Fitting a GARCH(1,1) model and performing the pricing algorithm, we get that the parameters are:

λ	r	α_1	β_1	γ_1
9.058e-02	-6.09107e-04	1.621e-05	9.905e-01	152.4823

TABLE 5.1: Parameters for the GARCH(1,1) model.

And they yield the following result:

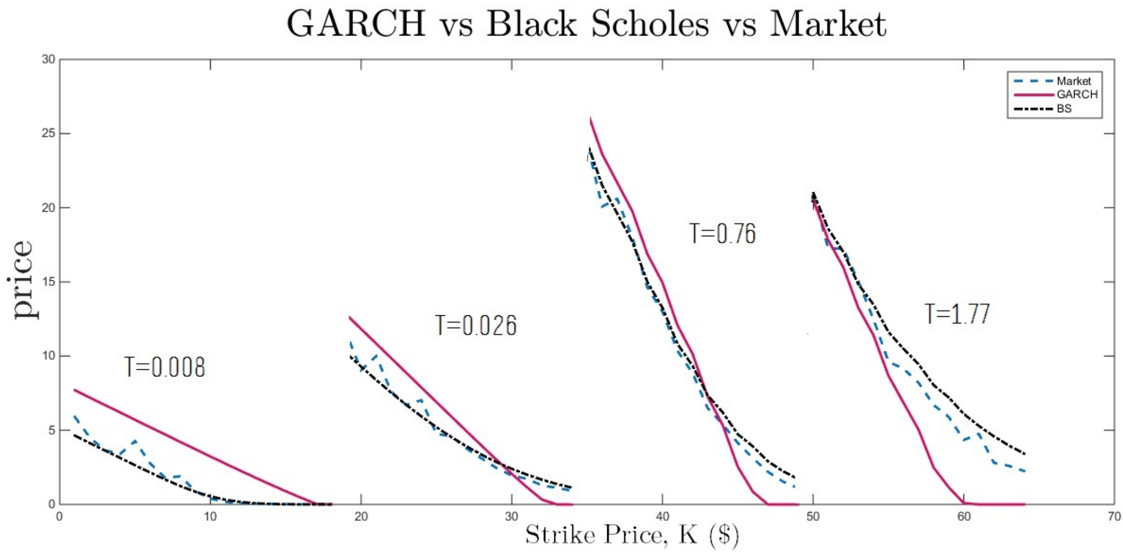

FIGURE 5.4: GARCH model Vs Black Scholes and Vs Market price. We can see that the GARCH approximation is good, however, it can bee improved.

As we can see, the GARCH model seems to follow the market and Black Scholes price very closely. However, it isn't the most accurate and it is exceeded by the Black-Scholes model. Regardless of this, we can't ignore how interesting it is for it to behave on a similar manner as the market and the Black Scholes model. This method can be extended and refined in order to give better results, as we shall see.

5.4 Modifications: IGN and Hyperbolic Distributions

We have seen that the traditional GARCH(1,1) model gives "fair enough" results. However, as noted by Chorro et al (2010), the returns are better modeled using generalized hyperbolic innovations, instead of just normal innovations. That is, the term z_t is a random number from a generalized hyperbolic distribution, instead than from a normal distribution. Studies have been made mainly using the hyperbolic distribution and the inverse normal Gaussian distribution. Said distribution is given by its PDF:

$$f(\gamma, \beta, \alpha, \delta, \mu, x) = \frac{(\gamma/\delta)^{\lambda}}{\sqrt{2\pi}K_{\lambda}(\delta\gamma)} e^{\beta(x-\mu)} \times \frac{k_{\lambda-1/2}(\alpha\sqrt{\delta^2 + (x-\mu)^2})}{(\sqrt{\delta^2 + (x-\mu)^2/\alpha})^{1/2-\lambda}}, \tag{5.18}$$

with parameters α and λ, as well as with scale parameter δ and asymmetry parameter β. $\gamma = \sqrt{\alpha^2 - \beta^2}$. K_{λ} is the modified Bessel function of the third kind. This is a super class of distributions of which the hyperbolic, student's-t, Laplace and inverse normal Gaussian distribution are derived from.

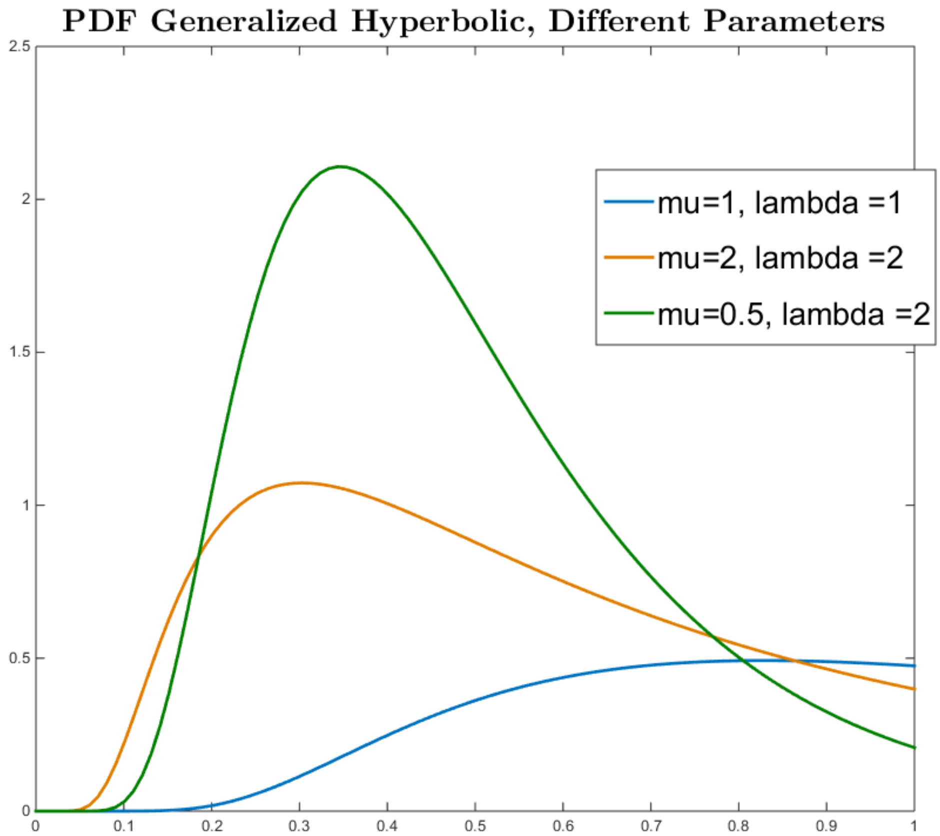

FIGURE 5.5: Different GH-PDF plotted using different parameters.

The reason to choose this rather *cumbersome* distribution is that the innovations for the GARCH model not always behave better assuming that they are normal. The proposal of this method comes from empirical evidence (Chorro, 2012). A more in depth discussion is given by Chorro (2012). The implementation of the method is rather similar to the Heston-Nandi and Duan approach, the only major difference is the distribution of the innovations, which takes a generalized hyperbolic distribution.

I shall limit this text to run and simulate the proposed model for a normal inverse Gaussian (NIG) distribution, which has the form:

$$f_{NIG}(x;\alpha,\beta,\mu,\delta) = \frac{\alpha\delta K_1\left(\alpha\sqrt{\delta^2+(x-\mu)^2}\right)}{\pi\sqrt{\delta^2+(x-\mu)^2}}e^{\delta\gamma+\beta(x-\mu)}, \tag{5.19}$$

where K_1 denotes a modified Bessel function of the third kind, μ is the location, α is the tail heaviness, β is the asymmetry parameter and δ is the scale parameter. Lastly, $\gamma = \sqrt{\alpha^2 - \beta^2}$. To test the modified model, I use a normal inverse Gaussian distribution. The parameters for this GARCH (1,1) model are given the table bellow[1]. The data is the same as used before; the Yahoo data:

Parameter	Value
r	-6.091074e-04
α_1	-5.243848e-01
β_1	5.402940e-01
ω	1.130127e-05
α	7.749328e-02
β	9.206399e-01
δ	-4.268686e-02
μ	6.597487e-01
γ_1	152.4823

TABLE 5.2: Parameters for the GARCH(1,1) model.

Running the model, we get that the results are:

[1]Parameters obtained on `R`

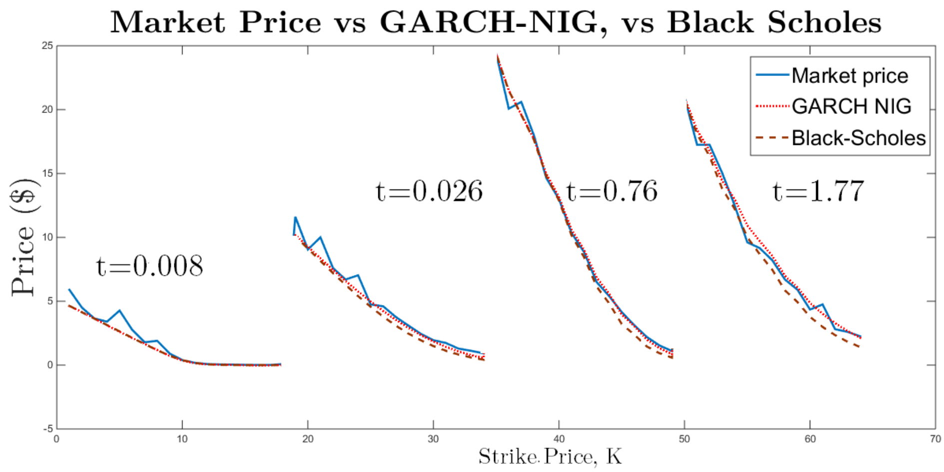

FIGURE 5.6: Comparison of the modified GARCH and the Black Scholes. Time is given in years.

We can see that the modification of the GARCH gives a much better result compared to the one obtained on figure 5.4. The justification is that the returns are not always normally distributed and adding a skewness to the distribution, such as the one portrayed on figure 5.5. Implementing this "skewness" captures better the nature of the market, including leverage effects, as explained by Chorro.

5.5 Discussion

We have then seen how the GARCH models are also quite useful to price options, although, they are not as efficient when used with normal innovations. These type of models are much more useful when different innovations are used, because the skewness of the hyperbolic distributions capture the essence of the market on a more "loyal" manner. This presents a trade off because we are compelled to use a more complex model and therefore we increase the computational cost overall.

In general, using results from time series analysis mixed with concepts from the classical option pricing theory is a rather new and interesting concept, and it is something that should be studied in more depth. I was very satisfied that the model worked for this asset, mainly because in the literature they only use the S&P 500 to test these models, and, being such an important economic indicator, the S&P500 is affected by several factors, many more that could affect an asset.

Part III

Final Remarks

Chapter 6

Discussion

As stated at the beginning of this project, financial models are always in constant change. Coming up with different ideas as well as implementations, discussions and modifications of the already existing methods is always desirable.

At the first part of this project, I presented some modifications to the Black Scholes model. Some of them are rather empirical and are based on how the volatility behaves throughout the price space or time. However, it has been shown that relaxing the assumption of a constant volatility yields, in general, better results under certain circumstances. Therefore, it is our duty to continue exploring these type of models and applying them under optimal circumstances, so we can have better predictions and obtain better prices. To my opinion, the most remarkable method is the one on which the volatility is modeled as a Markov process in the strike price state. Not only it yields better results, it is a rather elegant approach to model the volatility dynamics. I will strive to continue research on this topic.

Regarding the Heston model, I've taken a practical approach to it and discussed the method as well as the developments and modifications that it has been through lately. In general this model produces rather good results and I am satisfied by simulating and comparing them. It is also interesting to see how a process that is stochastic in nature yields such accurate results. In particular, I believe that the FFT approach is the best for this situation.

Regarding the GARCH models, which are relatively new, we can see that they model the option prices fairly well when the normal innovations are used, and that they are yield very accurate prices when different (NIG) innovations are used. This is because the real nature of the market doesn't yield normally distributed returns, but rather a more skewed distribution.

We can summarize the results and the decision making scheme for what model to choose using the following diagram:

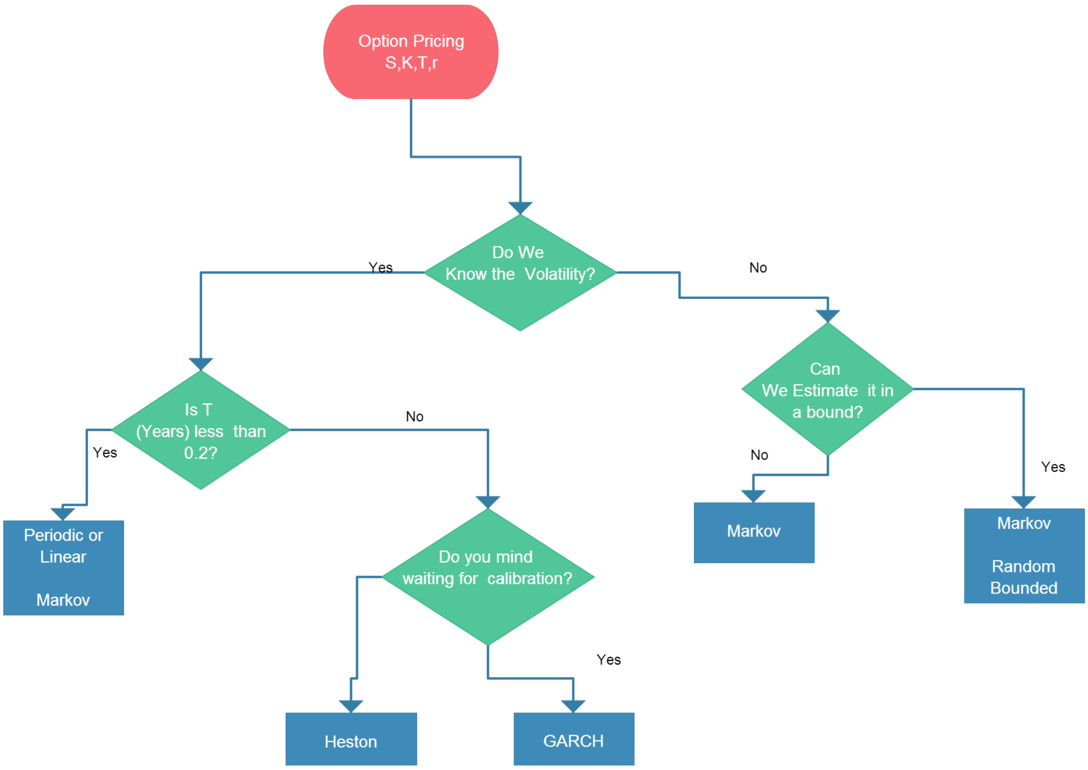

FIGURE 6.1: Decision making flow chart for the choice of model

As stated at the beginning of this document, it is important to test and evaluate models on liquid assets (assets that are being traded right now) so we can model and have better predictions for future assets. Also, having a wide variety of methods increases our "financial toolbox", therefore being able to make better decisions and reduce risk.

Bibliography

Agnolucci, P. (2009). Volatility in crude oil futures: a comparison of the predictive ability of GARCH and implied volatility models. Energy Economics, 31(2), 316-321.

Albrecher, H., Mayer, P., Schoutens, W., & Tistaert, J. (2006). The little Heston trap. status: published, 6(05).

Avellaneda, M., Levy, A., Parás, A. (1995). Pricing and hedging derivative securities in markets with uncertain volatilities. Applied Mathematical Finance, 2(2), 73-88.

Avellaneda, M., Paras, A. (1996). Managing the volatility risk of portfolios of derivative securities: the Lagrangian uncertain volatility model. Applied Mathematical Finance, 3(1), 21-52.

Attaway, S. (2013). Matlab: a practical introduction to programming and problem solving. Butterworth-Heinemann.

Carr, P., Madan, D. (1999). Option valuation using the fast Fourier transform. Journal of computational finance, 2(4), 61-73.

Chorro, C., Guégan, D., & Ielpo, F. (2012). Option pricing for GARCH-type models with generalized hyperbolic innovations. Quantitative Finance, 12(7), 1079-1094.

Chorro, C., Guegan, D., & Ielpo, F. (2008). Option pricing under GARCH models with generalized hyperbolic innovations (II): data and results (No. b08047). Université Panthéon-Sorbonne (Paris 1), Centre d'Economie de la Sorbonne.

Duan, J. C., Gauthier, G., & Simonato, J. G. (1997). An analytical approximation for the GARCH option pricing model. École des hautes études commerciales, Groupe de recherche en finance.

Duan, J., Gauthier, G., Simonato, J., & Sasseville, C. (2006). Approximating the GJR-GARCH and EGARCH option pricing models analytically. Journal of Computational Finance, 9(3), 41.

Elliott, W. D. (2015). Analyzing Options Market Toxicity and the Black-Scholes Formula in the Presence of Jump Diffusion as Simulated with Agent-Based Modeling. Undergraduate Economic Review, 11(1), 14.

Friz, P., Gatheral, J. (2005). Valuation of volatility derivatives as an inverse problem. Quantitative Finance, 5(6), 531-542.

Gatheral, J. (2011). The volatility surface: a practitioner's guide (Vol. 357). John Wiley & Sons.

Gatheral, J., Hsu, E. P., Laurence, P., Ouyang, C., & Wang, T. H. (2012). Asymptotics of implied volatility in local volatility models. Mathematical Finance, 22(4), 591-620.

Guernsey, S. (2014). Analyzing a Non-linear Black-Scholes Option Pricing Model: The Cost of Constant Volatility

Hatfield, G. (2009). A Note Regarding "Risk Neutral" and "Real World" Scenarios—Dispelling a Common Misperception

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. Review of financial studies, 6(2), 327-343.

Heston, S. L., Nandi, S. (2000). A closed-form GARCH option valuation model. Review of Financial Studies, 13(3), 585-625.

Hull, J. C. (2006). Options, futures, and other derivatives. Pearson Education India.

Lewis, A. L. (2000). Option valuation under stochastic volatility. Option Valuation under Stochastic Volatility.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. Journal of financial economics, 3(1), 125-144.

Mikhailov, S., Nogel, U. (2004). Heston's stochastic volatility model: Implementation, calibration and some extensions.

Moodley, Nimalin. (2005). The Heston Model. A practical Approach.

Qiu, Y., Lorenz, J. (2009). A non-linear Black-Scholes equation. International Journal of Business Performance and Supply Chain Modelling, 1(1), 33-40.

Rouah, F. D. (2013). The Heston Model and Its Extensions in Matlab and C#. John Wiley & Sons.

Sauer, Tim. (2011). Numerical Analysis. Wiley.

Schöbel, R., Zhu, J. (1999). Stochastic volatility with an Ornstein–Uhlenbeck process: An extension. European Finance Review, 3(1), 23-46.

Shumway, R. H., Stoffer, D. S. (2010). Time series analysis and its applications: with R examples. Springer Science & Business Media.

Stein, Elias M., Rami Shakarchi. 2005. Real analysis: measure theory, integration, and Hilbert spaces. Princeton, N.J.: Princeton University Press.

Wilmott, P. (1995). The mathematics of financial derivatives: a student introduction. Cambridge University Press.