University of Utrecht

TWIN Mathematics and Physics

Bachelor Thesis

Identifying stock market bubbles

Author:
C.A.H. Vermeulen
5732115

Supervisor:
Karma Dajani

A mathematical approach to predicting the stock market

June 16, 2019

Abstract

In this thesis we look at a mathematical way of noticing changes in the financial market. Big changes would result in economic bubbles. With the use of *Python* and *Mathematica* we simulated stocks, compared calculated option prices to real data and calculated the implied volatility and the illiquidity parameter. We did this not with the use of the ordinary Black-Scholes model, but with the use of the lesser known Kou model. In contrary to the Black-Scholes model the Kou model incorporates sudden changes of a stock price in this model, which we call jumps. We found that the Kou model is also as good as or even better approximation for the stock market than the Black-Scholes model. The illiquidity parameter is a self made parameter representing the time it takes to sell or buy something without causing its price to change. This stands closely connected with crises and when calculated from the Kou model, laid next to the graph of the $SP500$ index price shows signs of destabilization and uncertainty which leads to a crash in the stock market.

Contents

1 Introduction

There has always been a need for predictions in our world. Humanity does not want their future to depend on fate. We want to make our own future through any means necessary. This also happened, and is still happening, in the world of finance and stock markets. Through predictions we can make money, if you play your cards right. Knowing more than the others is an advantage what many people seek to find. But in a world so volatile as the financial world models for predictions can never fully determine what will happen next. There are occasions caused by human emotions that leads to uncertainty and panic. If this happens on a large scale the financial market will notice some irregularities, which can form an economic bubble.

In this thesis we are going to look at two stock predicting models, namely the Merton model and the Kou model. Our focus will be more on the latter because of its and ability to explain two major empirical phenomena that occur in the distribution of stocks and that the model is fairly new which makes the research more challenging. We will show a comparison of these two models and aim to draw some conclusions.

In chapters 3 to 6 we will provide the necessary mathematical background needed for developing an understanding of how and why these models work. This is needed otherwise we cannot substantiate our inferences. Both models mentioned before are build up from stochastic processes, especially by a Brownian motion, and rely on the Itô calculus for giving a sensible expression.

After this we are going to derive the Merton and Kou model in chapters 7 to 9, two alternatives to price options in the stock market. Using as base the geometric Brownian motion and the famous Black-Scholes equation, we are going to add "jumps", hoping for a more precise approximation of options.

At last in chapter 10, we finally come to the aim of this thesis. Showing the results we got from implementing these models in Python with the use of collected data we are going to calculate a parameter which will us something about bid ask spread. Through the mean squared error function we will derive this "illiquidity" parameter which can be seen as a financial trouble parameter. It is a trend that this becomes big before a economic bubble. So if everything goes right we will have with certainty a parameter that can predict an economic bubble

2 Background on the stock market

In this thesis we are going to use some terms which will be unknown for some readers. In this chapter these terms will be discussed and explained for a better understanding of this thesis and the stock market in its entirety.

Options

An option is a contract which gives the buyer of this contract the right to sell or buy an underlying asset, which is most of the time a stock, at a specified strike price before or on a specified time (expiration time). It depends on the kind of option which gives the owner of the contract to buy or sell a certain stock. A call option gives the owner the right to buy an asset at a specific price, where a put option gives the owner the right to sell an asset at a specific price. Hereby has the seller of the contract the obligation to fulfill the transaction if the buyer chooses to "exercise" his right.

Bid and Ask prices

The bid and ask prices are related to the option price. people are trading stocks and thereby trading options continuously. The price a seller wants for his option is in a sane world always higher than the price a buyer is willing to pay for an option. So in the financial world we talk about bid and ask prices. The bid price is the highest price that a buyer is willing to pay, whereas the ask price is the lowest price that the seller is willing to accept for his assets.

Volatility

The volatility will be frequently used in formulas for option and/or stock pricing. The volatility or σ is the degree of variation a trading price has over time as measured by the standard deviation. The higher the volatility of a certain option, the more uncertain you are about the future of the price. Frequently traded assets thereby have low volatility because of the low risk whereas not frequently traded assets have high volatility.
In chapters 7 to 9 we will consider *Implied volatility.* This means that the volatility is being derived from the market price of an option, using the option price, index price and strike price of the option. This can be used to look forward in time, predicting which way an asset is leaning towards.

Liquidity

Lastly, liquidity or its counterpart illiquidity, represents how fast you can buy or sell an asset without causing a drastic change in the asset's price. For an asset to be very liquid, selling quickly will not reduce the price much. But to sell an illiquid asset quickly, you need to cut its price by some amount, otherwise it would difficult to sell. Liquidity could be used to represent in which state the market is in. Especially when you look at economic bubbles it is good to look at the liquidity of the market. When the economy is good and prosperous the liquidity of a market is low. But when people are frightened and the faith in the economy declines, people tend to sell everything they got before their assets will be worth nothing. If everyone does this panic move, the market becomes illiquid, due to the overkill of certain assets in the market. This also happened in 2008 when the housing market in America collapsed causing people to dump all their shares, trying to make a little bit of cash before they get worthless.

3 Conditional Expectation

In this chapter we look at the conditional expectation which is an extension of the expectation value of a certain random variable. It will give the estimated average value of the random variable under certain conditions. Formally, we consider a random variable X defined on a probability space. If X is $\mathcal{G}$-measurable then we can easily determine X with the information $\mathcal{G}$ provides. If X is independent of $\mathcal{G}$ then information of $\mathcal{G}$ is not enough to evaluate X. In the intermediate case we can use $\mathcal{G}$ to estimate X but not evaluate precisely. The conditional expectation of X given $\mathcal{G}$ is this estimate.

3.1 Basic definitions

Definition 3.1 (Conditional expectation given an event) *Consider a random variable X, which is either non-negative or integrable, on a probability space $(\Omega, \mathcal{F}, P)$ and an event $A \in \mathcal{F}$, with strictly positive probability $\mathbb{P}(A) > 0$, then the conditional expectation of X given A is defined by,*

$$\mathbb{E}[X|A] = \frac{\mathbb{E}[\mathbb{I}_A X]}{\mathbb{P}(A)} = \frac{\int_A X d\mathbb{P}}{\mathbb{P}(A)} \tag{1}$$

Definition 3.2 (Conditional expectation given a σ-algebra $\mathcal{G}$) *Consider a probability space and let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$, and X be a random variable, either non-negative or integrable. The conditional expectation of X given $\mathcal{G}$is then a new random variable $E[X|\mathcal{G}]$ satisfying the following conditions:*

1. $\mathbb{E}[X|\mathcal{G}]$ *is G-measurable*
2. $\int_A \mathbb{E}[X|\mathcal{G}] d\mathbb{P} = \int_A X d\mathbb{P}$

Theorem 3.3 *Given a probability space and let $\mathcal{G}$ and $\mathcal{H}$ be sub-σ-algebras of a σ-algebra $\mathcal{F}$. Then the following hold:*

1. *(Linearity) For constants a,b*
   $\mathbb{E}[aX + bY|\mathcal{G}] = a\mathbb{E}[X|\mathcal{G}] + b\mathbb{E}[Y|\mathcal{G}]$
2. *(Taking out what is known)If X is $\mathcal{G}$-measurable*
   $\mathbb{E}[XY|\mathcal{G}] = X\mathbb{E}[Y|\mathcal{G}]$
3. *(Tower property) IF $\mathcal{H} \subset \mathcal{G}$*
   $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[H]$
4. *(Average of average)*
   $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[X]$
5. *(Independence)If X is independent of $\mathcal{G}$*
   $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$
6. *(Positivity) If $X \geq 0$*
   $\mathbb{E}[X|\mathcal{G}] \geq 0$
7. *(Jensen's inequality) If ϕ is convex*
   $\mathbb{E}[\phi(X)|\mathcal{G}] \geq \phi(\mathbb{E}[X|\mathcal{G}])$

Definition 3.4 (Conditional expectation given a random variable) *Given X, Y random variables on a probability space $(\Omega, \mathcal{F}, P)$. Then we define the conditional expectation of X given the sub-σ-algebra of Y, $\sigma(Y)$, as $\mathbb{E}[X|Y]$.This is then $\sigma(Y)$-measurable and with the assumption that X is integrable satisfies for any $A \in \sigma(Y)$,*

$$\int_A \mathbb{E}[X|Y] d\mathbb{P} = \int_A X d\mathbb{P} \tag{2}$$

3.2 Notes

If we now go back to probability theory, a conditional expectation is itself a random variable and measurable with respect to the σ-algebra generated by the condition. This way of looking is necessary for the conditional expectations that arise in martingale theory, which we will discuss shortly in Chapter 4.

4 Stochastic processes

In this chapter we discuss stochastic processes and their fundamental importance for financial mathematics. A stochastic process or otherwise referred to as a random process is a collection of random variables indexed by some mathematical set and can be divided in discrete-time or continuous-time.

Definition 4.1 (Stochastic Process) *1. A stochastic process is a sequence of random variable $X_t : \Omega \to \mathbb{R}$ parameterized by time t belonging to an index set $I \subset \mathbb{R}$. In other words, when a stochastic process $X : I \times \Omega \to \mathbb{R}$ is given, $X(t, \cdot) = X_t : \Omega \to \mathbb{R}$ is a measurable mapping for each $t \in I$.*

2. *If I is a discrete set, then a process $\{X_t\}_{t\in I}$ is called a discrete time stochastic process, and if I is an interval then $\{X_t\}_{t\in I}$ is called a continuous time stochastic process.*

3. *For each $\omega \in \Omega$ the mapping $t \mapsto X_t(\omega)$ is called a sample path.*

4. *If almost all sample paths of a continuous time process are continuous, then we call the process a continuous process.*

5. *The filtration $\mathcal{F}_t$ generated by a process X_t, i.e., $\mathcal{F}_t = \sigma(\{X_s : 0 \le s \le t\})$ is called a natural filtration for X_t.*

If we talk about stochastic processes in the future, we always imply that it is defined on some probability space $(\Omega, \mathcal{F}, P)$

Definition 4.2 (Adapted process) *Consider a filtration $\{\mathcal{F}_t\}_{t\in I}$ and a stochastic process $\{X_t\}_{t\in A}$. If X_t is measurable with respect to $\mathcal{F}_t$ for every t, then $\{X_t\}_{t\in I}$ is said to be adapted to the filtration $\{\mathcal{F}_t\}_{t\in I}$.*

4.1 Martingales

Definition 4.3 (Martingale) *Suppose that a stochastic process $\{X_t\}_{t\in I}$ is adapted to a filtration $\{\mathcal{F}_t\}_{t\in I}$, and that X_t is integrable for every t, i.e., $\mathbb{E}[|X_t|] < \infty$. If*

$$X_s = \mathbb{E}[X_t|\mathcal{F}_s] \tag{3}$$

for arbitrary $s \le t$, then $\{X_t\}_{t\in I}$ is called a martingale with respect to $\{\mathcal{F}_t\}_{t\in I}$. If

$$X_s \le [X_t|\mathcal{F}_s] \tag{4}$$

for $s \le t$, then it is called sub-martingale, and if

$$X_s \ge [X_t|\mathcal{F}_s] \tag{5}$$

for $s \le t$, then it is a super-martingale.

Theorem 4.4 *If $\{X_t\}_{t\ge 0}$ is a martingale with respect to a filtration $\{\mathcal{F}_t\}_{t\in I}$, where $\mathcal{F}_0 = \{\emptyset, \Omega\}$, then $\mathbb{E}[X_t] = \mathbb{E}[X_0]$ for every t.*

4.2 Random walks

A random walk is a particular stochastic process which is at the foundation of our search to the optimal model for the stock market. Random walks are as followed defined:

Definition 4.5 *Let $Z_n, n \geq 1$, be a sequence of i.i.d. random variables defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $\mathbb{P}(Z_n = 1) = p$ and $\mathbb{P}(Z_n = -1) = 1 - p$ for some $0 \leq p \leq 1$. A one-dimensional random walk $X_0, X_1, \cdots$ is defined by $X_0 = 0, X_n = Z_1 + \cdots + Z_n, n \geq 1$. Let $\mathcal{F}_0$ be the trivial σ-algebra $\{\emptyset, \Omega\}$ and let $\mathcal{F}_n = \sigma(Z_1, \cdots, Z_n)$ be the sub-σ-algebra generated by $Z_1, \cdots, Z_n$. If $p = \frac{1}{2}$, the process $\{X_n\}_{n \geq 0}$ is called a symmetric random walk.*

Definition 4.6 *A stochastic process $\{X_t\}_{t \geq 0}$ is called a Lévy process if it satisfies the following conditions*

1. *$X_0 = 0$ almost surely*
2. *All increments $X_t - X_s$ are independent of $\mathcal{F}_t$ for any $0 \leq s < t \leq T$. X_t has independent increments.*
3. *For any $s < t$ the increment $X_t - X_s$ is equal in distribution as X_{t-s}, X_t has stationary increments*
4. *For every $o \leq t \leq T$ and $\epsilon > 0$: $\lim_{h \to 0} P(|X_{t+h} - X_t| > \epsilon) = 0$*

with the meaning of stationary increments is that the probability distribution of a increments only depends on the length such that every increment with the same length is evenly distributed.

Theorem 4.7 (Quadratic Variation) *Let $\{X_t\}_{0 \leq t \leq T}$ be a continuous martingale. The quadratic variation process of X_t, denoted by $[X, X]_t$ or $[X]_t$, is defined by:*

$$[X, X]_t = \lim_{n \to \infty} \sum_{j=1}^{n} |X_{t_j} - X_{t_{j-1}}|^2 \tag{6}$$

The convergence used for the quadratic variation is the convergence in mean square, otherwise called the L^2-convergence. In the next chapter we will calculate the quadratic variation of the Brownian Motion. We can base on that outcome a "new" calculus, namely Itô calculus.

5 Brownian Motion

In this chapter we look at the concept of Brownian motion and how it is defined and how it can be used for our purposes in the stock market. Brownian motion is a common phenomenon in physics and nature. In mathematics it is described by the Wiener process; a continuous-time stochastic process. That is why we write "W_t" for Brownian motion in further definitions and proofs.

5.1 Basic principles

To obtain the Brownian motion as a stochastic process we take the limit of the scaled random walk as $n \to \infty$. The Brownian motion retains the same properties as the random walk, and we will define it as a stochastic process $W_t, t \geq 0$ with the following properties:

1. W_t is a almost surely continuous function
2. W_t has independent increments
3. Each of these increments is normally distributed: $W_t - W_s\ N(0, t-s)$ (for $0 \leq s \leq t$)

If additional $P(W_0 = 0) = 1$ applies then the Brownian motion is called *standard.*

Theorem 5.1 (Quadratic Variation of Brownian Motion) *Let W_t be a Brownian motion. Then $[W,W]_t = t$ almost surely for all $t \geq 0$.*

Here $[W,W]_t$ is defined as in theorem 4.7.

Proof Let W_t be a Brownian motion, and define the partition $0 < t_1 < \cdots < t_n < t$. We define the sampled quadratic variation corresponding with this partition as

$$\sum_{j=0}^{n-1}(W(t_{j+1}) - W(t_j))^2 \tag{7}$$

We need to show that this sampled quadratic variation converges to t as $n \to \infty$. We know already that

$$\mathbb{E}[(W(t_{j+1}) - W(t_j))^2] = Var[W(t_{j+1}) - W(t_j)] = t_{j+1} - t_j \tag{8}$$

Without further showing we know that

$$Var[(W(t_{j+1}) - W(t_j))^2] = 2(t_{j+1} - t_j)^2 \tag{9}$$

If we now take a standard normal random variable as

$$Y_{j+1} = \frac{W(t_{j+1}) - W(t_{j+1})}{\sqrt{t_{j+1} - t_j}} \tag{10}$$

And let us choose a large value for n and take $t_j = \frac{jT}{n}$, such that $t_{j+1} - t_j = \frac{T}{n}$ for all j. Then we can use the law of large numbers which implies that $\sum_{j=0}^{n-1} \frac{Y_{j+1}^2}{n}$ converges to $\mathbb{E}Y_{j+1}^2$ as $n \to \infty$, which will become 1. So we get

$$\lim_{n\to\infty}\sum_{j=0}^{n-1}(W(t_{j+1}) - W(t_j))^2 = \lim_{n\to\infty} T * \frac{Y_{j+1}^2}{n} = T * \mathbb{E}Y_{j+1}^2 = T \tag{11}$$

So here comes the almost surely term in the theorem. Every term $(W(t_{j+1}) - W(t_j))^2$ in the summation can be different from $t_{j+1} - t_j$ but if we sum over a lot of terms they average out to zero.

This proof can also be summarized by

$$dW(t)dW(t) = dt \tag{12}$$

and this result will be used in the next chapters

6 Itô Calculus

In this chapter we look at the calculus named after Kiyoshi Itô which is often used in financial mathematics, and plays a big role in the derivation of the Black-Scholes formula. The most useful fact derived by Itô is the Itô formula, or Itô lemma, which approximates a function depending on time and Brownian motion in the same way as the Taylor series expansion. The only difference between these approximations is that in closeness of approximation the second order of the Brownian Motion will remain rather then be discarded. This will be proven with the approximation of the quadratic variation of the Brownian motion.

If we look at a normal second order Taylor expansion for a differentiable function for both t and W_t, with W_t a Brownian motion we have

$$df(t,x) = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dx + \frac{1}{2}(\frac{\partial^2 f}{\partial x^2}dx^2 + 2\frac{\partial^2 f}{\partial x \partial t}dxdt + \frac{\partial^2 f}{\partial t^2}dt^2) + \cdots \tag{13}$$

We now ignore all terms which have a order of dt-terms higher than 1, using also the fact that continuously differentiable functions have bounded variation, which makes the quadratic variation zero. Using also the fact that due to quadratic variation of a Brownian motion we have, $dW(t)dW(t) = dt$, we will get,

$$df(t,W_t) = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial W_t}dW_t + \frac{1}{2}\frac{\partial^2 f}{\partial W_t^2}dt \tag{14}$$

This is called the Itô-Doeblin formula in differential form.

Definition 6.1 (Itô-Doeblin formula for Brownian motion) *Let f(t,x) be a continuous function with partial derivatives $f_t(t,x), f_x(t,x) and f_{xx}(t,x)$ defined and continuous. Then with W_t a Brownian motion and $T \geq 0$*

$$f(T,W(T)) = f(0,W(0)) + \int_0^T f_t(t,W(t))dt + \int_0^T f_x(t,W(t))dW(t) + \int_0^T f_{xx}(t,W(t))dt \tag{15}$$

Definition 6.2 (Itô Process) *Let W(t) be a Brownian motion and let $\{\mathcal{F}_t\}_{t\geq 0}$ be an associated filtration. An Itô process is then a stochastic process of the form*

$$X(t) = X(0) + \int_0^t \Delta(u)dW(u) + \int_0^t \Theta(u)du \tag{16}$$

where X(0) is nonrandom, and $\Delta(t)$ and $\Theta(t)$ are adapted stochastic processes to $\{\mathcal{F}_t\}_{t\geq 0}$.
For simpler notation we write the Itô process in differential form

$$dX_t = \Delta(t)dW(t) + \Theta(t)dt \tag{17}$$

6.1 Lévy Characterization

A different way to characterize Brownian motions, named after the mathematician Lévy, uses martingale theory. The theorem is as followed

Theorem 6.3 (Lévy Characterization, one dimension) *Given a stochastic process $\{X_t\}_{t\geq 0}$ and the filtration $\{\mathcal{F}_t\}_{t\geq 0}$, the process X_t is a Brownian motion if and only if all of the following conditions hold:*

1. *$X_0 = 0$ with probability 1.*
2. *$\{X_t\}_{t\geq 0}$ is almost surely a continuous martingale with respect to $\{\mathcal{F}_t\}_{t\geq 0}$.*
3. *The quadratic variation $[X,X]_t = t$.*

7 Option pricing of the stock market

We first need to understand the concept of option pricing and the mathematics behind it before we can go to the more advanced models that we are going to use for calculations. Option pricing is an important way for an investor to determine what the worth of a stock is. So we are going to look at a more basic way of simulating stocks, namely by the use of the geometric Brownian motion and the Black-Scholes-Merton formula.

Definition 7.1 (Geometric Brownian Motion) *A stochastic process $\{S_t\}_{t\geq 0}$, usually the stock price, is said to follow a Geometric Brownian Motion if it satisfies the stochastic differential equation*

$$dS_t = S_t(\mu dt + \sigma dW_t), \tag{18}$$

with μ a drift term and σ is the volatility. The solution of this differential equation, for $t \geq 0$, is given by:

$$S_t = S_0 e^{(\mu - \frac{\sigma^2}{2})t + \sigma W_t} \tag{19}$$

In figure 1 we plotted several paths which follow the geometric Brownian motion. It is clear to see that the paths are continuous and tend a little bit upward, because of the drift is set to $\sigma = 0.3$.

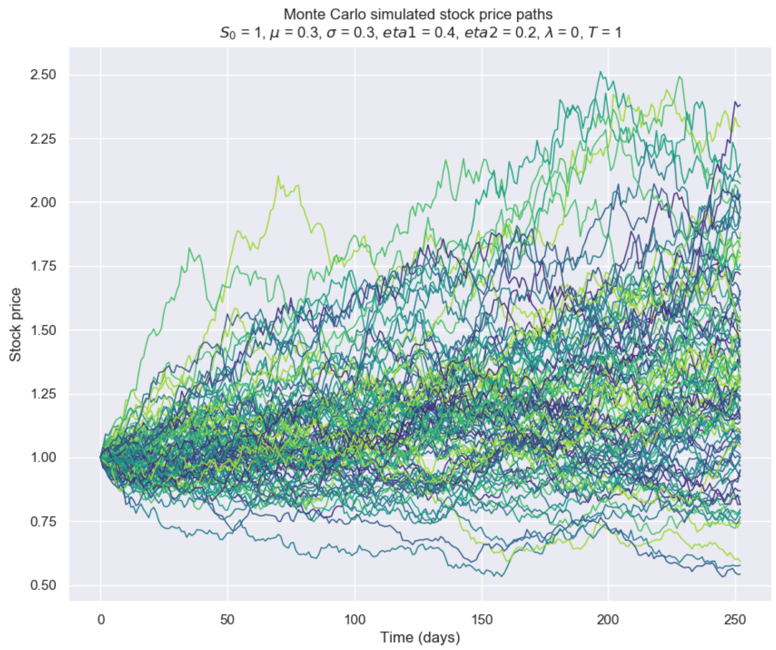

Figure 1: Monte Carlo simulated paths for geometric Brownian motion

Note: In the description of the figure stands $\lambda = 0$. This is the parameter that represents the number of jumps per time-interval and setting it to zero reduces any jump-diffusion model to the geometric Brownian motion. This comes from the fact that a jump-diffusion model is a geometric Brownian motion with some stochastic jump process added, but will be discussed in next chapter.

7.1 Black-Scholes Model

The first model we are going to look at is the Black-Scholes model. This model is the foundation of all other stock market models and is used to determine call and put options.

Definition 7.2 (Black-Scholes-Merton Formula) *The price of a European call option at time $0 \leq t \leq T$ with expiry T and strike price K is given by*

$$C_t = S_t N(d_1) - KE^{-r(T-t)} N(d_2) \tag{20}$$

where $N(x)$ is the cumulative standard normal distribution. For the European put option

$$P_t = KE^{-r(T-t)} N(-d_2) - S_t N(-d_1) \tag{21}$$

Where

$$d_1 = \frac{ln(\frac{S_t}{K}) + (r + \sigma^2)(T-t)}{\sigma(T-t)}, \;\; d_2 = \frac{ln(\frac{S_t}{K}) + (r - \sigma^2)(T-t)}{\sigma(T-t)} \tag{22}$$

Despite the success of the Black-Scholes model, it is certainly not perfect. There are some empirical phenomena which cannot be explained by this model. First there is the volatility "smile". This curve occur when solving the implied volatility through backwards calculating. However if the Black-Scholes model would be correct, the implied volatility has to be constant over time. The second phenomena we see is the asymmetric leptokurtic features of the distribution of the options. This means that this distribution is skewed to the left, has a higher peak and two heavier tails than the normal distribution has. These two phenomena will be discussed in next chapter with included figures for a better understanding.

8 Jump diffusion model

Before we look at the model we are going to use for modelling the illiquidity parameter, we will look at its predecessor which came first with the notion that stock prices can "jump". In previous chapter we discussed the important Black-Scholes model. This model is often also credited to Robert C. Merton due his research in expanding the mathematical understanding of the option-pricing. Merton proposed to add jump-diffusions to the geometric Brownian Motion which led to the beginning of jump-diffusion models and a new way of interpreting the stock market.

8.1 Merton model

The Merton Model is one of the simplest cases of a jump-diffusion model. In fact it is a combination of a Brownian Motion with a compound Poisson process. So we get a model which has an occasional jump and is continuous between the jump times. We start with the stochastic differential equation of Merton's jump-diffusion model

$$\frac{dS(t)}{S(t-)} = \mu dt + \sigma dW(t) + d\sum_{i=1}^{N(t)} Y(i) \tag{23}$$

The jumps displayed by Y_i are normally distributed, with a Poisson points process giving the frequency of a jump. It is not in our eyes an exciting model, for the jumps are, like the geometric Brownian motion, normally distributed and is the distribution function of the whole model simpler to calculate. We will look at the Kou model which will give us more of a challenge.

8.2 Kou Model

In this section we focus on the financial model proposed and named after Steven Kou. First mentioned in 2002, Kou proposed a different model for option-pricing than the more widely used models like the Black-Scholes model plus a jump-diffusion part. Instead of having a normal distribution, he looked at an asymmetric double exponential distribution for the jump that is added. According to Kou it will be the solution for some big problems we encounter by approximating the option prices. These are namely the leptokurtic feature of the prices and the famous volatility "smile". In this chapter we will not go in depth in these phenomena but show how this model explains both.

Our main goal is showing the usefulness of the Kou model, because we are going to use this model in the next chapter to numerically approximate options and compare the results with real-life data. We define the following differential equation,

$$\frac{dS(t)}{S(t-)} = \mu dt + \sigma dW(t) + d(\sum_{i=1}^{N(t)} (V_i - 1)) \tag{24}$$

Here $W(t)$ represents the standard Brownian Motion, μ the drift, $N(t)$ a Poisson process with rate λ and $\{V_i\}$ a sequence of i.i.d. non-negative random variables such that $\ln(Vi)$ follows an asymmetric double exponential distribution,with the density,

$$\begin{aligned} f_Y(y) = p * \eta_1 e^{-\eta_1 y}\mathbb{I}_{\{y\geq 0\}} + q * \eta_2 e^{-\eta_2 y}\mathbb{I}_{\{y<0\}}, \\ \eta_1 > 1,\ \eta_2 > 0,\ p, q \geq 0,\ p + q = 1 \end{aligned}$$

It is helpful to have a solution for the differential equation given by 24. We must modify the Itô-Doeblin formula, so it accepts jump processes. Let us first recall an Itô process given by 6.2. We add a right-continuous jump term J to this process to obtain a jump process, setting,

$$X(t) = X(0) + I(t) + R(t) + J(t) \tag{25}$$

with $I(t) = \int_0^t \Delta(u)dW(u)$ and $(t) = \int_0^t \Theta(u)du$. The continuous part of $X(t)$ is then defined as,

$$\begin{aligned} X^c(t) &= X(0) + I(t) + R(t) \\ dX^c(t) &= \Delta(t)dW(t) + \Theta(t)dt \end{aligned} \tag{26}$$

which is the jump process reduced to the Itô process. This is already defined in chapter 6 and using the Itô-Doeblin formula on this continuous part, we can write,

$$f(X^c(t)) = f(X^c(0)) + \int_0^t f'(X^c(u))dX^c(u) + \frac{1}{2}\int_0^t f''(X^c(u))\Delta^2(u)du \tag{27}$$

Between jumps, the jump part is equal to zero, so we have $dX(t) = dX^c(t)$ and thus because when there is a jump in X from $X(s-)$ to $X(s)$ there is also a jump in $f(X)$ from $f(X(s-))$ to $f(X(s))$. So between jumps it is fair to say that $df(X(s)) = df(X^c(s))$. When we integrate these differential from 0 to t, we must add all jumps together that occur in this time-interval. Therefor we state the following theorem.

Theorem 8.1 (Lévy-Itô formula for a jump process) *For a given jump-process given by*

$$X_t = X_0 + \int_0^t \Theta(u)du + \int_0^t \Delta dW(u) + \sum_{i=1}^{N(t)} \Delta J_i, \tag{28}$$

where b_t and σ_t are continuous processes, then for $f(x)$ a function for which the first and second derivatives are defined and continuous, then,

$$\begin{aligned} f(X(t)) = f(X(0)) &+ \int_0^t f'(X(u))dX^c(u) + \frac{1}{2}\int_0^t f''(X^c(u))\Delta^2(u)du \\ &+ \sum_{0<s\leq t} [f(X(s)) - f(X(s-))] \end{aligned} \tag{29}$$

If a jump occurs at time s, we write $f(X(s)) - f(X(s-)) = \Delta f(X(s))$ and if there is no jump at time s, we write of course $f(X(s)) - f(X(s-)) = 0$. We can combine these to get,

$$f(X(s)) - f(X(s-)) = \Delta f(X(s))\Delta N(s), \tag{30}$$

here $N(s)$ is a Poisson process and $\Delta N(s)$ is 1 if N has a jump at time s and 0 otherwise. Using this notation we can rewrite the theorem above to,

$$\begin{aligned} f(X(t)) = f(X(0)) &+ \int_0^t f'(X(u))dX^c(u) + \frac{1}{2}\int_0^t f''(X^c(u))\Delta^2(u)du \\ &+ \sum_{0<s\leq t} [\Delta f(X(s))\Delta N(s)] \\ &= f(X(0)) + \int_0^t f'(X(u))dX^c(u) + \frac{1}{2}\int_0^t f''(X^c(u))\Delta^2(u)du \\ &+ \int_0^t [\Delta f(X(s))\Delta N(s)] \end{aligned} \tag{31}$$

Now in differential form, this becomes,

$$\begin{aligned} df(X(t)) &= f'(X(t))dX^c(t) + \frac{1}{2}f''(X^c(t))\Delta^2(t)dt \\ &+ [\Delta f(X(t))\Delta N(t)] \\ &= f'(X(t))dX^c(t) + \frac{1}{2}f''(X^c(t))\Delta^2(t)dt \\ &+ f(X(t)) - f(X(t-)) \end{aligned}. \tag{32}$$

Now for our case, if we integrate function 23, we get the a price process of the form,

$$S(t) = S(0) + \int_0^t S(u)\mu du + \int_0^t S(u)\sigma dW(u) + \sum_{i=1}^{N(t)} S(i-) * (V(i) - 1), \tag{33}$$

where we used that $S(u)$ and $S(u-)$ only differ finitely many times, and when we integrate with respect to du, the differences do not matter anymore. Furthermore we notice that the first three elements are together the continuous part and the last element is the jump. Thus a way to rewrite 23 is,

$$dS(t) = S(t-)\mu dt + S(t-)\sigma dW(t) + (S(t-) * (V(t) - 1))\Delta N(t), \tag{34}$$

Implementing expression 33 in the Itô-Doeblin formula differential for a jump process gives,

$$\begin{aligned} df(S(t)) = &\frac{\partial f(S(t))}{\partial t}dt + \mu S(t)\frac{\partial f(S(t))}{\partial x}dt + \frac{\sigma^2 S(t)^2}{2}\frac{\partial^2 f(S(t))}{\partial S(t)^2}dt + \sigma S(t)\frac{\partial f(S(t))}{\partial S(t)}dW_t \\ &+ [f(S(t-) + S(t-) * (V(t) - 1)) - f(S(t-))] \end{aligned}. \tag{35}$$

Substituting $f(S(t)) = \ln(S(t))$ gives:

$$d(ln(S(t))) = \mu S_t \frac{1}{S(t)}dt - \frac{\sigma^2 S(t)^2}{2}\frac{1}{S(t)^2}dt + \sigma S(t)\frac{1}{S(t)}dW(t) + ln(S(t-)) + ln(V(t)) - ln(S(t-)) \tag{36}$$

which if we simplify and integrating (summing for the jumps) both sides, from 0 to t, results in the following equation:

$$ln(S(t)) = ln(S(0)) + (\mu - \frac{\sigma^2}{2})t + \sigma W(t) + \sum_{i=1}^{N(t)} ln(V(i)) \tag{37}$$

or

$$S(t) = S(0)e^{(\mu - \frac{\sigma^2}{2})t + \sigma W(t) + \sum_{i=1}^{N(t)} Y(i)} \tag{38}$$

with $Y(i) = \ln(V(i))$
In figure 2 we plotted some paths of 38 with the parameters $S_0 = 1, \mu = 0.3, \sigma = 0.3, \eta_1 = 0.4, \eta_2 = 0.2, \lambda = 3, T = 1$. You can clearly see the jumps happening, with bigger jumps upward than downward ($\eta_1 > \eta_2$), and with an average of $\lambda = 3$ jumps for the whole time-interval. The difference with the geometric Brownian motion paths are nice and clear to see.

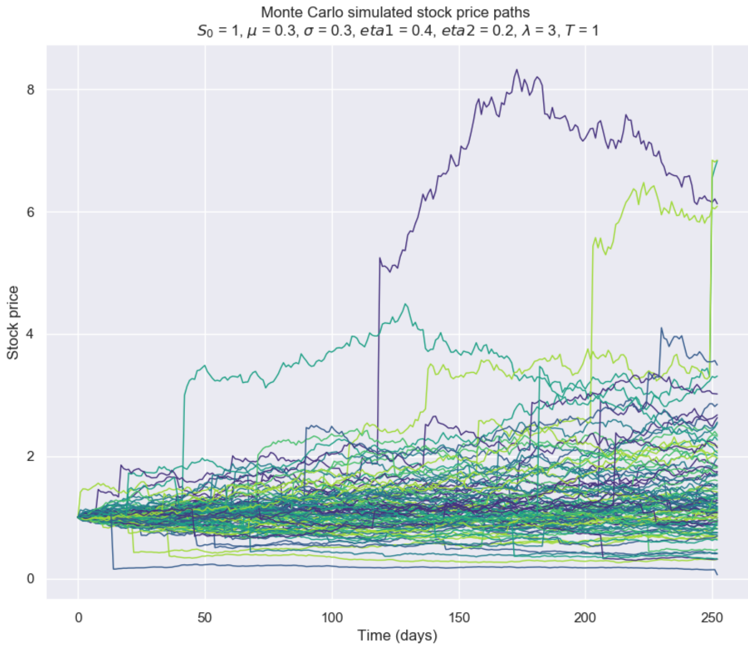

Figure 2: Monte Carlo simulated paths for the Kou model

8.3 Improvement

A jump-diffusion model solves two empirical phenomena we encounter in the behaviour of the stocks in the stock market. These are the leptokurticity of the return distribution and the volatility smile. A lot of research has been done on these two phenomena trying to understand and to correct the already existing models.

8.3.1 Leptokurticity

When looking at the return distribution of stocks or options we see something different from the normal distribution of the Black-Scholes equation based on a Brownian motion. We actually encounter asymmetric leptokurtic features. Leptokurticity is a form of kurtosis, which describes the shape of a probability distribution. "Lepto-" means "slender", and is exactly what we see in the return distribution. This distribution is more skewed to the left, has a higher tail and most importantly heavier/fatter tails than the normal distribution. In figure 3 we plotted with the use of the distribution of the Kou model found in chapter 9, the Kou distribution and normal distribution (with the mean and variance of the Kou model). The parameters are: $t = 1/250$ year, $\sigma = 0.2, \mu = 0.15, \lambda = 10, p = 0.30, \eta_1 = 50, \eta_2 = 25,$

The leptokurtic features are quite evident when looking at the figure. We see a higher peak, more slender distribution, and the tails are heavier, especially the left tail.

8.3.2 Volatility smile

In an ideal world, were the Black-Scholes model is a perfect model for the stock market, volatility will be constant. If you take different options, all at the same point in their life time, with different strikes but with the same expiration time, their volatilities should be the same. What in the real world happens is that the different volatilities form a curve, the volatility "smile". This means that

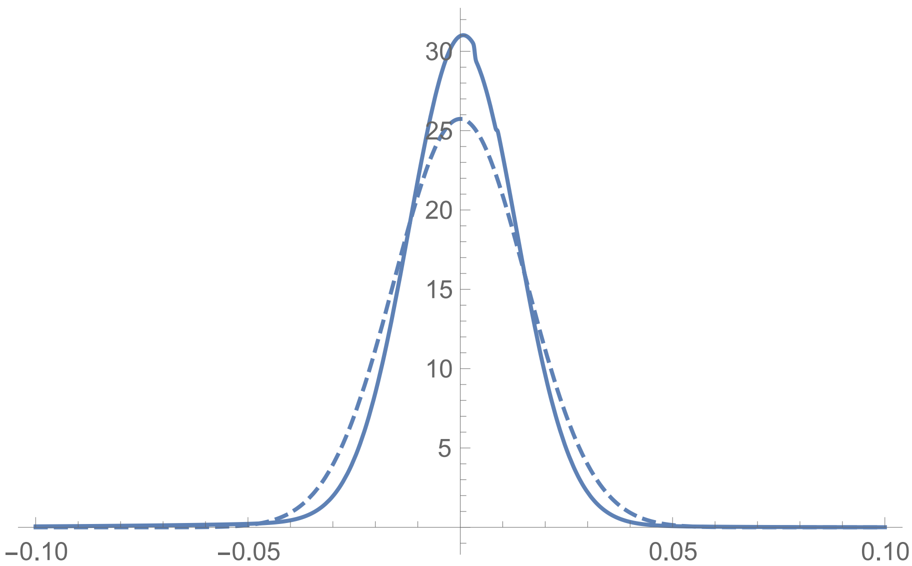

Figure 3: The Kou model distribution (smooth line) and the normal distribution (striped line) plotted with the same mean and variance in *Mathematica*

the implied volatility is a convex function of the strike prices. Due to the lack of different option data we could not produce our own volatility "smile". We show in figure 4 a implied volatility smile generated by a Wolfram Alpha demonstration with the use of the Merton jump diffusion model [15]. We see that, for an increasing strike price, the implied volatility follows a curve, which resembles a smile.

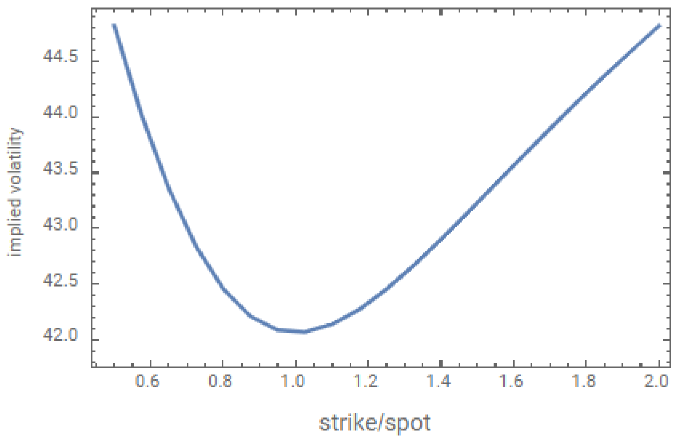

Figure 4: The implied volatility plotted against the strike price divided by the current/spot price.

9 Numerical approach

In this chapter we take a look at the computations and functions needed for eventually using it for our goal of simulating stocks. We need the cumulative density function of the stock using the Kou model. This is actually hard to accomplish because of the fact that there is no closed function of the addition of a normal distributed variable with an asymmetric double exponential distributed variable. Thus we introduce some other functions, especially the Hh-function, defined in equation 70, to make up for this.

We follow the same path that Kou took for obtaining an expression for calculating a call option which follows the Kou model, that we calculate in 9.2. Thus we used two papers of his and combined this to give a good overview but we will not go deep into some derivations or made-up functions. Because that is not very interesting to discuss in this paper, and a link to Kou's paper will be provided, [6] [7].

9.1 How are we going to predict a economic bubble

The big question is now "how are we going to predict an economic bubble?". We first need a lot of data, especially data from a time period which contains an economic bubble. This data is not obtainable on the internet because we need level 1 data. This stands for the best bid/ask option price at any given time. The best bid price is the highest price a buyer is willing to pay for an option whereas the best ask price is the lowest price that a seller is willing to offer for an option. Companies pay around ten to fifteen thousand euros per year for level 1 data, but fortunately I got a small piece of the vast amount of bid/ask prices from my nephew. This could then be used for some calculations and the results will be shown in chapter 10.

All this programming and calculating is done to obtain in the end a parameter called, the illiquidity parameter. We talked about illiquidity in chapter 2, and this parameter represents a sort of destabilization of the stock market. The way this parameter is calculated is by minimizing the "predicted" bid/ask price against the "real" bid/ask price. The "predicted" bid/ask price is calculated by the use of a distortion function.

There are many different distortion functions out there that can be used. Some of them have very different outcomes which are not preferred but there are still a lot to choose from. A question that then arises is, "which one is the best to use?". The mathematicians Madan and Cerny concluded that different distortion functions provide relatively similar results and give good approximations for bid and ask prices [14]. There is still further research needed on this topic, but for our thesis we make use of the *minmaxvar* distortion function. This function is obtained by first using the *maxvar* procedure and then the *minvar* procedure. Without showing a more detailed derivation we will give the *minmaxvar* as,

$$\Psi^{\gamma}(u) = 1 - (1 - u^{\frac{1}{1+\gamma}})^{1+\gamma}, u \in [0,1],\ \gamma \geq 0 \tag{39}$$

For an expression of the bid/ask prices, we first define the expectation value of a random variable with distribution function $\Psi^{\gamma}(F_X)(x)$,

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x d\Psi^{\gamma}(F_X(x)), \tag{40}$$

in other words it is the expected value of X under a new probability measure $\mathbb{P}^{\gamma}(X)$, having density $\psi^{\gamma}(F_X)(x)$ (the pdf), the derivative of $\Psi^{\gamma}(F_X)(x)$ (the cdf) to u, with respected to the original measure $\mathbb{P}$. It is actually the same as the definition of the continuous case for the expectation value, namely,

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f(x) dx = \int_{-\infty}^{\infty} x dF(x) \tag{41}$$

but now we have distorted the cumulative distribution function, and the integral still the Riemann-Stieltjes integral. The bid price is essentially the price that a seller gets for selling an asset to the market. It is the maximum a buyer is willing to pay for it. For bid price b, the buyer is satisfied when $X-b$, where X is the cash flow of the asset, is profitable. We say that $X-b$ is acceptable at the level of γ, therefore the bid price is obtained, by setting the expectation value bigger or equal to zero,

$$\int_{-\infty}^{\infty} xd\Psi^{\gamma}(F_{X-b}(x)) \geq 0 \Longleftrightarrow -b + \int_{-\infty}^{\infty} xd\Psi^{\gamma}(F_X(x)) \geq 0 \tag{42}$$

Thus taking the maximum gives us a function,

$$b_{\gamma}(X) = \int_{-\infty}^{\infty} xd\Psi^{\gamma}(F_X(x)), \tag{43}$$

so we see that the bid price is the distorted expectation of the cash flow X. In our thesis we only look at call options so we calculate only the bid/ask prices for these kind of option. The derivations for the put options go the same way. For a call option we know that the payoff is given by,

$$C_T = max\{(S_T - K), 0\} \tag{44}$$

Inserting this in the expression for the bid price gives us the following integral, where due to the definition of C_T the integration interval is reduced to $[0, \infty]$,

$$\begin{aligned} b_{\gamma}(C) &= \int_{0}^{\infty} xd\Psi^{\gamma}(F_{C_T}(x)) \\ &= \int_{K}^{\infty} (x-K)d\Psi^{\gamma}(F_{S_T}(x)), \end{aligned} \tag{45}$$

using now integration by parts gives us,

$$b_{\gamma}(C) = (x-K)(\Psi^{\gamma}(F_{S_T}(x)) - 1)|_K^{\infty} + \int_{K}^{\infty} (1 - \Psi^{\gamma}(F_{S_T}(x)))d(x-K), \tag{46}$$

here the first term vanishes if we insert the boundaries, so we are left with, the expression for a bid price,

$$b_{\gamma}(C) = \int_{K}^{\infty} (1 - \Psi^{\gamma}(F_{S_T}(x)))d(x-K) \tag{47}$$

We do the same for the ask price, where the ask price is the minimum price a seller is willing to receive, or in other words what a trader needs to pay for it to purchase. The seller always wants profit so $a - X$, where X is the cash flow, needs to be profitable, so we get,

$$\int_{-\infty}^{\infty} xd\Psi^{\gamma}(F_{a-X}(x)) \geq 0 \Longleftrightarrow a + \int_{-\infty}^{\infty} xd\Psi^{\gamma}(F_{-X}(x)) \geq 0 \tag{48}$$

Taking the minimum of this inequality, we are left with,

$$a_{\gamma}(X) = -\int_{-\infty}^{\infty} xd\Psi^{\gamma}(F_{-X}(x)) \tag{49}$$

so the ask price is the negative of the distorted expectation of the cash flow $-X$. It is strange to see that it seems that the strike price is negative and thus less than the bid price, which in practice is always the opposite. Further calculations will show that the negative sign will disappear. We first note that for a call option and $x > 0$,

$$\begin{aligned} F_{-C}(x) &= \mathbb{P}(-(S_T - K) \leq x) = \mathbb{P}((S_T - K) \geq -x) \\ &= \mathbb{P}(S_T \geq K - x) = 1 - \mathbb{P}(S_T \leq K - x) = 1 - F_{S_T}(K - x) \end{aligned} \tag{50}$$

Because the distribution F_X in the formula for the ask price has a negative stochastic process, we integrate from ∞ to 0,

$$\begin{aligned}
a_\gamma(X) &= -\int_{-\infty}^{0} x d\Psi^\gamma(F_{-C_T}(x)) \\
&= -\int_{-\infty}^{0} x d\Psi^\gamma(1 - F_{S_T}(K-x)) \\
&= -\int_{0}^{\infty} x d\Psi^\gamma(1 - F_{S_T}(K+x)) \\
&= -\int_{K}^{\infty} (x-K) d\Psi^\gamma(1 - F_{S_T}(x)) \\
&= -(x-K)(\Psi^\gamma(1 - F_{S_T}(x)))|_K^\infty + \int_K^\infty \Psi^\gamma(1 - F_{S_T}(x)) d(x-K) \\
&= \int_K^\infty \Psi^\gamma(1 - F_{S_T}(x)) dx,
\end{aligned} \tag{51}$$

using again integration by parts and inserting the boundaries will give us,

$$\begin{aligned}
a_\gamma(X) &= -(x-K)(\Psi^\gamma(1 - F_{S_T}(x)))|_K^\infty + \int_K^\infty \Psi^\gamma(1 - F_{S_T}(x)) d(x-K) \\
&= \int_K^\infty \Psi^\gamma(1 - F_{S_T}(x)) dx
\end{aligned} \tag{52}$$

Thus,

$$\begin{aligned}
b_\gamma(C) &= \int_K^\infty (1 - \Psi^\gamma(F_{S_t}(x))) dx \\
a_\gamma(C) &= \int_K^\infty \Psi^\gamma(1 - F_{S_t}(x)) dx
\end{aligned} \tag{53}$$

The goal then is, is to fit these predictions for the bid and ask prices to the real data by minimizing the difference for γ. This is done by minimizing the total squared error,

$$TSE(\gamma) = \sum_{i=1}^{\tau} ((bid_i - b_{i,\gamma})^2 + (ask_i - a_{i,\gamma})^2) \tag{54}$$

Then minimizing this for $\gamma \geq 0$ will give us the illiquidity parameter and tells us in essence how far away the predicted prices from the actual prices lay. When in investors panic and dump all their assets onto the market, the bid and ask prices will drop drastically, increasing γ.

9.2 Kou distribution calculation

The Kou model is hard to calculate. This comes from the fact that it is based on two different distributions which cannot really be united. Fortunately there is a function that can unite these and is called the Hh function, a special function of mathematical physics.
This derivation comes from Kou himself [6] and will be used to find an expression for the distribution function. First we begin with an expression for the calculation of a call option. This is defined as,

$$\Phi(\mu, \sigma, \lambda, p, \eta_1, \eta_2, a, T) := \mathbb{P}[Z(t) \geq a] \tag{55}$$

With $Z(t) = \mu t + \sigma W(t) + \sum_{i=1}^{N(t)} Y_i$.

$$C = e^{-rT}\mathbb{E}^*[(S_T - K)^+] \tag{56}$$

$$C = \mathbb{E}^*[e^{-rT}(S_T - K) * \mathbb{I}_{S_T \geq K}] \tag{57}$$

$$C = \mathbb{E}^*[e^{-rT} S_T * \mathbb{I}_{S_T \geq K}] - Ke^{-rT}\mathbb{P}^*[S_t \geq K] \tag{58}$$

Hereby is $\mathbb{E}^*$ defined with the risk neutral probability measure $\mathbb{P}^*$. That is why the substituting takes place of $\mu = r - \lambda\zeta$

For the calculation of the first term on the right hand side of the equation, we need to make a change of measure. So we introduce a new probability $\widetilde{\mathbb{P}}$,

$$\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}^*} = e^{-rT}\frac{S(T)}{S(0)} = e^{-rT}e^{(r-\frac{\sigma^2}{2}-\lambda\zeta)T+\sigma W(T)+\sum_{i=1}^{N(T)} Y_i} \tag{59}$$

This is well-defined, because $\mathbb{E}^*[e^{-rt}\frac{S(t)}{S(0)}] = 1$ Now we by the theorem of Girsanov for jump processes, $\widetilde{W(t)} := W(t) - \sigma t$ is a new Brownian motion under the measure $\widetilde{\mathbb{P}}$

$$\begin{aligned} X(t) &= (r - \frac{1}{2}\sigma^2 - \lambda\zeta)t + \sigma W(t) + \sum_{i=1}^{N(T)} Y_i \\ &= (r + \frac{1}{2}\sigma^2 - \lambda\zeta)t + \sigma\widetilde{W}(t) + \sum_{i=1}^{N(T)} Y_i \end{aligned} \tag{60}$$

Due to the change of probability measure, some variables changes. First the Poisson process rate becomes: $\widetilde{\lambda} = \lambda\mathbb{E}^*[e_i^Y] = \lambda(1+\zeta)$. The new density of the jump process becomes,

$$\begin{aligned} \frac{1}{\mathbb{E}^*[e^Y]}e^Y f_Y(y) &= \frac{1}{\mathbb{E}^*[e^Y]}e^y p\eta_1 e^{-\eta_1 y}\mathbb{I}_{y\geq 0} + \frac{1}{\mathbb{E}^*[e^Y]}e^y q\eta_2 e^{\eta_2 y}\mathbb{I}_{y<0} \\ &= p\frac{1}{\mathbb{E}^*[e^Y]}\frac{\eta_1}{\eta_1 - 1}(\eta_1 - 1)e^{-(\eta_1-1)y}\mathbb{I}_{y\geq 0} + q\frac{1}{\mathbb{E}^*[e^Y]}\frac{\eta_2}{\eta_2+1}(\eta_2+1)e^{(\eta_2+1)y}\mathbb{I}_{y<0} \end{aligned} \tag{61}$$

For this still to be a double exponential density we must make the substitutions: $\widetilde{\eta_1} = \eta_1 - 1, \widetilde{\eta_2} = \eta_2 + 1$,

$$\widetilde{p} = p[\frac{p\eta_1}{\eta_1 - 1} + \frac{q\eta_2}{\eta 2+1}]^{-1}\frac{\eta_1}{\eta_1 - 1} \tag{62}$$

$$\widetilde{q} = q[\frac{p\eta_1}{\eta_1 - 1} + \frac{q\eta_2}{\eta_2+1}]^{-1}\frac{\eta_2}{\eta_2 + 1} \tag{63}$$

So we had,

$$\mathbb{E}^*[e^{-rT}S(T) * \mathbb{I}_{S(T)\geq K}] \tag{64}$$

$$= S(0)\mathbb{E}^*[e^{-rT}\frac{S(T)}{S(0)} * \mathbb{I}_{S(T)\geq K}] \tag{65}$$

$$= S(0)\widetilde{\mathbb{P}}[S(T) \geq K] \tag{66}$$

$$= S(0)\Phi(r + \frac{1}{2}\sigma^2 - \lambda\zeta, \sigma, \widetilde{\lambda}, \widetilde{p}, \widetilde{\eta_1}, \widetilde{\eta_2}, ln(\frac{K}{S(0)}), T) \tag{67}$$

So we now need to calculate Φ. We first need to decompose the jump variable Y_i, as:

$$\sum_{i=1}^{n} Y_i = \begin{cases} \sum_{i=1}^{k}\xi^+, & \text{with } P_{n,k}, \text{ for } k = 1, 2, \ldots, n \\ -\sum_{i=1}^{k}\xi^-, & \text{with } Q_{n,k}, \text{ for } k = 1, 2, \ldots, n \end{cases}$$

Here ξ^+ and ξ^- are i.i.d. exponential random variables, representing respectively a jump up and jump down, with rate η_1 and η_2, and occur with probabilities $P_{n,k}$ and $Q_{n,k}$. We will give a representation for $P_{n,k}$ and $Q_{n,k}$ without going further into it.

$$P_{n,k} = \sum_{i=k}^{n-1}\binom{n-k-1}{i-k}\binom{n}{i}(\frac{\eta_1}{\eta_1+\eta_2})^{i-k}(\frac{\eta_2}{\eta_1+\eta_2})^{n-i}p^i q^{n-i} \tag{68}$$

$$Q_{n,k} = \sum_{i=k}^{n-1} \binom{n-k-1}{i-k} \binom{n}{i} (\frac{\eta_1}{\eta_1+\eta_2})^{n-i} (\frac{\eta_2}{\eta_1+\eta_2})^{i-k} p^{n-i} q^i \tag{69}$$

We now introduce a new function, taken from the handbook of mathematical function of Abramowitz and Stegun [8]

Definition 9.1 (Hh function) *For every $n \geq 0$, the Hh function is a non-increasing function defined by:*

$$Hh_n(x) = \int_x^\infty Hh_{n-1}(y)dy = \frac{1}{n!} \int_x^\infty (t-x)^n e^{\frac{-t^2}{2}} dt, \quad n = 0, 1, 2, \ldots \tag{70}$$

$$Hh_{-1}(x) = e^{-x^2/2}, \quad Hh_0(x) = \sqrt{2\pi} N(-x) \tag{71}$$

where N(x) is again the cumulative normal distribution function.

An approximation of the Hh function which makes it implementing in python easier and the computation time shorter is given by,

$$Hh_n(x) = 2^{-n/2} \sqrt{\pi} e^{-x^2/2} * (\frac{{}_1F_1(\frac{1}{2}n + \frac{1}{2}, \frac{1}{2}, \frac{1}{2}x^2)}{\sqrt{2}\Gamma(1+\frac{1}{2}n)} - x \frac{{}_1F_1(\frac{1}{2}n + 1, \frac{3}{2}, \frac{1}{2}x^2)}{\Gamma(\frac{1}{2}+\frac{1}{2}n)}) \tag{72}$$

where ${}_1F_1 * (a, b, c)$ is a confluent hypergeometric function, and $\Gamma(x)$ the gamma function, both functions already implemented in a Python library, which makes the coding more pleasant.
We will see when we calculate the probability that it is important to evaluate the integral $I_n(c, \alpha, \beta, \delta)$,

$$I_n(c, \alpha, \beta, \delta) := \int_c^\infty e^{\alpha x} Hh_n(\beta x - \delta) dx, \ n \geq 0, \tag{73}$$

for $\alpha, \beta, c \in \mathbb{R}$

$$I_n(c, \alpha, \beta, \delta) = \begin{cases} -\frac{e^{\alpha c}}{\alpha} \sum_{i=0}^n (\frac{\beta}{\alpha})^{n-i} Hh_i(\beta c - \delta) + (\frac{\beta}{\alpha})^{n+1} \frac{\sqrt{2\pi}}{\beta} e^{\frac{\alpha\delta}{\beta} + \frac{\alpha^2}{2\beta^2}} N(-\beta c + \delta + \frac{\alpha}{\beta}), & \text{if } \beta > 0 \text{ and } \alpha \neq 0 \\ -\frac{e^{\alpha c}}{\alpha} \sum_{i=0}^n (\frac{\beta}{\alpha})^{n-i} Hh_i(\beta c - \delta) - (\frac{\beta}{\alpha})^{n+1} \frac{\sqrt{2\pi}}{\beta} e^{\frac{\alpha\delta}{\beta} + \frac{\alpha^2}{2\beta^2}} N(\beta c - \delta - \frac{\alpha}{\beta}), & \text{if } \beta < 0 \text{ and } \alpha < 0 \end{cases}$$

We now have everything we the calculation of Φ. With the use of $pi_n := \mathbb{P}(N(T) = n) = e^-()^n/n!$ we make a decomposition for $P(Z(T) \geq a)$ with $Z(T) = \mu T + \sigma\sqrt{T}Z + \sum_{i=1}^n Y_i$, with Z a normal random variable with distribution $N(0, \sigma^2)$. This decomposition is then,

$$\begin{aligned} \mathbb{P}(Z(T) \geq a) &= \sum_{n=0}^\infty \pi_n \mathbb{P}(\mu T + \sigma\sqrt{T}Z + \sum_{i=1}^n Y_i \geq a) \\ &= \pi_0 \mathbb{P}(\mu T + \sigma\sqrt{T}Z \geq a) \\ &+ \sum_{n=1}^\infty \pi_n \sum_{k=1}^n P_{n,k} \mathbb{P}(\mu T + \sigma\sqrt{T}Z + \sum_{i=1}^k \xi_i^+ \geq a) \\ &+ \sum_{n=1}^\infty \pi_n \sum_{k=1}^n Q_{n,k} \mathbb{P}(\mu T + \sigma\sqrt{T}Z - \sum_{i=1}^k \xi_i^- \geq a) \end{aligned} \tag{74}$$

First we need to know the probability density function of our variables. Because we have two different variables with different densities which are also independent we can use convolution of functions the calculate the density. We take for the density of $Z + \sum_{i=1}^n \xi^+$,

$$\begin{aligned} f_{Z+\sum_{i=1}^n \xi^+}(t) &= \int_{-\infty}^\infty f_{\sum_{i=1}^n \xi^+}(t-x) f_Z(x) dx \\ &= e^{-t\eta_1} (\eta_1)^n \int_{-\infty}^t \frac{e^{x\eta_1}(t-x)^{n-1}}{(n-1)!} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-x^2}{2\sigma^2}} dx \\ &= e^{-t\eta_1} (\eta_1)^n e^{\frac{(\sigma\eta_1)^2}{2}} \int_{-\infty}^t \frac{(t-x)^{n-1}}{(n-1)!} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\sigma^2\eta_1)^2}{2\sigma^2}} dx \end{aligned} \tag{75}$$

substituting $y = (x - \sigma^2\eta_1)/\sigma$ gives

$$\begin{aligned} f_{Z+\sum_{i=1}^n \xi^+}(t) &= e^{-t\eta_1}(\eta_1)^n e^{\frac{(\sigma\eta_1)^2}{2}} \int_{-\infty}^{t/\sigma-\sigma\eta_1} \frac{(t\sigma - y - \sigma\eta_1)^{n-1}}{(n-1)!} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-y^2}{2\sigma^2}} dy \\ &= \frac{e^{\frac{(\sigma\eta_1)^2}{2}}}{\sqrt{2\pi}} (\sigma^{n-1}\eta_1^n) e^{-t\eta_1} Hh_{n-1}(-t/\sigma + \sigma\eta_1) \end{aligned} \tag{76}$$

The derivation of $f_{Z-\sum_{i=1}^n \xi^-}(t)$ is the same. Our last step is now integrating the obtained probability density function which will give us an expression for $P(Z + \sum_{i=1}^n \xi^+ \geq x)$

$$P(Z+\sum_{i=1}^n \xi^+ \geq x) = \frac{e^{\frac{(\sigma\eta_1)^2}{2}}}{\sqrt{2\pi}} (\sigma^{n-1}\eta_1^n) \int_x^\infty e^{-t\eta_1} Hh_{n-1}(-t/\sigma+\sigma\eta_1) dt = \frac{e^{\frac{(\sigma\eta_1)^2}{2}}}{\sqrt{2\pi}} (\sigma^{n-1}\eta_1^n) I_{n-1}(x, -\eta_1, -1/\sigma, -\sigma\eta_1) \tag{77}$$

And also for this the computation is the same for $P(Z - \sum_{i=1}^n \xi^- \geq x)$
These computations were made for a normal random variable with distribution $N(0, \sigma^2)$, but now if you take into account the drift and volatility, we get essentially a normal random variable with distribution $N(\mu T, \sigma^2 T)$. Having this all in the back of our heads we finally get an expression for $P(Z(T) \geq a)$,with $Z(T)$ given, which we can use for our programming, we have

$$\begin{aligned} P(Z(T) \geq a) = &\frac{e^{\frac{(\sigma\eta_1)^2 T}{2}}}{\sigma\sqrt{2\pi T}} \sum_{n=1}^\infty \pi_n \sum_{k=1}^n P_{n,k}(\sigma\sqrt{T}\eta_1)^k I_{k-1}(a - \mu T, -eta_1, -\frac{1}{\sigma\sqrt{T}}, -\sigma\eta_1\sqrt{T}) \\ &+ \frac{e^{\frac{(\sigma\eta_2)^2 T}{2}}}{\sigma\sqrt{2\pi T}} \sum_{n=1}^\infty \pi_n \sum_{k=1}^n Q_{n,k}(\sigma\sqrt{T}\eta_2)^k I_{k-1}(a - \mu T, eta_2, \frac{1}{\sigma\sqrt{T}}, -\sigma\eta_2\sqrt{T}) \\ &+ \pi_0 N(-\frac{a - \mu T}{\sigma\sqrt{T}}) \end{aligned} \tag{78}$$

So we have the following expression for a call option given the Kou model,

$$\begin{aligned} C(0) = &S(0)\Phi(r + \frac{1}{2}\sigma^2 - \lambda\zeta, \sigma, \widetilde{\lambda}, \widetilde{p}, \widetilde{\eta_1}, \widetilde{\eta_2}, log(\frac{K}{S(0)}), T) \\ &- Ke^{-rT}\Phi(r - \frac{1}{2}\sigma^2 - \lambda\zeta, \sigma, \lambda, p, \eta_1, \eta_2, log(\frac{K}{S(0)}), T) \end{aligned} \tag{79}$$

and for a call option at any time, $0 \leq t < T$,

$$\begin{aligned} C(t) = &S(t)\Phi(r + \frac{1}{2}\sigma^2 - \lambda\zeta, \sigma, \widetilde{\lambda}, \widetilde{p}, \widetilde{\eta_1}, \widetilde{\eta_2}, log(\frac{K}{S(t)}), T - t) \\ &- Ke^{-r(T-t)}\Phi(r - \frac{1}{2}\sigma^2 - \lambda\zeta, \sigma, \lambda, p, \eta_1, \eta_2, log(\frac{K}{S(t)}), T - t) \end{aligned} \tag{80}$$

10 Results

It took some time implementing the Kou model in python, which was my preferred program to use for this thesis. At first I looked at simulating plain stock for a certain drift and volatility, after that I needed data for the more advanced calculations, like the implied volatility. Not knowing that the data is not obtainable on the internet, my nephew who works at a financial company provided me with a chunk of level 1 data. This contained daily bid ask prices of the SP 100, the stock market index which is like the average of the 100 biggest companies with common stock listed on the NASDAQ and NYSE stock market. Although it was not much I could do some programming and get nice results.

10.1 Implied volatility

For the calculation of option prices you need volatility. Because of the fact that volatility changes over time and is not constant, contrary to the assumption in the Black-Scholes model, we need to calculate the implied volatility. This is done by solving the expression of a call option for σ. Normally this is done with the use of the Black-Scholes model but for a better approximation we used the Kou model. Solving the expression given by (80), resulted in the following figure, For

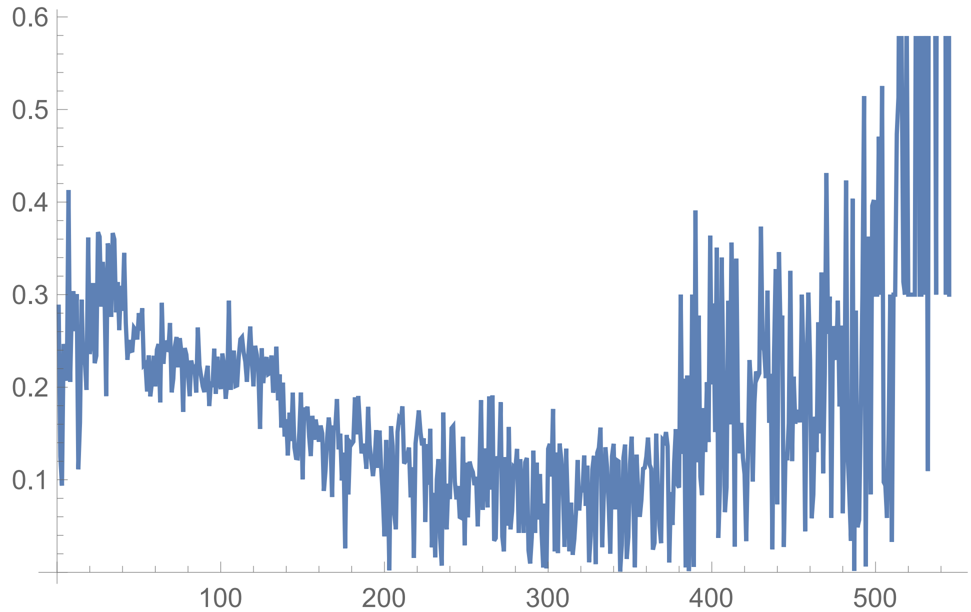

Figure 5: Implied volatility from the Kou model

some reason the Black-Scholes implied volatility gave exactly the same graph as the Kou implied volatility, so we included only one graph. What we see, is a curve in the implied volatility and more erratic for increasing t on the x-axes. This can be explained because of the root finding technique. When t tends more to the right, the time between t and the expiration time, $T - t$, becomes smaller, and gives more room for σ to be a solution for solving (80).

Unfortunately I had not the time or the resources to try to manufacture the illiquidity parameter in Python but the result would have looked something like this, Figures 6 and 7 come from the literature [1] that helped me the most during my thesis and where I got the inspiration from to investigate this matter. What you see in is the two year time elapse from 2008 to 2010, when the American housing bubble took place.

We see some differences between the illiquidity parameters calculated by the two different models.

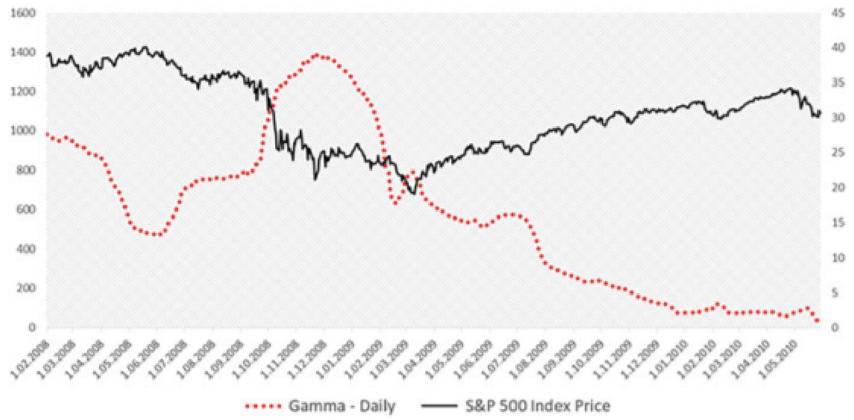

Figure 6: The daily illiquidity parameter with the SP 500 index

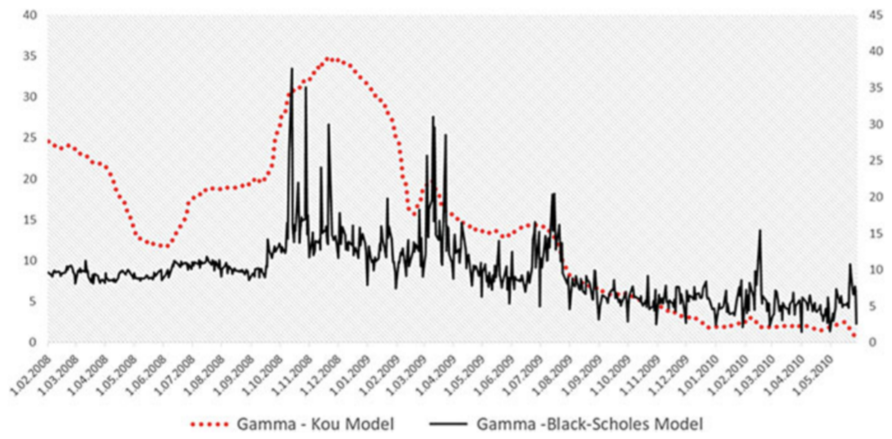

Figure 7: The daily illiquidity parameter for the Kou model and the Black-Scholes model

First we see that the graph is much smoother if you use the Kou model. This comes from the fact that the Kou model does take into account the jumps which are proved to occur in the movement of assets. The Black-Scholes model does not take this into account and therefor has to compensate this which makes it more erratic. The second difference, is that the illiquidity parameter based on Kou starts high whereas for the Black-Scholes stays constant for a period of time. This can be explained due to an incident that happened on the 9th of August 2007. That day BNP Paribas announced that it ceased paying off redemption's on investments funds due to problems in the housing market. This was a moment that made the investors realize that their derivative contracts (like options) were worth a lot less than they had imagined. Hence an increase for Kou's illiquidity parameter, but the Black-Scholes one does not show a sign, which we think comes from the effect that it cannot incorporate sudden changes (jumps) in the market.

After a couple of months the market started to calm down after the news of BNP Paribas, but then the bankruptcy of the Lehman Brothers happened. In June they already reported a second-quarter loss of $2.8billion, and on 15th of September 2008 the Lehman Brothers went bankrupt. This results for both models a big increase in the illiquidity parameter, because everybody lost their faith in the banks. Till that moment the banks were considered "too big to fail", but with the report of the bankruptcy of the Lehman Brothers, panic got the better of investors and everyone wanted to dump their contracts. At the last moment most banks were saved by the government

but the damage was already done.
After the first peak, there are two smaller ones. These come from the financial uncertainty around the big companies General Motors and Chrysler. Due to their actions and the fact that the American and Canadian governments had to give a financial bailout of $85 billion in order to help the organizations to restructure, the uncertainty in the market started to increase again.

10.2 Conclusions and outlook

Figures 6 and 7 are obtained the same way as we tried to accomplish during this thesis, but failed to do so. But we can say with certainty that the illiquidity parameter has some representation of the destabilization or uncertainty of the financial market. The events that led to the housing crisis of 2008 can be found in the parameter, but there is more research needed for saying that it can truly predict a economic bubble.

The difference between the Black-Scholes model and the Kou model is not really noticeable, the implied volatilities look the same, and we notice when we program both models in Python (code included in the appendix), that both are a fairly good approximation for the pricing of options. What we could do in the future is looking at fitting the jump up and jump down parameters with their respective probabilities such that the Kou model becomes a better approximation for the option prices.

There are also other improvements we can accomplish in this research. First, we need to look at other historical economic bubble and check if the same thing happens to the illiquidity parameter as we suspect. Then through time we can say that the method holds. Second, we need more data to compare results. The data we worked with had only 100 options with daily prices. These prices actually change within millisecond, so we also need more computing power. Our code took 22 hour for the calculation of the implied volatility, so finding the illiquidity parameter would take days. A computer that can process big data is essential in this branch.

Our own contribution came mostly from the coding which provided me with more insight in stock market and the Kou model. To try to model the Kou model took more time than imagined and that took the place of the time we could spend on the literature. We did not even come to the calculation of the illiquidity parameter, which in our eyes was the main goal but was in the end a bridge too far. The research for the Kou model and the additional knowledge needed for understanding the background information is too much spread out, because of the fact that it is not a fairly known model. The bid and ask prices calculations are part of a new theory of Conic Finance, not discussed in this thesis, which goes even further into the pricing of bid and ask option prices. We would recommend to look into this to anyone that is interested in this branch of financial mathematics.

11 Acknowledgments

At the end of this thesis I would like to take the time to thank some people who helped me during this period. First of all my bachelor supervisor Karma Dajani for being very patient and supporting. I have not been the easiest person and I have really taken the time the write this thesis but her door was always open if I needed some guidance. I am really grateful for her help.

Furthermore I would like my parents for always being there for me when all the formulas and programming were a bit too much for me, even though this thesis is something complete different from their branch of work.

And lastly my friends Bart Koppelman and Joris de Jong, for being supportive when we were working on our Thesis at the university.

12 Appendix

12.1 Python code for the simulation of call options and implied volatilities

Listing 1: Insert code directly in your document

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from scipy.stats import norm
import math
from mpmath import hyp1f1, gamma

df = pd.read_excel (r'C:\Users\Gebruiker\Documents\Universiteit\Scriptie\data\sp100dataCALL400000.xls
dff = pd.read_excel (r'C:\Users\Gebruiker\Documents\Universiteit\Scriptie\data\^OEXaangepast.xlsx')
ddf = pd.read_excel (r'C:\Users\Gebruiker\Documents\Universiteit\Scriptie\data\heelveeldata.xlsx')
data = df.values
data1 = dff.values
data2 = ddf.values
stock = data1[:,[1,7]]
bidask = data[:,[2,3,4,6]]

x = np.arange(stock[:,1].size, dtype=int)
K = np.full_like(x, 400.0)
s_t = (stock[:,0]/1000000).astype(float)
optionprice = (bidask[:,1]+bidask[:,2])/2
sigma = np.tile(0.001,stock[:,1].size)
rnt = np.tile(0.1/100.0,stock[:,1].size)
T_t = np.tile(0.5,stock[:,1].size)
T_t = (stock[:,1]/365.0).astype(float)

et1 = np.tile(10.0,stock[:,1].size)
et2 = np.tile(5.0,stock[:,1].size)
lab = np.tile(1.0,stock[:,1].size)
pp = np.tile(0.4,stock[:,1].size)

tau = np.log(K/s_t)
P_implied = np.zeros(s_t.size)

for i in range(s_t.size):
    a = 0.0
    b = 1.0
    n = 10

    for k in range(n):
        sigma[i] = (a+b)/2
        d_1 = (math.log(s_t[i]/K[i])+(rente[i]+(sigma[i]**2)/2)*T_t[i])/(sigma[i]*(np.sqrt(T_t[i])))
        d_2 = (math.log(s_t[i]/K[i])+(rente[i]-(sigma[i]**2)/2)*T_t[i])/(sigma[i]*(np.sqrt(T_t[i])))
        P_implied[i] = s_t[i]*norm.cdf(d_1) - K[i]*math.exp(-rente[i]*T_t[i])*norm.cdf(d_2)
        if optionprice[i]-P_implied[i] < 0:
            a = a
            b = (a+b)/2
        else:
            a = (a+b)/2
            b = b
    if sigma[i] > (1-2**(-n+2)):
        sigma[i] = None
    if sigma[i] < 2**(-n+2):
        sigma[i] = None

zeta = pp*eta1/(eta1 - 1) + (1 - pp)*eta2/(eta2 + 1) - 1
temp1 = rente + (sigma**2)/2 - lambda*zeta
temp2 = temp1 - sigma**2
```

```python
vo = 100
def fact(n):
    if n == 0:
        return 1
    else:
        return n * fact(n-1)

factorial = np.zeros(vo).astype(np.object)
for i in range(vo):
    factorial[i] = fact(i)

def nCk(n,k):
    return factorial[n]/(factorial[k]*factorial[n-k])
nck=np.zeros((vo,vo))
for i in range(vo):
    for j in range(i+1):
        nck[i,j] = nCk(i,j)

def Hhfunction(x,n):
    return np.array(2**(-n/2)*np.sqrt(np.pi)*np.e**((-x**2)/2)*(hyp1f1((1/2)*n+(1/2),(1/2),(1/2)*x**

c1 = tau - temp1*T_t
c2 = tau - temp2*T_t
a11 = -(eta1[1]-1)
a12 = (eta2[1]+1)
a21 = -eta1[1]
a22 = eta2[1]

imatrix11 = np.zeros((16,s_t.size))
imatrix12 = np.zeros((16,s_t.size))
imatrix21 = np.zeros((16,s_t.size))
imatrix22 = np.zeros((16,s_t.size))

Hh11function = np.zeros(16).astype(np.object)
Hh12function = np.zeros(16).astype(np.object)
Hh21function = np.zeros(16).astype(np.object)
Hh22function = np.zeros(16).astype(np.object)
Hhl = np.frompyfunc(Hhfunction,2,1)

def Pn1(n,i,p,et1,et2):
    if (i == n and n >= 1):
        return (p)**n
    else:
        po = 0.0
        for j in range(i,n):
            po += (p**(j))*((1-p)**(n-j))*nck[n-i-1,j-i]*nck[n,i]*(et1/(et1+et2))**(j-i) \\
            * (et2/(et1+et2))**(n-j)
        return po

def Qn1(n,i,p,et1,et2):
    if (i == n and n >= 1):
        return (1-p)**n
    else:
        pu = 0.0
        for j in range(i,n):
            pu += (1-p)**j *(p)**(n-j)*nck[n-i-1,j-i]*nck[n,i]*(et2/(et1+et2))**(j-i) \\
            * (et1/(et2+et1))**(n-j)
        return pu

def pin(n,l,T):
    return np.e**(-(l*T))*((l*T)**n)/factorial[n]

pmatrix1 = np.zeros((16,16))
pmatrix2 = np.zeros((16,16))
qmatrix1 = np.zeros((16,16))
qmatrix2 = np.zeros((16,16))
```

```
pip1 = np.zeros(16).astype(np.object)
pip2 = np.zeros(16).astype(np.object)

for n in range(15):
    pip1[n] = pin(n,(lambda*(1+zeta)),T_t)
    pip2[n] = pin(n,lambda,T_t)
    for i in range(16):
        pmatrix1[n,i] = Pn1(n,i,pp[1]*eta1[1]/((1+zeta[1])*(eta1[1]-1)),eta1[1]-1,eta2[1]+1)
        pmatrix2[n,i] = Pn1(n,i,pp[1],eta1[1],eta2[1])
        qmatrix1[n,i] = Qn1(n,i,pp[1]*eta1[1]/((1+zeta[1])*(eta1[1]-1)),eta1[1]-1,eta2[1]+1)
        qmatrix2[n,i] = Qn1(n,i,pp[1],eta1[1],eta2[1])



simga = np.zeros(s_t.size)


b1 = -1/(sigma*np.sqrt(T_t))
b2 = 1/(sigma*np.sqrt(T_t))
g11 = -sigma*np.sqrt(T_t)*(eta1-1)
g12 = -sigma*np.sqrt(T_t)*(eta2+1)
g21 = -sigma*np.sqrt(T_t)*eta1
g22 = -sigma*np.sqrt(T_t)*eta2

for n in range(16):
    Hh11function[n] = Hhl(b1*c1-g11,n)
    Hh12function[n] = Hhl(b2*c1-g12,n)
    Hh21function[n] = Hhl(b1*c2-g21,n)
    Hh22function[n] = Hhl(b2*c2-g22,n)

for x in range(s_t.size):
    for n in range(0,15):
        j11 = 0.0
        j12 = 0.0
        j21 = 0.0
        j22 = 0.0
        for i in range(0,n):
            jo11 = (b1[x]/a11)**(n-1-i)*Hh11function[i][x]
            j11 += jo11
            jo12 = (b2[x]/a12)**(n-1-i)*Hh12function[i][x]
            j12 += jo12
            jo21 = (b1[x]/a21)**(n-1-i)*Hh21function[i][x]
            j21 += jo21
            jo22 = (b2[x]/a22)**(n-1-i)*Hh22function[i][x]
            j22 += jo22
        imatrix11[n,x] = -((np.e**(a11*c1[x]))/a11)*j11 - \\
        ((b1[x]/a11)**(n+1))*(np.sqrt(2*np.pi)/b1[x])*np.e**((a11*g11[x]/b1[x])+(1.0/2.0)*(a11/b1[x])
        imatrix12[n,x] = -((np.e**(a12*c1[x]))/a12)*j12 + \\
        ((b2[x]/a12)**(n+1))*(np.sqrt(2*np.pi)/b2[x])*np.e**((a12*g12[x]/b2[x])+(1.0/2.0)*(a12/b2[x])
        imatrix21[n,x] = -((np.e**(a21*c2[x]))/a21)*j21 - \\
        ((b1[x]/a21)**(n+1))*(np.sqrt(2*np.pi)/b1[x])*np.e**((a21*g21[x]/b1[x])+(1.0/2.0)*(a21/b1[x])
        imatrix22[n,x] = -((np.e**(a22*c2[x]))/a22)*j22 + \\
        ((b2[x]/a22)**(n+1))*(np.sqrt(2*np.pi)/b2[x])*np.e**((a22*g22[x]/b2[x])+(1.0/2.0)*(a22/b2[x])

bt11 = np.e**((sigma*(eta1-1))**2 *(T_t/2))/(sigma*np.sqrt(2*np.pi*T_t))
bt12 = np.e**((sigma*(eta2+1))**2 *(T_t/2))/(sigma*np.sqrt(2*np.pi*T_t))
bt21 = np.e**((sigma*eta1)**2 *(T_t/2))/(sigma*np.sqrt(2*np.pi*T_t))
bt22 = np.e**((sigma*eta2)**2 *(T_t/2))/(sigma*np.sqrt(2*np.pi*T_t))

def cdist1(mu,et1,et2,la,p,sig,aa,T,i):
    ff = 0.0
    gg = 0.0
    for n in range(1,15):
        for k in range(1,n+1):
            ff += bt11[i]*pip1[n][i]*pmatrix1[n,k]*((sig*np.sqrt(T)*et1)**k)*imatrix11[k-1,i]
            gg += bt12[i]*pip1[n][i]*qmatrix1[n,k]*((sig*np.sqrt(T)*et2)**k)*imatrix12[k-1,i]
    return ff + gg + np.e**(-la*T)*norm.cdf(-(aa-mu*T)/(sig*np.sqrt(T)))
```

```
def cdist2(mu,et1,et2,la,p,sig,aa,T,i):
    hh = 0.0
    ii = 0.0
    for n in range(1,15):
        for k in range(1,n+1):
            hh += bt21[i]*pip2[n][i]*pmatrix2[n,k]*((sig*np.sqrt(T)*et1)**k)*imatrix21[k-1,i]
            ii += bt22[i]*pip2[n][i]*qmatrix2[n,k]*((sig*np.sqrt(T)*et2)**k)*imatrix22[k-1,i]
    return hh + ii + np.e**(-la*T)*norm.cdf(-(aa-mu*T)/(sig*np.sqrt(T)))


def callprice(et1,et2,la,p,sig,rr,st,k,t,i):
    return st* \\
    cdist1(temp1[i],et1-1,et2+1,la*(1+zeta[i]),pp[i]*et1/((1+zeta[i])*(et1-1)),sig,np.log(k/st),t,i)
    - k*np.e**(-rr*t)*cdist2(temp2[i],et1,et2,la,pp[i],sig,np.log(k/st),t,i)

callprijs = np.zeros(s_t.size)

for i in range(s_t.size):
    if math.isnan(sigma[i]):
        callprijs[i] = None
    else:
        callprijs[i] = callprice(et1[i],et2[i],lab[i],pp[i],sigma[i],rnt[i],s_t[i],K[i],T_t[i],i)
```

12.2 Mathematical definitions and theorems

Theorem 12.1 (Filtration) *Let Ω be a measurable space with a σ-algebra $\mathcal{F}$. Consider a collection of sub-σ-algebras $\{\mathcal{F}_t\}_{t\in I}$ of F, indexed by $I \subset \mathbb{R}$. (For example, $I = \{0, 1, \cdots, n\}$, $I =$, $I = [0, T]$ and $I = [0, \infty)$. The parameter or index t represents time in general.) If $\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$ for $s, t \in I$ such that $s \leq t$, then $\{\mathcal{F}_t\}_{t\in I}$ is called a filtration. Unless stated otherwise, we assume that 0 is the smallest element in I and $\mathcal{F}_0 = \{\emptyset, \Omega\}$*

Theorem 12.2 (Poisson distribution) *Let X be a random variable. Then X is said to be Poisson distributed when the probability distribution with parameter λ is given by,*

$$P(X = n) = \frac{\lambda^n}{n!} e^{-\lambda} \tag{81}$$

Theorem 12.3 (Poisson point process) *Take the sequence $\{\tau_i\}$ with $i \geq 1$ of independent exponential random variables, all with the mean $\frac{1}{\lambda}$ and let $S_n = \sum_{k=1}^{n} \tau_k$. Then the process,*

$$N(t) = \sum_{n\geq 1} \mathbb{I}_t \geq S_n \tag{82}$$

is called a Poisson process and counts the number of occurrences or "jumps" that occur at or before time t.

Theorem 12.4 (Gamma Density) *For a random variable S_n defined as: $S_n = \sum_{i=1}^{n} x_i$, with x_i a random variable with an exponential distribution, $f(t) = \lambda e^{-\lambda t}$, $t \geq 0$. Then S_n has the gamma density, given by,*

$$g_n(s) = \frac{(\lambda s)^{n-1}}{(n-1)!} \lambda e^{-\lambda s}, \ s \geq 0 \tag{83}$$

Proof
We proof this by induction. We first establish the base case, $n = 0$, we get,

$$g_1(s) = \lambda e^{-\lambda s}, \ s \geq 0 \tag{84}$$

This proves that for $n = 0$ the gamma density is reduced to the exponential distribution. Now we look at $S_{n+1} = S_n + x_{n+1}$. Using convolution because of the fact that S_n and x_{n+1} are independent we get,

$$\begin{aligned}
g_{n+1}(s) &= \int_0^s g_n(y) f_{x_{n+1}}(s-y) dy \\
&= \int_0^s \frac{(\lambda y)^{n-1}}{(n-1)!} \lambda e^{-\lambda y} * \lambda e^{-\lambda(z-y)} \\
&= \frac{\lambda^{n+1} e^{-\lambda s}}{(n-1)!} \int_0^s y^{n-1} dy \\
&= \frac{\lambda^{n+1} e^{-\lambda s}}{(n-1)!} \frac{s^n}{n} = \frac{(\lambda s)^n}{(n)!} \lambda e^{-\lambda s} = g_{n+1}(s)
\end{aligned} \tag{85}$$

Definition 12.5 (Almost surely) *When in probability theory an event is said to happen almost surely, it means that it happens with probability one. In another words, the set of all possible exceptions may be non-empty, but the set has zero probability.*

12.3 Black-Scholes option price

Given the Black-Scholes PDE, we need to find the solution for a call and/or put option

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0 \tag{86}$$

This needs to hold for $S > 0, t \in [0, T)$ with S the stock price, t the time, T expiration time and V the option price.
First we use some transformations, namely: $x = ln(\frac{S}{K}), \tau = T - t \in [0, T]$ and a new function $Z(x, \tau) = V(Ke^x, T - \tau)$, which gives the PDE:

$$\frac{\partial Z}{\partial \tau} - \frac{1}{2}\sigma^2 \frac{\partial^2 Z}{\partial x^2} + (\frac{\sigma^2}{2} - r)\frac{\partial Z}{\partial x} + rZ = 0 \tag{87}$$

Now we need to transform to the heat equation for "easier" calculation. The new function is: $u(x, \tau) = e^{\alpha x + \beta \tau} Z(x, \tau)$, and $\alpha, \beta \in \mathcal{R}$ are chosen so that the PDE for u will become the heat equation. The PDE for u:

$$\frac{\partial u}{\partial \tau} - \frac{\sigma^2}{2}\frac{\partial^2 u}{\partial x^2} + A\frac{\partial u}{\partial x} + Bu = 0 \tag{88}$$

with

$$A = \alpha\sigma^2 + \frac{\sigma^2}{2} - r,\ B = (1 + \alpha)r - \beta - \frac{\alpha^2\sigma^2 + \alpha\sigma^2}{2} \tag{89}$$

For the PDE for u to become the heat equation, we need $A = B = 0$, so we set

$$\alpha = \frac{r}{\sigma^2} - \frac{1}{2}, \beta = \frac{r}{2} + \frac{\sigma^2}{8} + \frac{r^2}{2\sigma^2} \tag{90}$$

So the solution for $u(x, \tau)$ for the PDE is given by the Green formula

$$u(x, \tau) = \frac{1}{\sqrt{2\sigma^2 \pi \tau}} \int_{\infty}^{\infty} e^{-\frac{(x-s)^2}{2\sigma^2 \tau}} u(s, 0) ds. \tag{91}$$

The initial condition for u is:

$$u(x, 0) = e^{\alpha x} V(Ke^x, T) = \begin{cases} e^{\alpha x}(S - K)\ if x > 0 \\ 0\ otherwise \end{cases} \tag{92}$$

Putting functions 91 and 92 together, gives us,

$$u(x, \tau) = \frac{1}{\sqrt{2\sigma^2 \pi \tau}} \int_{\infty}^{\infty} e^{-\frac{(x-s)^2}{2\sigma^2 \tau}} e^{\alpha s}(S - K) ds. \tag{93}$$

We can write this in terms of $N(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-x^2/2} dx$, the cumulative distribution function of a normalized normal distribution. Then with backwards substitutions to $V(S, t)$, we get the following expression,

$$V(S, t) = SN(d_1) - Ke^{-r(T-t)} N(d_2), \tag{94}$$

where N is defined as above, and $d_1 = \frac{ln(S/K) + (r - \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}$ and $d_2 = d_1 - \sigma\sqrt{T - t}$

References

[1] Azar Karimov, *Identifying Stock Market Bubbles: Modeling Illiquidity Premium and Bid-Ask Prices of Financial Securities*, 2017. ISBN 978-3-319-65008-1.

[2] Steven E. Shreve, *Stochastic Calculus for Finance II: Continuous-Time Models*, 2004. ISBN 978-0-387-40101-0.

[3] Geon H. Choe, *Stochastic Analysis for Finance with Simulations*, 2016. ISBN 978-3-319-25589-7.

[4] Rama Cont, Peter Tankov, *Financial Modelling with Jump Processes*, 2004. ISBN 978-1-584-88413-2

[5] Peter Tankov, Ekaterina Voltchkova, *Jump-diffusion models: a practitioner's guide*, Banque et Marchés, No. 99, March-April 2009.

[6] S.G. Kou, *A Jump-Diffusion Model for Option Pricing*, Management Science. Vol. 48, 1086-1101, 2002.

[7] S.G. Kou, Hui Wang, *Option Pricing Under a Double Exponential Jump Diffusion Model*, Management Science. Vol. 50, 1178-1192, 2004.

[8] M. Abramowitz, I. Stegun, *Handbook of Mathematical Functions*, 1964. ISBN 978-1-614-27617-3

[9] Paul Glasserman, *Monte Carlo Methods in Financial Engineering*, 2004. ISBN 978-0-387-00451-8

[10] Ernst Eberlein, *Jump-type Lévy processes* in *Handbook of Financial Time series*, 439-455, ISBN 978-3-540-71296-1

[11] S.G. Kou, Hui Wang, *Advances in Applied Probability*, 2003.

[12] Zoran Ivanovski, Toni Stojanovski, Zoran Narasanov, *Volatility and Kurtosis of Daily Stock Returns at MSE*, UTMS Journal of Economicsm, vol. 6(2), pages 209-221, 2015.

[13] Beáta Stehlíková, *Black-Scholes model: Derivation and Solution*, 2014/2015, Comenius University, Bratislava.

[14] Madan, Cherny, *Illiquid Markets as a Counterparty: An introduction to Conic Finance*, Robert H. Smith School Research Paper No. RHS 06-115, 2010.

[15] Peter Falloon, *Implied Volatility in Merton's Jump Diffusion Model*, https://demonstrations.wolfram.com/ImpliedVolatilityInMertonsJumpDiffusionModel/, 2011.