Option volatility study from a data analysis perspective

*Linghua **Zhang***

Thesis submitted for the

Degree of Master of Philosophy

University of Sussex

June 2016

University of Sussex

Linghua Zhang

Degree of Master of Philosophy

Option volatility study from a data analysis perspective

Summary

In this research, we will investigate both financial option pricing models and link the theory to real market performance studies. By combining traditional option pricing theory and real market data analysis, we propose that, in the real world, some behaviour of the financial option price is strongly associated with the local maximum or minimum of asset price.

Firstly, we analyse some mathematical formulas and theorems to understand how to simulate the random process of asset price movement. Based on these foundations we discuss Black-Scholes option pricing model, stochastic volatility models and numerical methods to price options.

Secondly, we utilise Monte-Carlo simulation to learn about the mechanisms of European option pricing with different models. Subsequently, regression analysis is presented in preparation for studying real market data analysis.

Thirdly, we use nine years of real market data to reveal the relationship among variables involved in pricing European options. It will be concluded that the implied Black-Scholes risk calculated using real world call options and put options correlates with asset prices in opposing ways. For call options: with dominant probability, the instantaneous implied option risk and the asset price have a negative correlation; whereas with dominant probability, one-day earlier implied option risk and the asset price have a positive correlation. Put options are the exact opposite.

Finally, we conclude that when the real market option prices are undervalued, they have the ability to catch local extreme values of asset prices statistically.

Acknowledgements

My greatest thanks must go to my supervisor, Dr. Qi Tang. I would like to thank his supporting, caring, understanding, sharing and guiding during my study at the University of Sussex. He was very kind and patient when I have any questions. His positive attitude towards life and humorous lectures also impressed me a lot. Without his help, this thesis could not have been completed.

Furthermore I would like to express my thanks to my parents, especially my mother. She always supported me whatever I want to do, wherever I want to go. I am so grateful to be her daughter.

In addition, I would like to thank all the people who have worked with me. They helped me to understand and realize more points that I have never thought about. I think my learning ability increased a lot thanks to their help.

Additionally, I wish to express my thanks to the examiners Dr. Bertram During and Dr. Katerina Tsakiri. Thanks their suggestions for my thesis.

Finally, I wish to sincerely thank all my friends. They always cheer me up on my bad days and make me feel positive.

Contents

1 Introduction

In this dissertation, we study real world financial option prices and their corresponding Black-Scholes implied volatilities. It is apparent that these real world implied volatilities, if plotted along the strike price and time to maturity as a surface, the structure becomes coarse and do not follow the theory of standard volatility surface which is assumed to be smooth according to standard text books. It is therefore interesting for us to investigate how these implied volatilities interact with the underlying asset prices.

In the subsections 1.1-1.4 of the introduction, we discuss the standard definition and related properties of financial options. In Section 2, we discuss financial backgrounds and related mathematical tools required for our study. In Section 3, we discuss the Black-Scholes option pricing model and some stochastic volatility models presented in various studies. We also examine the corresponding numerical methods for pricing options. In Section 4, we demonstrate Monte-Carlo simulation of different models and regression analysis. In Section 5, we discuss some properties of the implied volatility of real world option data, establishing links between implied volatility and asset price movements.

1.1 Introduction to financial derivatives

Financial market instruments (Etheridge, 2004) can be divided into two distinct types. The first type consists of those representing a fraction of a real underlying asset: shares (fraction of a company), bonds (a nominal sum of money), commodity contracts (a certain quantity of a particular metal, agricultural product etc.), and foreign currencies. The second type consists of their derivatives, mostly comprising promises to deliver some kind of value in the future dependent on the behaviour of the corresponding underlying assets.

A statement in the cover story of *The Economist* magazine on the 14 May 1994 expertly describes the concept of financial derivatives: financial derivatives are contracts, which give one party a claim on an underlying asset or the cash value of the underlying asset at some point in the future, and bind a counter-party to meet the corresponding liability. The contract might be described by a nominal amount of currency, a number of units of a security, a defined quantity of a physical commodity, a stream of cash payments, or the value of a market index. It might bind both parties equally, or offer one party an option to exercise it or not. It might provide for assets or obligations to be swapped in a predefined formula. It might also be a bespoke derivative combining several elements. Derivatives can be traded either on the stock exchanges or simply over the counter between two or several counter parties; their current market prices usually depend partly on the movement of the prices of the underlying assets after the contracts are created.

From a mathematical perspective, the price of a financial derivative is a function of the underlying asset price as well as a possible number of other variables, such as interest rates, time to maturity, volatility of markets or other factors.

Financial derivatives can be classified into four categories: forwards, futures, options and swaps. In this paper, options will be examined.

Definition 1.1.1: Financial options

A financial option is a contract written on an underlying asset. This contract gives the buyer the right but not the obligation to buy or sell the underlying asset on a specified price, K, before/on specified time, T, and gives the seller the obligation to fulfill the corresponding rights of the buyer.

Option buyers need to pay option premiums to the sellers (writers) to compensate the option sellers' duty to fulfill the obligation when the underlying price moves in the favour of the buyers, while allowing the buyers to abandon the contract should the underlying price move against them.

If an option is written to buy the underlying asset, it is called a *call option*. If an option is written to sell the underlying asset, it is called a *put option*.

The differences between call options and put options are given in Table 1.1.1.

Type	Trade side	Expectation for underlying asset	Premium	Duty	Maximum profit	Maximum loss
Call	Buyer	Increase	Pay	Right, no obligation	Infinite	Premium
	Seller	Decrease	Collect	Obligation, no right	Premium	Infinite
Put	Buyer	Decrease	Pay	Right, no obligation	Infinite	Premium
	Seller	Increase	Collect	Obligation, no right	Premium	Infinite

Table 1.1.1 Differences between call and put options

The difference between a *European option* and an *American option* is that a *European option* written at $t = 0$ can only be exercised at the maturity time $t = T$. An *American option* written at $t = 0$ can be exercised at any time between $t = 0$ and $t = T$.

Hence, a *European call (put) option* gives the buyer the right, but not the obligation to purchase (sell) one unit of the underlying asset at a specified time, *T*, for a specified price, *K*.

In the real financial world, only European options have comprehensive data available because they have well defined OE (option expiring) days. Thus, we focus on European options in this dissertation.

Definition 1.1.2: The exercise date/the maturity

The exercise date or the maturity is the time T at which the option contract expires.

Definition 1.1.3: The strike price

The strike price is the price K on which the option holder has the right to buy or sell.

A call option is defined to be *in the money*, if the spot price is greater than the exercise price *K*, it is *at the money* if the spot price is equal to *K*, and it is *out of the money* if the spot price is less than *K*.

For a put option, it is *in the money* if the stock price is lower than the exercised price *K*, it is *at the money* if the spot price is equal to *K*, and it is *out of the money* if the spot price is greater than *K*.

	Call options	Put options
In-the-money	Strike price<Asset price	Strike price>Asset price
At-the-money	Strike price=Asset price	Strike price=Asset price
Out-of-the-money	Strike price>Asset price	Strike price<Asset price

Table 1.1.2 Classification of options in(out)-of-the-money

1.2 History of option trading

Thompson (2007) stated that the Dutch parliament considered a decree (originally sponsored by the Dutch tulip investors who had lost money because of a German setback during the Thirty Years' War) that changed the way tulip contracts functioned: on 24 February 1637, the self-regulating guild of Dutch florists, in a decision that was later ratified by the Dutch Parliament, announced that all futures contracts written after 30 November 1636 and before the re-opening of the cash market in the early Spring, were to be interpreted as option contracts. They did this by simply relieving the futures buyers of the obligation to buy the future tulips, forcing them merely to compensate the sellers with a small fixed percentage of the contract price.

Before this parliamentary decree, the purchaser of a tulip contract – known in modern finance as a futures contract – was legally obliged to buy the bulbs. The decree changed the nature of these contracts, so that if the current market price fell, the purchaser could opt to pay a penalty and forgo the receipt of the bulbs, rather than pay the full contracted price. This change in law meant that, in modern terminology, the futures contracts had been transformed into options contracts.

Alexander (2008, p. 137) highlights that the first exchange listed options in the world were on the Marche a Prime in France. At that time, about 10 per cent of trading on shares was carried out in this market, where shares were sold accompanied with a three-month at-the-money put option. Due to the existence of this market, in 1900 Louis Bachelier devised a formula to evaluate options premiums based on arithmetic Brownian motion. Subsequently, during the 1930s, gold options were independently traded in Germany. However as they are difficult to value, these options were not popular during the period.

Decades later in 1973, the Chicago Board of Options Exchange (CBOE) was founded and became the first modern, comprehensive marketplace for trading listed options. In the same year, Black and Scholes published a price formula (Black and Scholes, 1973; Merton, 1973, p. 639) revealing how to value a financial option based on geometric Brownian motion. It was the first systematical tool that received the public's approval and is still widely used as a referencing valuation tool today.

In recent years, financial options are traded either over the counter (OTC) or on official stock exchanges.

1.3 The use of financial options

i. Speculation

If an investor believes that a particular share price is going to rise within a period T to a level much higher than K, he/she can buy a call option with exercise price K and expiry date T with the intention to make a profit. For example we suppose K is 25, the share price S_0 today is 25, an option on $K = 25$ and $T = 1$ year costs 1. If the share price at the expiry date S_T goes up to 27, then the investor who buys this option for $K = 25$ and holds it until $t = T$ will make a 100 per cent profit (profit = 2, cost = 1, by ignoring minor factors such as trading costs and interest costs).

ii. Hedging

Suppose that an investor already owns a particular share as a long-term investment and maybe in a situation which is inconvenient to sell (e.g. the holding is large). In this case, the investor may wish to insure against a temporary fall in the share price. Accordingly they can buy a put option to protect financial losses caused by the asset price decreasing. If the underlying asset price decreases, the investor can make a profit from the put option to compensate the loss from holding the underlying asset.

1.4 Introduction to Black-Scholes implied volatility on European options

Option volatility is a measure of the rate and magnitude of the change of underlying prices. Black-Scholes implied volatility is calculated by inverting the Black-Scholes formula (Black and Scholes, 1973) when option price and all other factors are provided.

According to this perspective, suppose that the fixed risk-free interest rate is r, strike price is fixed at K, and maturity is fixed at T, then the market price $f(S,\sigma,t)$ of a standard European call or put option can be calculated from the market price S of the underlying asset using the following formula:

$$f(S,\sigma,t) = \omega N(\omega d_1)S - \omega N(\omega d_2)K\exp(-r(T-t)) \,, \tag{1.4.1}$$

with $d_1 = \frac{1}{\sigma\sqrt{T-t}}\left[\ln\left(\frac{S}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-t)\right]$ and $d_2 = d_1 - \sigma\sqrt{T-t}$, where $=1$ for a call option and $\omega = -1$ for a put option, and $N(\cdot)$ is the cumulative normal distribution density function. The unknown value of σ that satisfies Equation (1.4.1) is the implied volatility. It is straightforward to find this value using MATLAB.

The Black-Scholes model (Black and Scholes, 1973) assumes that the variance rate of the return on the stock is constant for all possible values of strike price and maturity dates. Meanwhile, implied volatility using real world option prices always show different values for different strikes and maturity dates (Chen and Xu, 2013). Rubinstein (1994, p. 776) and Bakshi *et al*. (1997, p. 2022) used real world Standard & Pool's 500 (S&P) option data to calculate implied volatility and confirmed that implied volatility is not constant. They concluded that the implied volatility of S&P 500 options show a 'smile' pattern across the strike price. Thereafter, some researchers explored further this phenomenon (Bates 1996, p. 169, Dumas, Fleming & Whaley 1998, p. 2061) and found that, after 1987, the implied volatility of S&P 500 options was monotone with the moneyness or the strike price, and therefore it exhibited a so-called volatility ‘sneer’ instead of a volatility ‘smile’.

Considering the aforementioned observations, an increasing number of researchers began to investigate implied volatility for financial options determined by option prices. In recent years, researchers have introduced some alternative volatility models which have demonstrated that implied volatility had some mathematical forms other than a constant. The models include, for example, the jump diffusion model devised by Merton (1976, p. 132), the stochastic volatility models (Hull and White 1987, p. 288; Chesney and Scott 1989, p. 268; Stein and Stein 1991, p. 744; Heston 1993, p. 331), and the deterministic local volatility model (Dupire, 1994, p. 128; Derman and Kani 1994; Rubinstein 1994, p. 785).

2 Background and foundation

In this chapter, concepts concerning financial options will be introduced on two fronts — the mathematical theory front and the financial theory front. As will be explained, the value of financial options is a function of the underlying price, volatility of the underlying assets, exercise price, interest rate and time to maturity. In Section 2.1, the necessary mathematical tools that are needed to derive the formula for evaluating values of financial options is introduced. In Section 2.2, financial terminologies and concepts for defining option values are discussed.

2.1 Mathematical background

In this section we introduce some mathematical definitions. Firstly, random walk and Brownian motion or Wiener process, which are used to simulate the movement of the underlying asset price, will be presented. By applying Itô's formula (Itô, 1944) to the random walk model, we deduce the mathematical formula of geometric Brownian motion, which plays a pivotal role in analysing and simulating stock prices.

Definition 2.1.1: Random Walk

A random walk is a mathematical description of a path that consists of a succession of random steps. Feller (1971, p. 24) described it as follows:

Let $X(1), X(2), \ldots, X(N)$ *be independent random variables with values -1 or 1 in equal probability. A random walk is the sequence of random variables*

$$S(0) = 0,\ S(n) = \sum_{i=1}^{n} X(i)\,, n = 1,2,\ldots,N. \qquad (2.1.1)$$

We simulate a random walk for 100 steps in the following graph so that we can understand intuitively what it means.

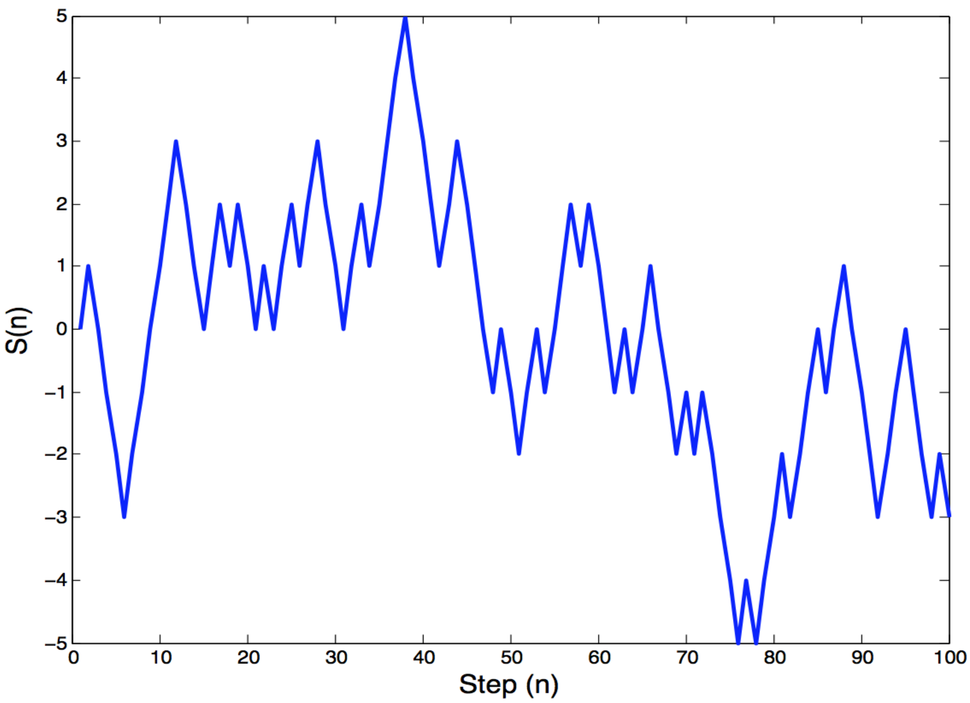

Figure 2.1.1 Random Walk

The steps range from 0 to 100 means n varies from 0 to 100.

Now we introduce the mathematical definitions required to describe random walk in a rigorous setting:

First, we specify a probability triple $(\Omega, \mathcal{F}, \mathbb{P})$, where Ω is a set, called the sample space; $\mathcal{F}$ is a collection of subsets of Ω, called events; and $\mathbb{P}$ specifies the probability of each event $A \in \mathcal{F}$. The events collection $\mathcal{F}$ is a σ −field, that is, $\Omega \in \mathcal{F}$ and $\mathcal{F}$ is closed under the operations of countable union and taking complements. The probability $\mathbb{P}$ must satisfy the usual axioms of probability (Etheridge, 2002):

- $0 \leq \mathbb{P}[A] \leq 1$ for all $A \in \mathcal{F}$,
- $\mathbb{P}[\Omega] = 1$,
- $\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B]$ for any disjoint A, B $\in \mathcal{F}$,
- if $A_n \in \mathcal{F}$ for all $n \in \mathbb{N}$ and $A_1 \subseteq A_2 \subseteq \cdots$ then $\mathbb{P}[A_n] \longrightarrow \mathbb{P}[\cup_n A_n]$ as $n \to \infty$.

A collection of $\{\mathcal{F}_n\}_{n\geq 0}$ where $\mathcal{F}_n \subseteq \mathcal{F}_{n+1} \subseteq \cdots \subseteq \mathcal{F}$ is called a filtration and if a filtration is given, the quadruple $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$ is called a filtered probability space (Etheridge, 2002).

Definition 2.1.2: Random variables (Etheridge, 2002)

A real-valued random variable X is a real-valued function on Ω that is $\mathcal{F}$ −measurable. In the case of a discrete random variable this simply means that for any $x \in R$,

$$\{\omega \in \Omega: X(\omega) = x\} \in \mathcal{F},$$

so that $\mathbb{P}$ assigns a probability to the event $\{X = x\}$. For a general real-valued random variable, we require that for any $x \in R$

$$\{\omega \in \Omega: X(\omega) \leq x\} \in \mathcal{F},$$

so that we can define the distribution function, $F(x) = \mathbb{P}[X \leq x]$.

Definition 2.1.3: Stochastic processes

A real-valued stochastic process is just a sequence of real valued functions, $\{X_n\}_{n\geq 0}$, on Ω. We say that it is adapted to the filtration $\{\mathcal{F}_n\}_{n\geq 0}$ if X_n is $\mathcal{F}_n$ −measurable for each n (Etheridge, 2002).

Definition 2.1.4: Brownian Motion/ Wiener process

A real-valued stochastic process $\{W(t)\}_{t\geq 0}$ is a $\mathbb{P}$ −Brownian motion (or $\mathbb{P}$ −Wiener process) if for some real constant σ, under $\mathbb{P}$ (Etheridge, 2002),

- *for each $s \geq 0$ and $t > 0$ the random variable $W(t+s) - W(s)$ follows the normal distribution with mean zero and variance $\sigma^2 t$,*
- *for each $n \geq 1$ and any time sequence $0 \leq t_0 \leq t_1 \leq \cdots \leq t_n$, the random variables $\{W(t_r) - W(t_{r-1})\}$ are independent,*
- *$W(0) = 0$,*
- *$W(t)$ is continuous in $t \geq 0$, which means $\lim_{s\to t} E\left[\frac{|W(s)-W(t)|}{1+|W(s)-W(t)|}\right] = 0$.*

We simulate a Brownian motion in Figure 2.1.2 for 1000 steps so that we can understand intuitively what it means.

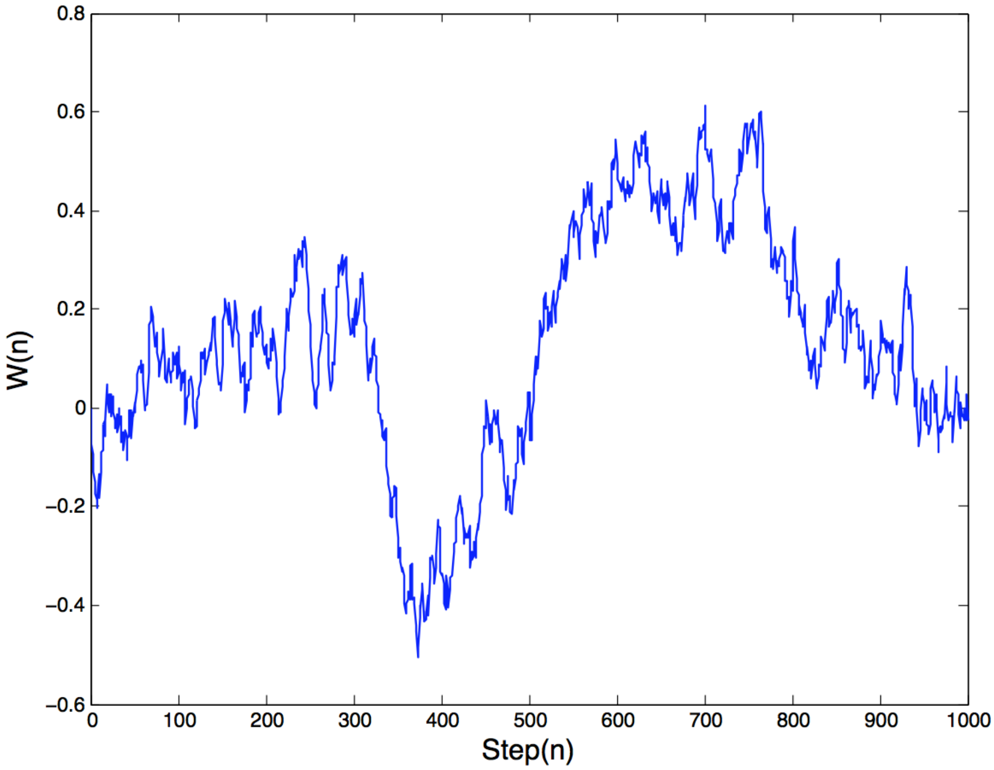

Figure 2.1.2 Brownian motion/Wiener process

Definition 2.1.5: Stochastic differential equation

A typical stochastic differential equation is of the form (Bichteler, 2002)

$$dX(t) = \mu(X(t), t)dt + \sigma(X(t), t)dW(t), \tag{2.1.2}$$

where the function μ is referred to as the drift coefficient; the function σ is called the diffusion coefficient; W denotes a Brownian motion/Winner process.

The integral form of the differential equation (2.1.2) can be written as

$$X(t) - X(0) = \int_0^t \mu(\mathrm{X(s)}, \mathrm{s})ds + \int_0^t \sigma(\mathrm{X(s)}, \mathrm{s})dW(s).$$

Definition 2.1.6: Itô's Formula

If X_t is a stochastic process, satisfying $dX_t = \mu_t dt + \sigma_t dW_t$, and f is a deterministic twice continuously differentiable function, then $Y_t \coloneqq f\,(X_t, t)$ is also a stochastic process and the differential equation of Y_t is given by

$$dY_t = \left(\frac{\partial f}{\partial X}\mu_t + \frac{\partial f}{\partial t} + \frac{1}{2}\sigma_t^2\right)dt + \frac{\partial^2 f}{\partial X^2}\sigma_t dW_t. \tag{2.1.3}$$

Definition 2.1.7: Geometric Brownian motion

Geometric Brownian motion is a continuous-time stochastic process in which the logarithm of the random variable follows a Brownian motion with drift.

A stochastic process $S(t)$ *is said to follow a Geometric Brownian motion if it satisfies the following stochastic differential equation*

$$dS(t) = \mu S(t)dt + \sigma S(t)dW(t), \mathit{for}\, t > 0, \tag{2.1.4}$$

with a constant drift μ, *and a constant volatility* σ, *where* W *is a Wiener process (also called Brownian motion)*.

For an arbitrary initial value S_0 *the above stochastic differential equation has the analytic solution* (Hull, 2009, p. 271)

$$S(t) = S_0 \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W(t)\right). \tag{2.1.5}$$

We simulate two geometric Brownian motions with different drifts and volatilities so that we can understand intuitively what it means. The simulation is displayed as below:

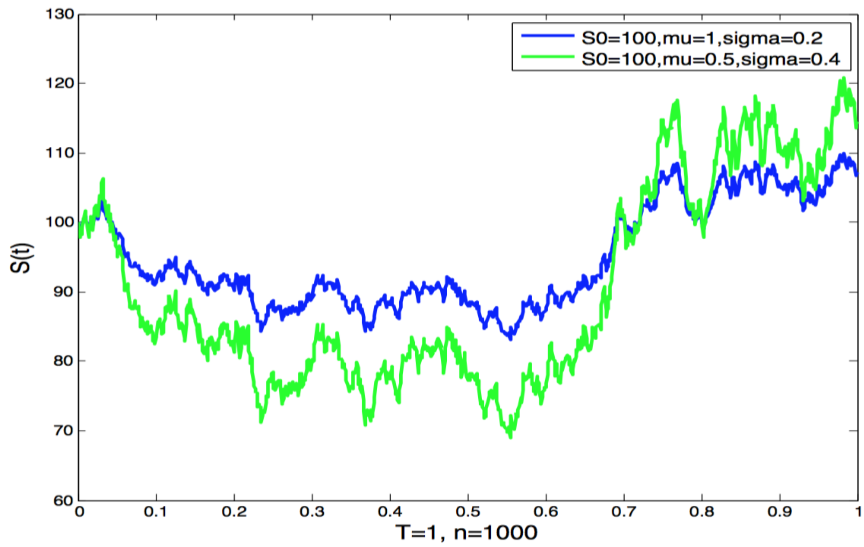

Figure 2.1.3 Geometric Brownian motion

2.2 Financial background

In this section, a brief introduction to some key financial terms is presented in order to make them more meaningful, including no-arbitrage argument for the derivation of option price formula, option contracts, some relationships between call options and put options and the bounds on prices of options.

2.2.1 No-arbitrage argument

Schachermayer(2002) argues that, '*the principle of no-arbitrage formalises a very convincing economic argument: in a financial market it should not be possible to make a profit with zero net investment or without bearing any risk'*.

For the purposes of this study, no-arbitrage is divided into weak no-arbitrage and strong no-arbitrage. They are introduced with the following definitions and examples.

Definition 2.2.1: Weak no arbitrage

In an investment scheme, let p *be the price to pay at time* $t = 0$, C_k *be the payoff at time* $k = 1,2, \dots T$. *Weak no arbitrage assumption means that when* $C_k \geq 0$, *for all* $k \geq 1$, *we must have* $p \geq 0$.

Justification: Suppose $p < 0$.

Since $C_k \geq 0$ for all $k \geq 1$, the buyer receives $-p > 0$ at time $t = 0$, and then does not lose money thereafter. This brings potential profit for no investment (receiving money at the beginning). The seller can increase p as long as $p < 0$, and still have buyers available because the riskless profit opportunity still exists.

Hence p could not be less then zero.

Definition 2.2.2: Strong no arbitrage

In an investment scheme, let p *be the price to pay at time* $t = 0$, C_k *be the payoff at time* $k = 1,2, \dots T$. *Strong no arbitrage assumption means that when* $C_k \geq 0$ *for all* $k \geq 1$ *and* $C_l > 0$ *for some* $l \geq 1$, *we must have* $p > 0$.

Justification: Suppose p ≤ 0.

Since $C_k \geq 0$ for all $k \geq 1$, and for some $l \geq 1$, $C_l > 0$, which means that the buyer makes profit at some time when $C_l > 0$. The buyer receives $-p \geq 0$ at time $t = 0$, and then makes profits thereafter. This brings potential profit for no investment (receiving money at some time $t = l$). The seller can increase p as long as $p \leq 0$, and still have buyers available because the riskless profit opportunity still exists.

Hence p must be greater than zero.

Type A and Type B arbitrage

Definition 2.2.3: *Type A arbitrage is a security or portfolio that produces immediate positive reward at* $t = 0$ *and has non-negative value at* $t = 1$.

Example 2.2.1: Suppose V_t is price of a security at time *t.* The security with initial cost $V_0 < 0$ and at time $t = 1$ value $V_1 \geq 0$ is an example of type A arbitrage.

Definition 2.2.4: *Type B arbitrage is a security or portfolio that has a non-positive initial cost which has positive probability of yielding a positive payoff at* $t = 1$ *and zero probability of producing a negative payoff at* $t = 1$.

Example 2.2.2: Suppose V_t is price of a portfolio at time *t.* The portfolio with initial cost $V_0 \leq 0$, and $V_1 \geq 0$ and $E[V_1] \neq 0$ is an example of type B arbitrage.

2.2.2 Price of options

In this section, we discuss the price of a European option. Theoretically, the option price includes two components: the intrinsic value and the time value.

Intrinsic value

The payoff of a European call option at expiration T depends on the spot price of the underlying asset at $t = T$. If the spot price S_T, is not greater than K, the buyer does not exercise the option, because it is cheaper for him/her to buy in the spot market. The payoff from the option is going to be zero. If the price at $t = T$ is strictly greater than K, the buyer then exercises the option, and the option allows him/her to buy the underlying asset at price K which is cheaper, and the buyer can immediately sell that in the spot market to get S_T. Therefore the payoff from a European call option, at expiration T is going to be $\max(S_T - K, 0)$. The intrinsic value of a call option at some time t, less than expiration T, is simply defined as $\max(S_t - K, 0)$ (Lin, Zheng, Cai & Xiong, 2012).

For a put option, the buyer exercises when the price S_T is less than K, because it allows the buyer to sell at a higher price. The payoff that he/she receives is the difference between the exercise price K and spot price S_T. If S_T is greater than or equal to K, then the buyer does not exercise the option. It is better for him/her to sell in the spot market meaning that the payoff that the buyer gets is *0*. Therefore the payoff from the European put option at expiration T is $\max(K - S_T, 0)$. The intrinsic value of a put option at some point $t \leq T$ is defined as $\max(K - S_t, 0)$ (Lin, Zheng, Cai & Qiu, 2012).

Intrinsic value for call option can be written as:

$$\max(S_t - K, 0) = \begin{cases} S_t - K & \qquad S_t > K, \quad 0 \leq t \leq T \\ & if \\ 0 & \qquad S_t \leq K, \quad 0 \leq t \leq T. \end{cases} \tag{2.2.1}$$

Intrinsic value for put option can be written as:

$$\max(K - S_t, 0) = \begin{cases} 0 & \qquad S_t > K, \quad 0 \leq t \leq T \\ & if \\ K - S_t & \qquad S_t \leq K, \quad 0 \leq t \leq T. \end{cases} \tag{2.2.2}$$

Time value

The market price of a financial option is usually greater than its intrinsic value. The difference between market price and the intrinsic value is called the time value.

The time value is related to the expected value of the underlying asset. For a call option, the higher the probability of spot price at expiry date is greater than the strike price, the higher the time value of the call option has. For put options, it is the exact opposite.

The time value is also related to the length of time until the expiry date. The longer the time remaining until expiration, the higher the time value is. The time value becomes smaller as the expiry date approaches. Finally, at the expiry date, the time value is reduced to zero.

So, before the expiry date, the option price can be written as

Option price = Intrinsic value + Time value.

At the expiry date the option price can be written as

Option price = Intrinsic value.

2.2.3 Put-call parity

Proposition 2.1: European put-call parity at time t for non-dividend paying stock is

$$P(S,t) + S(t) = C(S,t) + Ke^{-r(T-t)},$$

where $P(S,t)$ is the European put option price with strike price K and maturity T; $C(S,t)$ is the corresponding European call option price; $S(t)$ is the underlying share price at time t.

Proof: Construct the following trading strategy as a portfolio:

- Buy one unit of the underlying asset with price $S(t)$ at time t and sell it back at time T with price $S(T)$.
- Buy a European put option with strike K and expiration T, pay option price $P(S,t)$ at time t and keep to maturity.
- Sell a European call option with strike K and expiration T and keep to maturity.
- Borrow cash amount $Ke^{-r(T-t)}$ with continuously compounded interest rate r at time t, repay K at time T.

Portfolio constituents	Cash flow at time t	Cash flow at time $t = T$	
		$S(T) \leq K$	$S(T) \geq K$
Buy a stock	$-S(t)$	$S(T)$	$S(T)$
Buy a put option	$-P(S,t)$	$+\big(K - S(T)\big)$	0
Sell a call oprion	$C(S,t)$	0	$-\big(K - S(T)\big)$
Debit	$Ke^{-r(T-t)}$	$-K$	$-K$
Total	0	0	0

Table 2.2.4 Cash flow of portfolio

It is interesting to observe that:

Cash flow at time T: $\max(S(T) - K, 0) - \max(K - S(T), 0) + S(T) - K = 0.$

According to no-arbitrage argument, cash flow at time *t* should be equal to the cash flow at time T. It gives that cash flow at t equals zero. Hence, we have:

Cash flow at time t: $-S(t) - P(S,t) + C(S,t) + Ke^{-r(T-t)} = 0.$

Thus,

$$P(S,t) + S(t) = C(S,t) + Ke^{-r(T-t)}.$$

2.2.4 Bounds on prices of European options

Proposition 2.2: Upper bound on the price of a European call/put option price at time t are

$$C(S,t) \leq S(t) \text{ for call options,}$$

$$P(S,t) \leq Ke^{-r(T-t)} \text{ for put options.}$$

Proof: Since the European call option gives the buyer the right to buy one share of underlying asset for a certain price, the option can never be worth more than the asset. Hence, we have

$$C(S,t) \leq S(t).$$

For a European put option, the option price at the maturity cannot be worth more than strike price *K*, it means that the European put option price cannot be worth more than the present value of *K* today.

Hence, we have

$$P(S,t) \leq Ke^{-r(T-t)}.$$

Proposition 2.3: Lower bound for a European call/put option price are

$$C(S,t) \geq \max\{S(t) - Ke^{-r(T-t)}, 0\} \text{ for call options,}$$

$$P(S,t) \geq \max\{Ke^{-r(T-t)} - S(t), 0\} \text{ for put options.}$$

Proof: Using put-call parity, we have:

$$C(S,t) = \max\left\{P(S,t) + S(t) - Ke^{-r(T-t)}, 0\right\}$$

$$\geq \max\{S(t) - Ke^{-r(T-t)}, 0\}.$$

$$P(S,t) = \max\left\{C(S,t) + Ke^{-r(T-t)} - S(t), 0\right\}$$

$$\geq \max\{ Ke^{-r(T-t)} - S(t), 0\}.$$

3 Models

Option pricing models are systematic mathematical approaches (closed form formulas, partial differential equation problems or a system of equations with restrictive conditions) that can be used to calculate a theoretical value for financial option contracts. The most widely known work for valuing financial options with a closed mathematical formula is by Black and Scholes (1973) who established the so-called Black-Scholes formula. After the B-S formula, several other approaches were proposed (for example, Garman, 1976; Cox, Ingersoll and Ross, 1985, p. 379) incorporating different features or assumptions on option pricing.

In this chapter we introduce some of these option pricing models. In Section 3.1, we study the Black-Scholes equation and the corresponding closed form solution (Black and Scholes, 1973). In Section 3.2, we consider stochastic volatility and introduce the Hull-White (H-W) model (Hull and White, 1987) and the Heston model (Heston, 1993). Subsequently we present some numerical methods (binomial methods and finite difference methods) for computing option prices.

3.1 Black-Scholes option pricing model

In this section, Black-Scholes differential equation (Black and Scholes, 1973) and the corresponding closed form solutions are presented.

3.1.1 Black-Scholes model

To derive the Black-scholes option value formula (Black and Scholes, 1973), we make the following assumptions:

- The interest rate is a known constant through time.
- The instantaneous log return of the stock price is an infinitesimal random walk with drift; more precisely, it is a geometric Brownian motion, and it is assumed that its drift μ and volatility σ are constants.
- The stock pays no dividends.

- The option is a European option, that is, it can only be exercised at maturity.
- There is no arbitrage opportunity.
- It is possible to borrow and lend any amount, even fractional, of cash at the riskless rate. It is possible to buy and sell any amount, even a fraction of a share of the underlying stock. There are no transaction fees and taxes.

Under these assumptions, the value of the European option depends only on the stock price, strike price and time to maturity.

Assume that *S* is the stock price. Suppose that the stock price follows a geometric Brownian motion, as described by (2.1.4) we have

$$dS(t) = \mu S(t)dt + \sigma S(t)dW(t),$$

where μ is the constant drift of the asset returns, and σ is the constant volatility of the asset returns.

Suppose *V* is the value of the financial option with strike price *K* and maturity *T*, which is a function of the underlying asset price *S* and time *t*, $V(S,t)$. From Ito's formula, we have

$$dV = \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S}dS + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}dt, 0 < S < \infty, 0 < t < T, \qquad (3.1.1)$$

In the following, we follow the argument of Hull (2009, p. 287). Construct a portfolio *F* by buying a unit of *V* and selling (short)   units of *S*

$$F = V(S,t) - \theta S, \qquad (3.1.2)$$

According to Hull (2009, p. 287), we obtain,

$$dF = dV - \theta dS,$$

combine with (3.1.1), it implies

$$dF = \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S}dS + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}dt - \theta dS, 0 < S < \infty, 0 < t < T. \quad (3.1.3)$$

Choose θ, such that *F* is a riskless asset, that is, *F* is independent of *S*, we achieve this by letting

$$\frac{\partial V}{\partial S}dS - \theta S = 0, \tag{3.1.4}$$

in (3.1.3) and obtain

$$\theta = \frac{\partial V}{\partial S}. \tag{3.1.5}$$

$$\Rightarrow dF = \frac{\partial V}{\partial t}dt + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}dt. \tag{3.1.6}$$

In finance, θ defined by (3.1.5) is the so called delta hedging ratio.

Since the new portfolio F is a riskless asset, it should satisfy $dF = rFdt$, combining this with (3.1.6), we obtain

$$\frac{\partial V}{\partial t} + rS\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - rV = 0, 0 < S < \infty, 0 < t < T. \tag{3.1.7}$$

This partial differential equation (3.1.7) is called the Black-Scholes partial differential equation or Black-Scholes equation for short. We make three remarks about this equation (Wilmott, Howison & Dewynne, 1995):

- The Black-Scholes equation (3.1.7) is a linear backward parabolic partial differential equation.
- The delta hedging ratio given by (3.1.5) is the rate of change of the value of the option with respect to the underlying asset price S.
- The Black-Scholes equation (3.1.7) does not contain the growth rate μ of the underlying share.

The next step is to solve the Black-Scholes equation (3.1.7) and obtain solutions for the European call and put options.

3.1.2 Formulas for the European call and put options prices

For European call options, we have the equation

$$\frac{\partial C}{\partial t} + rS\frac{\partial C}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} - rC = 0, 0 < S < \infty, 0 < t < T, \tag{3.1.8}$$

where $C(S,t)$ is the value of the European call option with strike price *K* and maturity *T*. The final condition for Equation (3.1.8) is

$$C(S,T) = \begin{cases} S-K, & if\ S > K, \quad 0 < S < \infty \\ 0, & if\ S \le K, \quad 0 < S < \infty. \end{cases} \tag{3.1.9}$$

And the boundary condition is

$$\begin{cases} C(0,t) = 0, & 0 < t < T \\ C(S,t) \sim S, \ \ as\ S \to \infty, & 0 < t < T. \end{cases} \tag{3.1.10}$$

According to the final condition in (3.1.9) and the boundary condition in (3.1.10), the equation (3.1.8) can be explicitly solved (e.g. Wilmott, Howison & Dewynne, 1995) and the explicit solution is called the B-S pricing formula for European call options:

$$C(S,t) = N(d_1)S - N(d_2)Ke^{-r(T-t)}, \tag{3.1.11}$$

with

$$d_1 = \frac{1}{\sigma\sqrt{T-t}}\left[ln\left(\frac{S}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-t)\right], d_2 = d_1 - \sigma\sqrt{T-t}, \tag{3.1.12}$$

where $N(\cdot)$ is the cumulative distribution function of the standard normal distribution.

For European put options, the Black-Scholes equation (3.1.7) becomes

$$\frac{\partial P}{\partial t} + rS\frac{\partial P}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 P}{\partial S^2} - rP = 0, 0 < S < \infty, 0 < t < T, \tag{3.1.13}$$

where $P(S,t)$ is the value of a European put option with strike price *K* and maturity *T*.

We have the final condition for equation (3.1.13)

$$P(S,T) = \begin{cases} K-S, & if\ S < K, \quad 0 < S < \infty \\ 0, & if\ S \ge K, \quad 0 < S < \infty. \end{cases} \tag{3.1.14}$$

And the boundary conditions

$$\begin{cases} P(0,t) = Ke^{-r(T-t)}, & 0 < t < T \\ P(S,t) \sim 0, \ \ as\ S \to \infty, & 0 < t < T. \end{cases} \tag{3.1.15}$$

Solving (cf. Wilmott, Howison & Dewynne, 1995) the equation in (3.1.13) with the final condition in (3.1.14) and the boundary condition in (3.1.15), we obtain

$$P(S,t) = Ke^{-rT}N(-d_2) - SN(-d_1), \tag{3.1.16}$$

where $d_{1,2}$ are the same as (3.1.12) .

3.2 Stochastic volatility models

Two typical stochastic volatility models will be introduced in this section. The first is Hull-White stochastic model (Hull and White, 1987), whilst the second is the Heston model (Heston, 1993). In this section, we only study the concepts of the models. Because explicit solutions are no longer available as in Black-Scholes case, the Monte-Carlo simulations for both models are presented in Section 4.

3.2.1 Hull-White stochastic model

Hull and White (1987) considered a derivative asset f with a price that depends upon some security price S, time to maturity T, and its instantaneous variance, $V = \sigma^2$, which are assumed to obey the following stochastic processes:

$$dS(t) = \mu S(t)dt + \sigma(t)S(t)dW_1(t), \tag{3.2.1}$$

$$dV(t) = \varphi V(t)dt + \xi V(t)dW_2(t), \tag{3.2.2}$$

The variable μ is dependent on S, and t. The variables φ and ξ depend on σ and t. dW_1 *and* dW_2 *are Brownian motions/ Wiener processes and they are correlated with the correlation coefficient* ρ.

There are several assumptions for Hull-White model:

- The volatility V is uncorrelated with the stock price *S*.
- S, T and σ^2 are the only state variables which affect the price of the derivative security f.
- The risk-free rate, r, must be a constant or at least deterministic.

3.2.2 Heston model

The Heston model (Heston, 1993) assumes that the stock price $S(t)$ *and its instantaneous variance* $V(t)$ *satisfy the following stochastic differential equations (SDEs):*

$$dS(t) = \mu S(t)dt + \sqrt{V(t)}S(t)dW_1(t), \tag{3.2.3}$$

$$dV(t) = -\lambda(V(t) - \theta)dt + \eta\sqrt{V(t)}dW_2(t), \tag{3.2.4}$$

$$\langle dW_1, dW_2 \rangle = \rho dt. \tag{3.2.5}$$

The parameters in the above equations represent the following:

- *μ is the drift coefficient of stock price returns*
- *θ is the long-term mean of price variance*
- *λ is the speed of reversion of $V(t)$ to its long-term mean θ*
- *η is the volatility of volatility*
- *ρ is the correlation between Brownian motions*
- *$dW_1$and dW_2 are Brownian motions/ Wiener processes and they are correlated with the correlation coefficient ρ*

This process (3.2.3) uses the instantaneous variance $V(t)$, which is defined by the theory proposed in Cox, Ingersoll and Ross (1985, p.399). It is usually referred to as CIR process.

There are two assumptions for the Heston model (Heston, 1993):

- The interest rate is constant.
- There is no dividend payment.

3.3 Numerical methods

3.3.1 Binomial methods

Using discrete random walk models, we attempt to emulate the price movement of the underlying asset. Once the movement pattern of the underlying asset is set, we can use this pattern to price the derived option price. If the random walk consists of two possibilities, one is up and the other is down, then the model is a binomial model and the method is classified as binomial method. There are two assumptions (Wilmott, Howison and Dewynne, 1995) underlying the binomial methods:

- Suppose that the lifetime of the option is T, which is divided up into M time-steps of size $\Delta t = T/M$. The continuous random walk can be approximately utilised by a discrete random walk. The asset price S changes only at the discrete times $\Delta t, 2\Delta t, \dots, M\Delta t = T$. We suppose that the asset price is S^m at time $t = m\Delta t$, then the asset price at time $t = (m+1)\,\Delta t$, has two possibilities: moving up to uS with probability p $(0 < p < 1)$ or moving down to dS with probability $1 - p$ $(u > 1 > d > 0)$. The binomial tree is constructed by starting with the given value S, which is the asset price at $t = 0$, generating two possible asset prices at the first time step $t = \Delta t$, three possible values at the second time step $t = 2\Delta t$ until the maturity time of the security. Consequently at time $t = m\Delta t$, there are $m + 1$ possibilities for asset prices.

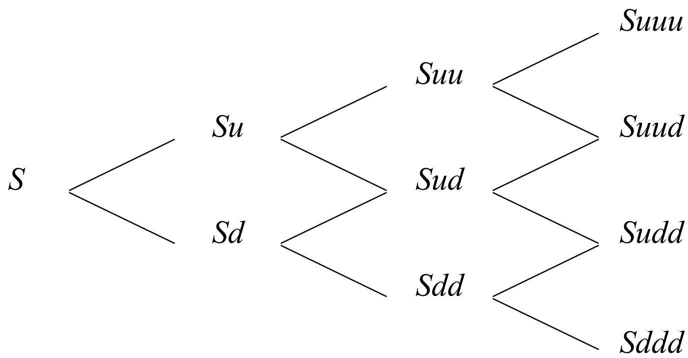

Figure 3.3.1 Binomial tree

- Assume it is a risk-neutral world (Wilmott, Howison & Dewynne, 1995) and thus the stochastic differential equation (2.1.4) is replaced by

$$\frac{dS(t)}{S(t)} = rdt + \sigma dW(t), \tag{3.3.1}$$

where r is the risk-free interest rate.

With these assumptions, we observe that the option value V^m at time $t = m\Delta t$ is the expected value of the option value at time $t = (m+1)\,\Delta t$, discounted by the risk-free interest rate r.

$$V^m = E[e^{-r\Delta t}V^{m+1}]. \tag{3.3.2}$$

Choosing the probability p of asset price moving up and $1 - p$ of asset price moving down, the moving up magnitude u and moving down magnitude d is such that the discrete random walk presented by the binomial tree and the continuous random walk (3.3.1) have the same mean and variance (Wilmott, Howison and Dewynne, 1995).

This means that the expected values and variances of a time-step under the continuous risk-neutral random walk (3.3.1) and the discrete binomial model are equal.

We have the expected value and the variance of S^{m+1}, given S^m, under the continuous random walk (3.3.1):

$$E_C[S^{m+1}|S^m] = \int_0^{\infty} S' p(S^m, m\Delta t; S', (m+1)\Delta t)\, dS' = e^{r\Delta t} S^m, \qquad (3.3.3)$$

$$Var_C[S^{m+1}|S^m] = e^{2r\Delta t}(e^{\sigma^2 \Delta t} - 1)(S^m)^2, \qquad (3.3.4)$$

where $p(S, t; S', t')$ is the probability density function

$$p(S, t; S', t') = \frac{1}{\sigma S' \sqrt{2\pi(t'-t)}} e^{-\left(log(S'/S)-(r-\frac{1}{2}\sigma^2)(t'-t)\right)^2/2\sigma^2(t'-t)}, \qquad (3.3.5)$$

for the risk-neutral random walk (3.3.1) (Wilmott, Howison & Dewynne, 1995).

For the discrete binomial random walk, the expected value of S^{m+1} under S^m is

$$E_B[S^{m+1}|S^m] = (pu + (1-p)d)S^m, \qquad (3.3.6)$$

$$Var_B[S^{m+1}|S^m] = (pu^2 + (1-p)\sigma^2 - e^{2r\Delta t})(S^m)^2. \qquad (3.3.7)$$

Let $E_C[S^{m+1}|S^m] = E_B[S^{m+1}|S^m]$, $Var_C[S^{m+1}|S^m] = Var_B[S^{m+1}|S^m]$, we obtain,

$$pu + (1-p)d = e^{r\Delta t}, \qquad (3.3.8)$$

$$pu^2 + (1-p)\sigma^2 = e^{(2r+\sigma^2)\Delta t}. \qquad (3.3.9)$$

For the three unknown values u, d and p, we have two equations (3.3.8) and (3.3.9). We require three equations to determine three unknown values. Hence, we need another equation. The choice of the third equation is somewhat arbitrary. Two frequently selected options for the third equation are:

$$u = \frac{1}{d}, \qquad (3.3.10)$$

or

$$p = \frac{1}{2}. \qquad (3.3.11)$$

In the case of (3.3.10), the unknown values u, d and p are determined by the equations (3.3.8), (3.3.9) and (3.3.10). We obtain

$$u = A + \sqrt{A^2 - 1}, d = A - \sqrt{A^2 - 1}, p = \frac{e^{r\Delta} - d}{u - d}, \qquad (3.3.12)$$

where $A = \frac{1}{2}\left(e^{-r\Delta t} + e^{(r+\sigma^2)\Delta t}\right)$.

Since $u = \frac{1}{d}$, Sud in Figure 3.3.1 becomes S. It is easy to observe that the binomial tree is vertically symmetrical.

Suppose the asset price at the beginning time is $S = 100$, $u = \frac{1}{d} = \frac{5}{4}$, the binomial tree is shown in Figure 3.3.2.

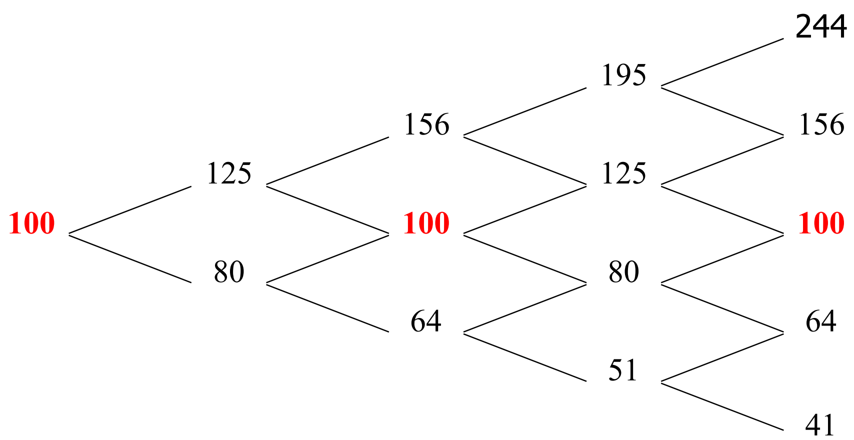

Figure 3.3.2: Binomial tree of underlying asset price when $u = \frac{1}{d}$

In the case of (3.3.11), the unknown values u, d and p are determined by the equations (3.3.8), (3.3.9) and (3.3.11). We obtain

$$u = e^{r\Delta t}\left(1 + \sqrt{e^{\sigma^2\Delta t} - 1}\right), d = e^{r\Delta t}\left(1 - \sqrt{e^{\sigma^2\Delta t} - 1}\right), p = \frac{1}{2}. \qquad (3.3.13)$$

Only if $u \cdot d = 1$, $Sud = S$. In general, the binomial tree will be slightly upwardly adjusted if $u \cdot d > 1$, or downwardly adjusted if $u \cdot d < 1$.

Let $p = 0.5$, $r = 0.2$, $\Delta t = 1$, $\sigma = 0.2$. We have $u = 1.9772$, $d = 0.4656$, $u \cdot d \approx 0.9206 < 1$. With asset price $S = 100$ at the beginning, according to the above discussion about $u \cdot d < 1$, we have a downwardly adjusted binomial tree which is shown in Figure 3.3.3.

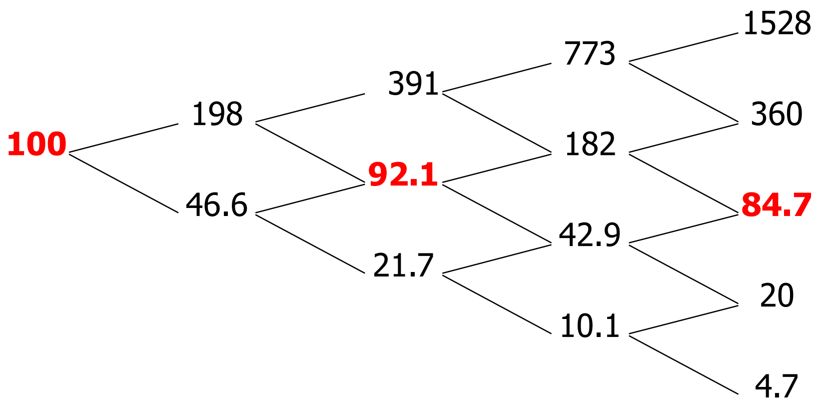

Figure 3.3.3: Binomial tree of underlying asset price when $p = \frac{1}{2}$ and $u \cdot d < 1$.

Let $p = 0.5$, $r = 0.3$, $\Delta t = 1$, $\sigma = 0.2$. We have $u = 2.1851$, $d = 0.5146$, $u \cdot d \approx 1.1244 > 1$. With asset price $S = 100$ at the beginning, according to the above discussion about $u \cdot d > 1$, we have a upwardly adjusted binomial tree which is shown in Figure 3.3.4.

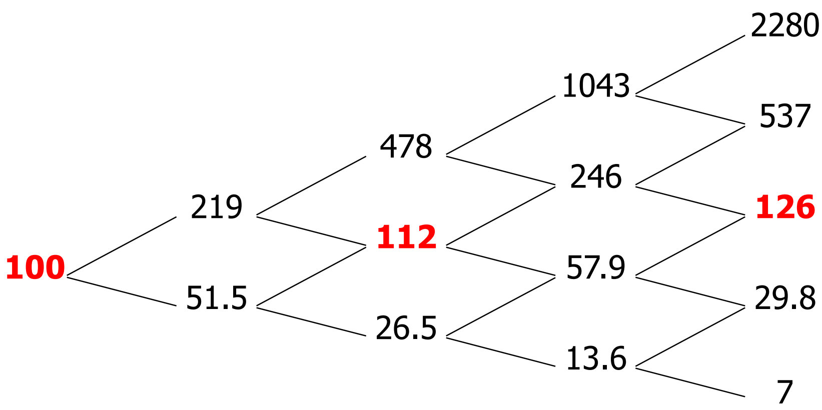

Figure 3.3.4: Binomial tree of underlying asset price when $p = \frac{1}{2}$ and $u \cdot d > 1$.

3.3.2 Pricing European option with binomial methods

Suppose V_n^M is the value of the European put option at $t = T = M\Delta t$, and the payoff function for the option depends only on the values of the underlying asset at maturity. The value of the European put option at the maturity can be priced as:

$$V_n^M = \max(K - S_n^M, 0), \; n = 0, 1, \dots, M, \tag{3.3.14}$$

where V_n^M is the n-th possible value of the European put option at time-step M, K is the strike price and S_n^M denotes the n-th possible value of the underlying asset at time-step M.

Since we have the binomial tree of the underlying asset price (Figure 3.3.1), we can calculate all values of S_n^m, $n = 0, 1, \ldots, m$, $m = 0, 1, \ldots, M$ and the corresponding probability p which represents the probability of S_n^{m-1} moving up to S_{n+1}^m and the probability $1 - p$ which is the probability of S_n^{m-1} moving down to S_n^m.

With the option price and the prices of underlying assets at maturity, we can calculate the expected value of the option at the time-step prior to the maturity $(M-1)\Delta t$ by discounting the values of maturity with the risk-free interest rate r.

$$e^{r\Delta t} V_n^{M-1} = pV_{n+1}^M + (1-p)V_n^M, 0 \le n \le M-1,$$

for the time-step $m\Delta t$, $0 \le m < M$, we have

$$e^{r\Delta t} V_n^m = pV_{n+1}^{m+1} + (1-p)V_n^{m+1}, 0 \le m < M, 0 \le n \le m. \qquad (3.3.15)$$

This gives

$$V_n^m = e^{-r\Delta t}(pV_{n+1}^{m+1} + (1-p)V_n^{m+1}), 0 \le m < M, 0 \le n \le m. \qquad (3.3.16)$$

We can calculate the values of V_n^m for each n and m, at last arriving at the current option price V_0^0.

				S_M^M
				...
		S_2^2	...	S_2^M
	S_1^1	S_1^2	...	S_1^M
S_0^0	S_0^1	S_0^2	...	S_0^M

Figure 3.3.2: The binomial tree of underlying asset price

				V_M^M $max(K - S_M^M, 0)$
				...
		V_2^2 $e^{-r\Delta t}(pV_3^3 + (1 - p)V_2^3)$	...	V_2^M $max(K - S_2^M, 0)$
	V_1^1 $e^{-r\Delta t}(pV_2^2 + (1 - p)V_1^2)$	V_1^2 $e^{-r\Delta t}(pV_2^3 + (1 - p)V_1^3)$	...	V_1^M $max(K - S_1^M, 0)$
V_0^0 $e^{-r\Delta t}(pV_1^1 + (1 - p)V_0^1)$	V_0^1 $e^{-r\Delta t}(pV_1^2 + (1 - p)V_0^2)$	V_0^2 $e^{-r\Delta t}(pV_1^3 + (1 - p)V_{20}^3)$	...	V_0^M $max(K - S_0^M, 0)$

Figure 3.3.3: The binomial tree of European put option price

We simulate the option pricing process using both (3.3.12) and (3.3.13) with different time to maturity and time steps. Thereafter researchers can compare the B-S option price with the price calculated by binomial option pricing methods (Table 3.3.1 and Table 3.3.2).

T	Black-Scholes	Binomial Method ($p = \frac{1}{2}$)				
		M=16	M=32	M=64	M=128	M=256
0.25	48.7578	48.7578	48.7578	48.7578	48.7578	48.7578
0.5	47.5310	47.5310	47.5310	47.5310	47.5310	47.5310
0.75	46.3197	46.3195	46.3195	46.3196	46.3196	46.3196
1	45.1253	45.1241	45.1246	45.1249	45.1252	45.1253

Table 3.3.1 Comparison of Black-Scholes values and binomial method ($p = \frac{1}{2}$) for a European put option with $K = 100, S = 50, r = 0.05,$ and $\sigma = 0.2$. Expiry time T is measured in years.

T	Black-Scholes	Binomial Method ($u = \frac{1}{d}$)				
		M=16	M=32	M=64	M=128	M=256
0.25	48.7578	48.7578	48.7578	48.7578	48.7578	48.7578
0.5	47.5310	47.5310	47.5310	47.5310	47.5310	47.5310
0.75	46.3197	46.3194	46.3195	46.3196	46.3196	46.3196
1	45.1253	45.1236	45.1245	45.1249	45.1251	45.1252

Table 3.3.2 Comparison of Black-Scholes values and binomial method ($u = \frac{1}{d}$) for a European put option with $K = 100, S = 50, r = 0.05,$ and $\sigma = 0.2$. Expiry time T is measured in years.

4 Simulation and regression

In this chapter we present the methods used to analyse the real market data. Boyle (1977, p. 329) and Boyle, Broadie, and Glasserman (1997, p. 1268) utilized Monte-Carlo simulation in derivative pricing and Michael, Fu and Laprise (1999, p. 57) conducted empirical testing to compare the different algorithms. We revisit Monte-Carlo simulation to explore how this method can be used for pricing options in Section 4.1.

In Section 4.1, we apply Monte-Carlo simulation in pricing European options, using a standard random walk model, the Hull-White model and the Heston model for the underlying stock prices. We compare the results using Monte-Carlo simulation with the Black-Scholes price obtained using the Black-Scholes formula. In Section 4.2, the regression method will be introduced.

4.1 Monte-Carlo simulation for pricing European options

Monte-Carlo simulation is one of the mathematical methods for pricing financial derivatives. Here we give three examples to understand the application of Monte-Carlo simulation on pricing European options.

4.1.1 Monte-Carlo simulation using standard Brownian model

Firstly, we use the standard random walk model to simulate the underlying asset price. We assume that the stock has no dividends and the price follows a geometric Brownian motion (cf. (2.1.4)):

$$dS(t) = \mu S(t)dt + \sigma S(t)dW(t), \text{for } t > 0.$$

Here S_t is the stock price at time t, W is a Brownian motion, μ is the assumed constant drift, σ is the assumed constant volatility.

The solution of the geometric Brownian motion is (cf. (2.1.5)):

$$S(t) = S_0 \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W(t)\right).$$

If we rewrite the above equation as a discrete time process：

$$S(t + \Delta t) = S(t)\exp\left(\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma\varepsilon(t)\sqrt{\Delta t}\right), \tag{4.1.1}$$

where $\varepsilon(t)$ is a random value that is normally distributed.

We assume that the underlying asset is a stock and the risk-free interest rate is *r*. Suppose that the lifetime of the option is T, Strike price is K. Then the Monte-Carlo simulation steps for this model are as follows:

a. Divide the lifetime of the option into M time-steps of the size $\Delta t = T/M$. Then simulating the stock price at every discrete times $\Delta t, 2\Delta t, \ldots, M\Delta t = T$ using formula (4.1.1).

b. Calculate the payoff at maturity date. The payoff for European call option and put option at maturity date is max $(S_T - K, 0)$, and max $(K - S_T, 0)$ respectively. Here S_T is the stock price at maturity T simulated in Step a.

c. Repeat steps a and b n times, and obtain n payoff values at the maturity date.

d. Calculate the average payoff value at the maturity date.

e. Discount the payoff value with interest rate r, and take the result as the simulated value of the financial option.

Simulation 4.1.1. Assume that the underlying stock price at time $t = 0$ is $S_0 = 100$; interest rate is $r = 0.02$, $\mu = 0.02$; the European call option has mature time of half a year: $T = 0.5$; Strike Price is $K = 100$; the constant volatility is $\sigma = 0.2$.

First of all, we divide the life-time of the option into 20 time steps, which gives $M = 20$, and $\Delta t = \frac{T}{M} = \frac{0.5}{20}$. Then we simulate the stock price using (4.1.1).

Secondly, we calculate the payoff p of the maturity. $p = \max(S_T - K, 0)$

Thirdly, we repeat the previous two steps for 1000 times, so we have 1000 stock price paths and 1000 payoffs.

Fourthly, we calculate the average value of the payoffs and we name it as a.

At last, the value of the call option is the discounted average payoff value with constant interest rate r: $C = a \cdot e^{-rT}$.

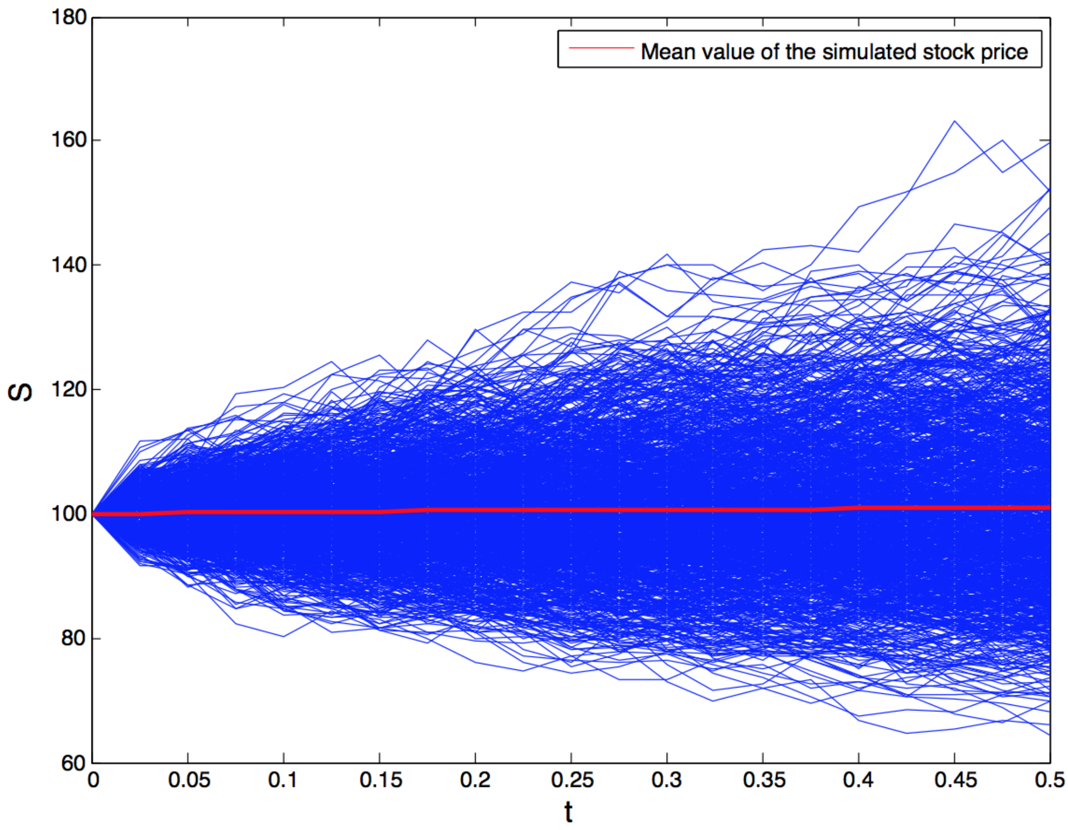

Figure 4.1.1 Simulation of stock price using standard random walk model

The average payoff at $t = T$ as we simulated using MATLAB is 5.9367, and its discounted value is 5.8777 which is the fair value of the option at $t = 0$. The simulation of stock price using standard random walk model is shown in Figure 4.1.1.

4.1.2 Monte Carlo simulation using Hull-White model

As we discussed in Chapter 3, the Hull-White stochastic model on stock price is (cf. (3.2.1), (3.2.2))

$$dS(t) = \mu S(t)dt + \sigma S(t)dW_1(t),$$

$$dV(t) = \varphi V(t)dt + \xi V(t)dW_2(t),$$

where *S(t)* is the stock price, $V(t) = \sigma(t)^2$ is the instantaneous variance of stock price. The parameter μ depends on stock price S, σ and t. The variables φ and ξ depend on σ and t. $dW_1(t)$ and $dW_2(t)$ are Brownian motions/Wiener processes and they are correlated with each other with the correlation coefficient ρ.

Assuming that μ, φ, and ξ are constants, the solutions of the Equations (3.2.1) and (3.2.2) are (Hull, 2009, p. 271)

$$S(t) = S_0 \exp\left(\left(\mu - \frac{\sigma(t)^2}{2}\right)t + \sigma(t)W_1\right), \tag{4.1.2}$$

and

$$V(t) = V_0 \exp\left(\left(\varphi - \frac{\xi^2}{2}\right)t + \xi W_2\right). \tag{4.1.3}$$

If we rewrite above equations as discrete time processes, we have

$$S(t+\Delta t) = S(t)\exp\left[\left(\mu - \frac{\sigma(t)^2}{2}\right)\Delta t + \sigma(t)\varepsilon_1(t)\sqrt{\Delta t}\ \right], \tag{4.1.4}$$

$$V(t+\Delta t) = V(t)\exp\left[\left(\varphi - \frac{\xi^2}{2}\right)\Delta t + \xi\varepsilon_2(t)\sqrt{\Delta t}\ \right], \tag{4.1.5}$$

where $V(t) = \sigma(t)^2$, $\varepsilon_1(t)$ and $\varepsilon_2(t)$ are random values that are normally distributed. And satisfies $\varepsilon_2(t) = \varepsilon_1(t)\rho + \varepsilon(t)\sqrt{1-\rho^2}$, where $\varepsilon(t)$ is a normally distributed random value.

We assume that the underlying asset is a stock and the risk-free interest rate is *r*. Suppose the lifetime of the option is T, Strike price is K. Then the Monte-Carlo simulation steps are as follows:

a. Divide the lifetime of the option into M time-steps $\Delta t = T/M$. Then simulate the stock price at every discrete times $\Delta t, 2\Delta t, \ldots, M\Delta t = T$ using formulas (4.1.4) and (4.1.5).
b. Calculate the payoff at the maturity date. The payoff for European call option and put option at maturity date is max $(S_T - K, 0)$, and max $(K - S_T, 0)$ respectively. Here S_T is the stock price at the maturity T simulated in Step a.
c. Repeat Steps a and b *n* times, and obtain *n* numbers of payoff at the maturity date.
d. Calculate average payoff value at the maturity date.
e. Discount payoff value with interest rate r, and take the result as simulated value of the financial option.

Simulation 4.1.2. Assume that the underlying stock price at time $t = 0$ is $S_0 = 100$; interest rate is $r = 0.02$; the constant drift $\mu = 0.02$; the European call option has half year life-time $T = 0.5$; the strike price K is 100; the other parameters of this model are set as $\varphi = 0.1$, $\xi = 0.1$, $\rho = 0.1$.

First of all, we divide the life-time of the option into 20 time steps, which gives $M = 20$, and $\Delta t = \frac{T}{M} = \frac{0.5}{20}$. Then we simulate the stock price using (4.1.4) and (4.1.5).

Secondly, we calculate the payoff p of the maturity: $p = \max(S_T - K, 0)$.

Thirdly, we repeat the previous two steps for 1000 times, so we have 1000 stock price paths and 1000 payoffs.

Fourthly, we calculate the average value of the payoffs $a = \frac{\sum_{i=1}^{n} p_i}{n}$.

At last, the value of the call option is the discounted average payoff value with interest rate r, $C = a \cdot e^{-rT}$.

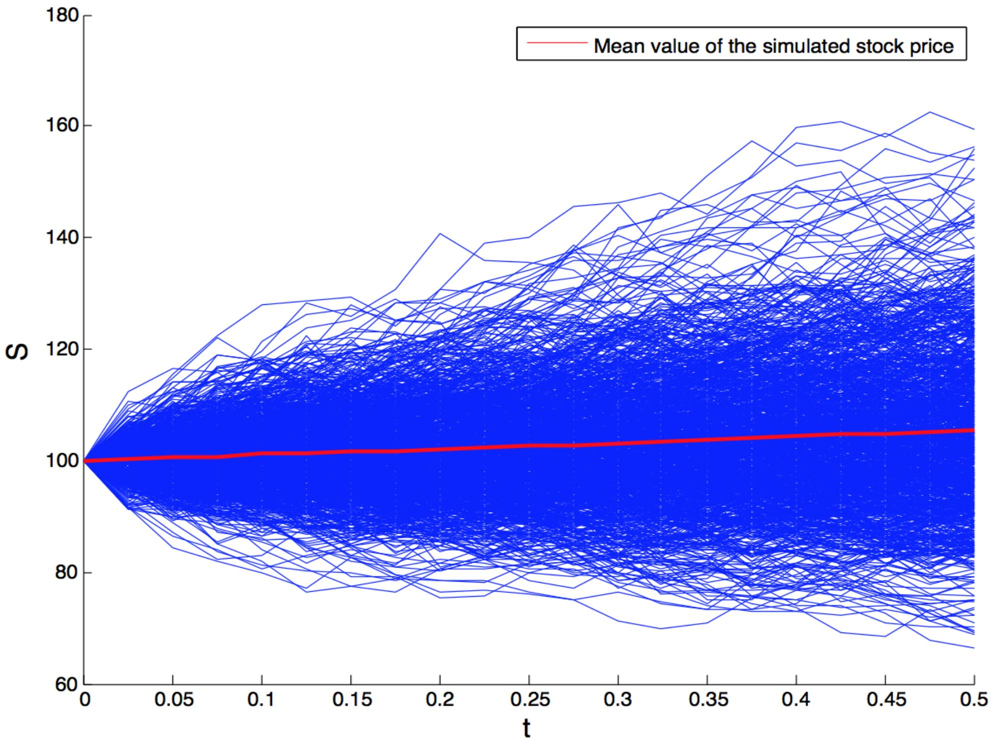

Figure 4.1.2 Simulation of stock price using Hull-White model

The average payoff at $t = T$ as we simulated using MATLAB is 6.3106, and its discounted value is 6.2478 which is the fair value of the option price at $t = 0$. The simulation of stock price using Hull-White model is shown in Figure 4.1.2.

4.1.3 Monte-Carlo simulation using Heston model

As we discussed in Chapter 3, the Heston stochastic model is given by (cf. (3.2.3), (3.2.4)):

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

$$dV(t) = -\lambda(V(t) - \theta)dt + \eta\sqrt{V(t)}dW_2(t),$$

$$\langle dW_1, dW_2\rangle = \rho dt,$$

where $S(t)$ is the stock price, $V(t) = \sigma(t)^2$ is the instantaneous variance of stock price; the parameter μ is the drift coefficient of stock price returns, θ is long-term mean of price variance, λ is the speed of reversion of V to its long-term mean θ, η is the volatility of volatility; dW_1 and dW_2 are Brownian motions/ Wiener processes and they are correlated with the correlation coefficient ρ.

Assuming that μ is a constant, the solution of Equation (3.2.3) is (Hull, 2009, p. 271):

$$S(t) = S_0\exp\left(\left(\mu - \frac{\sigma(t)^2}{2}\right)t + \sigma(t)W_1\right). \tag{4.1.6}$$

The differential equation of the volatility is

$$dV(t) = -\lambda(V(t) - \theta)dt + \eta\sqrt{V(t)}dW_2, \tag{4.1.7}$$

If we rewrite above equations as discrete time processes, we have

$$S(t + \Delta t) = S(t)\exp[\left(\mu - \frac{\sigma(t)^2}{2}\right)\Delta t + \sigma(t)\varepsilon_1(t)\sqrt{\Delta t}, \tag{4.1.8}$$

$$V(t + \Delta t) = \mathrm{V(t)} - \lambda(\mathrm{V(t)} - \theta)\Delta t + \eta\sqrt{V(t)}\varepsilon_2(t)\sqrt{\Delta t}, \tag{4.1.9}$$

where $V(t) = \sigma(t)^2$, $\varepsilon_1(t)$ and $\varepsilon_2(t)$ are random values that are normally distributed. And they satisfy $\varepsilon_2(t) = \varepsilon_1(t)\rho + \varepsilon(t)\sqrt{1-\rho^2}$, where $\varepsilon(t)$ is a normally distributed random value.

We assume that the underlying asset is a stock and the risk-free interest rate is r. Suppose that the lifetime of the option is T; the strike price is K. Then the Monte-Carlo simulation steps are as follows:

a. Divide the lifetime of the option into M time-steps of the size $\Delta t = T/M$. Then simulate the stock price at every discrete time $\Delta t, 2\Delta t, \ldots, M\Delta t = T$ using the formulas (4.1.8) and (4.1.9).
b. Calculate the payoff at maturity date. The payoff for European call option and put option at maturity date is max $(S_T - K, 0)$, and max $(K - S_T, 0)$ respectively. Here S_T is the stock price of maturity T simulated in Step a.
c. Repeat Steps a and b *n* times, and obtain *n* numbers of payoff at the maturity date.
d. Calculate average payoff value at maturity date $a = \frac{\sum_{i=1}^{n} p_i}{n}$.
e. Discount payoff value with interest rate r, and take the result as simulated value of the option price $C = a \cdot e^{-rT}$.

Simulation 4.1.3. Assume that the underlying stock price at time $t = 0$ is $S_0 = 100$; the interest rate is $r = 0.02$, $\mu = 0.02$; the European call option has half year life-time $T = 0.5$; the strike price *K* is 100; and all the other constant parameters are set as 0.1.

First of all, we divide the life-time of the option into 20 time steps, which gives $M = 20$, and $\Delta t = \frac{T}{M} = \frac{0.5}{20}$. Then we simulate the stock price with the formulas in (4.1.8) and (4.1.9).

Secondly, we calculate the payoff at the maturity. $p = \max(S_T - K, 0)$.

Thirdly, we repeat the previous two steps for 1000 times, so that we have 1000 stock price paths and 1000 payoffs.

Fourthly, we calculate the average value of the payoffs.

Finally, the value of the call option is the discounted average payoff value with interest rate r, $C = a \cdot e^{-rT}$.

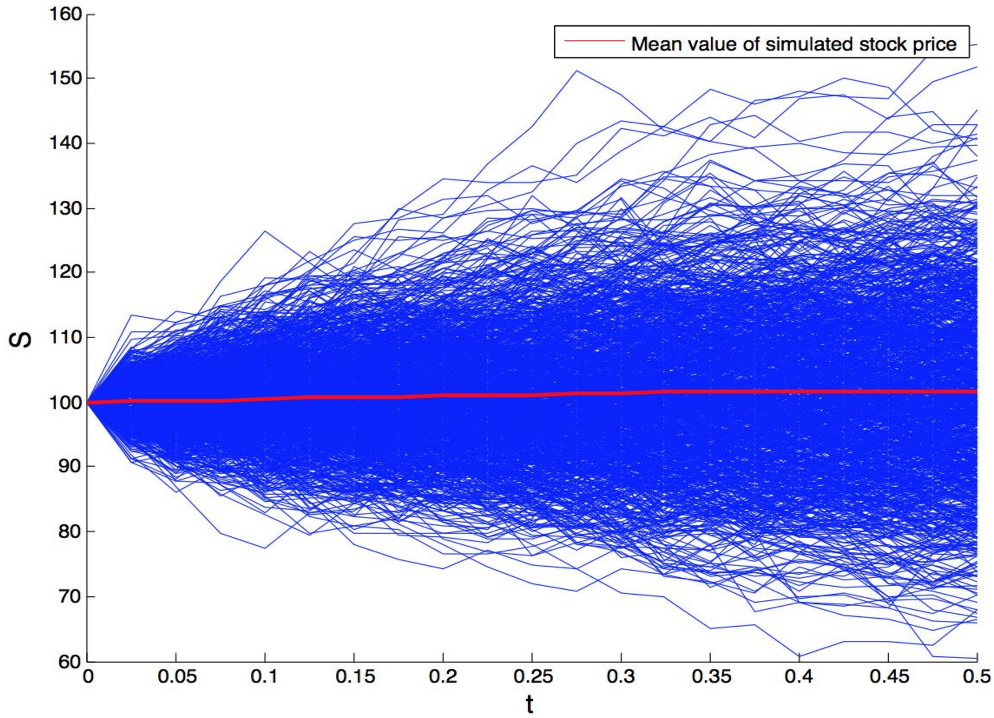

Figure 4.1.3 Simulation of stock price using Heston model

The average payoff at $t=T$ as we simulated in MATLAB is 6.3435, and its discounted value is 6.2803 which is the fair value of the option price at $t = 0$.

Therefore, the option prices for the European call option with the underlying stock price at time $t=0$ $S_0 = 100$, the interest rate $r = 0.02$, $\mu = 0.02$, half year life-time $T = 0.5$, the strike price $K = 100$ calculated by different models are listed in Table 4.1.1.

Model	Option price
B-S model	6.1207
Binomial tree model	6.0541
Standard Brownian model	5.8777
Hull-White model	6.2478
Heston model	6.3435

Table 4.1.1 Comparison of the option price calculated by different models

4.2 Regression

In the real world, the majority of asset pricing information can be expressed by economical variables such as risk, interest rate, price, sales volume, investment amounts. All of these are connected in some way. The most straightforward methods to describe their numerical relationships are regression analysis, correlation analysis and variance analysis. In this section, we mainly focus on regression analysis.

4.2.1 Linear regression

If the value of the dependent variable y is anticipated to have a relationship with the independent variable x or a set of independent variables $(x_1, x_2, \dots, x_n)$, the relationship between y and x or $(x_1, x_2, \dots, x_n)$ can be expressed by a function

$$y = f(x, u),$$

or

$$y = f(x_1, x_2, \dots, x_n, u).$$

where u is the error term.

For a set of different observed values (x_t, y_t) or $(x_{1t,} x_{2t}, \dots, x_{nt}, y_t)$, $t = 1,2, \dots, k$, we have

$$y_t = f(x_t, u_t), t = 1,2, \dots k,$$

or

$$y_t = f(x_{1t}, x_{2t}, \dots, x_{nt}, u_t), t = 1,2, \dots, k.$$

where u_t is the corresponding stochastic error term.

The simplest form of regression is a linear regression with one variable

$$y_t = b_0 + b_1 x_t + u_t.$$

The general linear regression formula with multiple variables are

$$y_t = b_0 + b_{11} x_{1t} + b_{12} x_{2t} + \dots + b_{1n} x_{nt} + u_t.$$

Parameter estimation

For a one-variable linear regression model

$$y_t = b_0 + b_1 x_t + u_t, \tag{4.2.1}$$

we should have $E(u_t) = 0$, which means that the expectation or mean of the random error term is zero.

Under above conditions, we take expectation on both sides of (4.2.1) to get

$$E(y_t) = b_0 + b_1 E(x_t), \tag{4.2.2}$$

with $b_0$and b_1 being regression parameters.

We need to answer the question that how can we estimate the values b_0 and b_1 in (4.2.2) to get the best possible approximation.

The most frequently used method is the Ordinary Least Squares Estimation (OLS).

Definition 4.2.1:Ordinary least squares (OLS)

The ordinary least square estimation method serves to estimate unknown parameters in a linear regression model, with the goal of minimising the differences between the observed responses in the observed response dataset *y* and responses predicted by the linear approximation of the data. In order to let observed regression function to be as close as possible to the real population observation, for each observed dataset (x_t, y_t), we let the difference between the fitted value $\hat{y}_t$ of the regression function $\hat{y}_t = \hat{b}_0 + \hat{b}_1 x_t$ and the observed value y_t, $e_t = y_t - \hat{y}_t$ to be as small as possible. As e_t has both positive and negative values, the sum $\sum e_t$ will not give an estimate of its size. Therefore mathematicians use residual sum of square $\sum e_t^2$ instead of $\sum e_t$. The specific steps of OLS are as follows.

$$\min \sum e_t^2 = \min \sum (y_t - \hat{y}_t)^2 = \min \sum (y_t - \hat{b}_0 - \hat{b}_1 x_t)^2,$$

To minimize $\sum e_t^2$, the undetermined coefficient $\hat{b}_0$ and $\hat{b}_1$ should satisfy

$$\begin{cases} \dfrac{\partial(\sum e_t^2)}{\partial \hat{b}_0} = \sum 2(y_t - \hat{b}_0 - \hat{b}_1 x_t)(-1) = -2\sum (y_t - \hat{b}_0 - \hat{b}_1 x_t) = 0 \\ \dfrac{\partial(\sum e_t^2)}{\partial \hat{b}_1} = \sum 2(y_t - \hat{b}_0 - \hat{b}_1 x_t)(-x_t) = -2\sum (y_t - \hat{b}_0 - \hat{b}_1 x_t) = 0, \end{cases}$$

thus,

$$\begin{cases} \sum y_t = n\hat{b}_0 + \hat{b}_1 \sum x_t \\ \sum x_t y_t = \hat{b}_0 \sum x_t + \hat{b}_1 \sum x_t^2 , \end{cases}$$

where n is the sample size.

This system can be solved as

$$\hat{b}_0 = \bar{y} - \hat{b}_1 \bar{x},$$

$$\hat{b}_1 = \frac{\sum (x_t - \bar{x})(y_t - \bar{y})}{\sum (x_t - \bar{x})^2}.$$

Hence the regression function should be

$$\hat{y}_t = \hat{b}_0 + \hat{b}_1 x_t.$$

Since the residual error is $e_t = y_t - \hat{y}_t$, we have

$$y_t = \hat{b}_0 + \hat{b}_1 x_t + e_t.$$

4.2.2 Measure goodness of fit

Equation $y_t = \hat{b}_0 + \hat{b}_1 x_t + e_t$ can be written as

$$y_t - \bar{y} = \hat{y}_t - \bar{y} + e_t,$$

where $\bar{y}$ is the mean value of the observed data y_t.

So

$$\sum (y_t - \bar{y})^2 = \sum (\hat{y}_t - \bar{y} + e_t)^2 = \sum (\hat{y}_t - \bar{y})^2 + \sum e_t^2 + 2 \sum e_t (\hat{y}_t - \bar{y}) . \quad (4.2.3)$$

Since

$$\sum e_t (\hat{y}_t - \bar{y}) = \sum e_t (\hat{b}_0 + \hat{b}_1 x_t - \bar{y}) = \sum e_t (\hat{b}_0 - \bar{y}) + \hat{b}_1 \sum e_t x_t = 0.$$

Equation (4.2.3) becomes

$$\sum (y_t - \bar{y})^2 = \sum (\hat{y}_t - \bar{y})^2 + \sum e_t^2. \tag{4.2.4}$$

The terms $\sum(y_t - \bar{y})^2, \sum(\hat{y}_t - \bar{y})^2, \sum e_t^2$ have their own explanations as follows:

Total sum of squares (TSS): square value of difference between observed value y_t and mean value y_t

$$TSS = \sum(y_t - \bar{y})^2. \tag{4.2.5}$$

Explained sum of squares (ESS): square value of difference of fitted value $\hat{y}_t$ and mean value $\bar{y}$

$$ESS = \sum(\hat{y}_t - \bar{y})^2. \tag{4.2.6}$$

Residual sum of squares (RSS): square value of errors, which is the square of difference between observed value y_t and the fitted value $\hat{y}_t$:

$$RSS = \sum e_t^2 = \sum(y_t - \hat{y}_t)^2. \tag{4.2.7}$$

(4.2.4) can now be written as

$$TSS = ESS + RSS. \tag{4.2.8}$$

Rewrite (4.2.8), we have

$$1 = \frac{ESS}{TSS} + \frac{RSS}{TSS}. \tag{4.2.9}$$

Now, $\frac{ESS}{TSS} = \frac{\sum(\hat{y}_t - \bar{y})^2}{\sum(y_t - \bar{y})^2}$ is defined as the **Coefficient of determination**: R^2 (R squared)

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS} = 1 - \frac{\sum e_t^2}{\sum(y_t - \bar{y})^2}. \tag{4.2.10}$$

Here, R^2 represents the goodness of fit of the regression. With smaller error $\sum \mathrm{e}_\mathrm{t}^2$, coefficient of determination R^2 has a larger value. It means that when $R^2 \longrightarrow 1, \sum e_t^2 \longrightarrow 0$.

So we have:

- $0 \leq R^2 \leq 1$.
- If $R^2 = 1$, then $\sum e_t^2 = 0$. In this case, the regression is perfectly fitted.
- If $R^2 = 0$, then $\sum e_t^2 = \sum (y_t - \bar{y})^2$. In this case, x and y have no linear relationship with each other.

5 Real life data analysis

In this chapter we analyse option volatility using real market data of SPY options. The discussions concentrate on five aspects: In Section 5.1, we explain the reasons why we choose the SPY option and the necessary steps required to filter and process data. In Section 5.2, we discuss the correlation between implied volatility and asset price. We mainly look at the results on the same time level and on different time levels (data with lags) in the corresponding data. In Section 5.3, we conduct the regression analysis to determine how real world data depends on various factors such as strike price, time to maturity and share price. The main purpose is to explain how the Black-Scholes implied volatility impact on real world data depends on its inputs in a linear pattern, i.e., to understand if volatility is positively or negatively proportional to the variables. In Section 5.4, we study what statistical character the underlying asset prices have when real world options are seriously undervalued (Black-Scholes formula implies that the risk is incalculable, or equivalently, implies negative volatility as discussed in (Gatheral, 2006, p. 21), what statistical character the underlying asset prices has. The results and implications are discussed in Section 5.5.

5.1 Introduction to data used

Choice of study object

We choose the S&P 500 (the Standard & Poor's 500) as our object of study. The S&P 500 is an American stock market index based on the market capitalisations and share prices of the 500 largest companies. SPY is an S&P 500 Exchange Traded Fund (ETF) and is the largest and most popular ETF tracking S&P 500. SPX is the S&P 500 Index. Although both SPY and SPX can be traded as the underlying asset for financial options, we choose SPY and its derived options as our object of study. The reason for this decision was the difference of the trading volume (popularity among traders).

Period	SPX trading volume	SPY trading volume	Volume Ratio SPY/SPX
Pre-Crisis (Year 2005- Bear Stein)	332,599,157	279,524,781	0.84
During Crisis (Bear Stein- end of year 2009)	233,942,487	547,829,314	2.34
After Crisis (Year 2010- Year 2013	527,382,841	1,705,122,316	3.23
Overall	1,093,924,485	2,532,476,411	2.32

Table 5.1.1 Trading volume comparison between SPY & SPX

From Table 5.1.1, it is clear that the trading volume in SPY is much larger when compared to that of SPX options in recent years.

All data concerned, such as asset price, strike price, option price (including last (the last trading price in a trading day), bid and ask), and expiry date come from *www.deltaneutral.com*, whom, according to the corporate website, is the data provider for *The Wall Street Journal*.

Other data such as interest rate (USD 1 year LIBOR), SPY end of day share price, dividend rate were downloaded from the *Thompson Reuters Data Service*.

Data filter

In principle, we used last option price to calculate implied volatility. However, there are a few exceptions (due to the large number of options, many were not traded for many consecutive days). Specifically, we considered the following abnormal situations and made the adjustments described below:

1) If the recorded last price is zero, and the corresponding ask price is less than asset price, last price is replaced by $\frac{bid+ask}{2}$,

2) If the recorded last price is zero, and the corresponding ask price is greater than asset price, the whole option is deleted (options prices should be lower than the corresponding underlying asset prices).

Observation of negative risk

As we discussed in Chapter 2, the upper bound and lower bound of European call options are

$$\max(S_t - Ke^{-r(T-t)}, 0) \le C_t \le S_t.$$

Hence, for in the money options, when the asset price is greater than strike price, the lower bound of the call option price is greater than zero. We drew a picture of option price according to the B-S formula with underlying asset price $S = 100$, strike price $K = 90$, interest rate $r = 0.2$, time to maturity $T = 0.2$, volatility changes from 0 to 10 to view the relationship between volatility and option price:

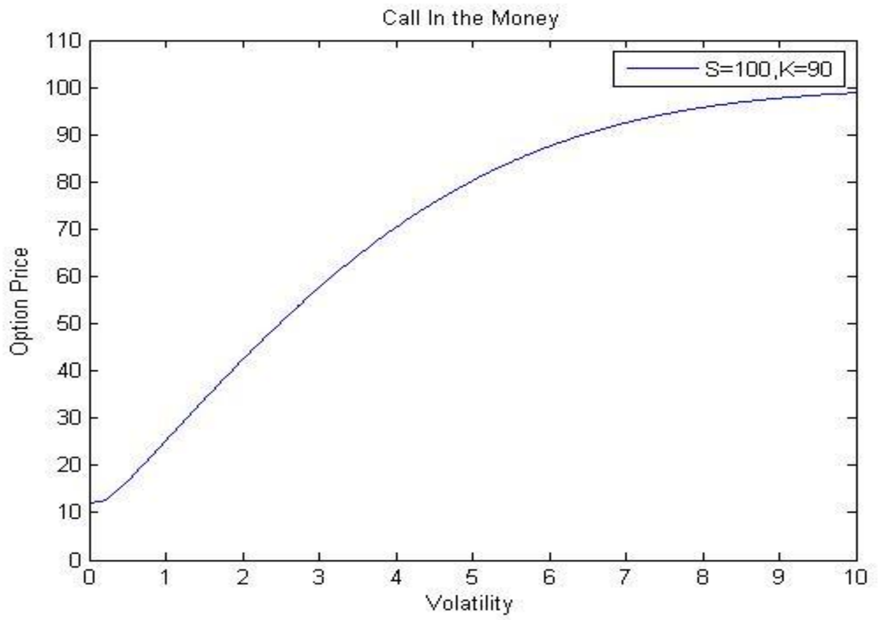

Figure 5.1.1. In the money volatility and option price relationship

From Figure 5.1.1 we see that the B-S option price is bounded form below by the zero risk price, a positive number, and bounded from above by the asset price. However, if we use real world data, it is often observed that the prices will fall below the theoretical 'zero risk value'. In this case, we cannot define the volatility because volatility is a monotonic increasing function of option price. Moreover, in the Black-Scholes formula, volatility is sigma squared which has to be positive. Therefore, we have to use a formula to estimate the volatility for contracts whose option price is less than the theoretical 'zero risk value':

We denote V as the B-S option price, V_0 as zero risk B-S option price, V_C as real world recorded option price, R_V as risk of the option when the option price is V, R_N as negative risk. Then we define

$$V = V_0 + |V_C - V_0|, \tag{5.1.1}$$

$$R_N = -R_V$$

to produce an artificial 'negative risk' for the B-S undervalued options.

The negative risk phenomenon has also been observed and commented by Gatheral (Gatheral, 2006, p. 21) as well as a number of other authors.

For out-of-the-money call options, since strike price is greater than asset price, which implies that $\max(S_t - Ke^{-r(T-t)}, 0) = 0$, the range of option price is from zero to the asset price, so there is no possible negative risk for out-of-the-money call option contracts. We drew a picture with strike price $K = 110$, whilst keeping asset price, interest rate and volatility the same as in Figure 5.1.1. The result is shown in Figure 5.1.2 below:

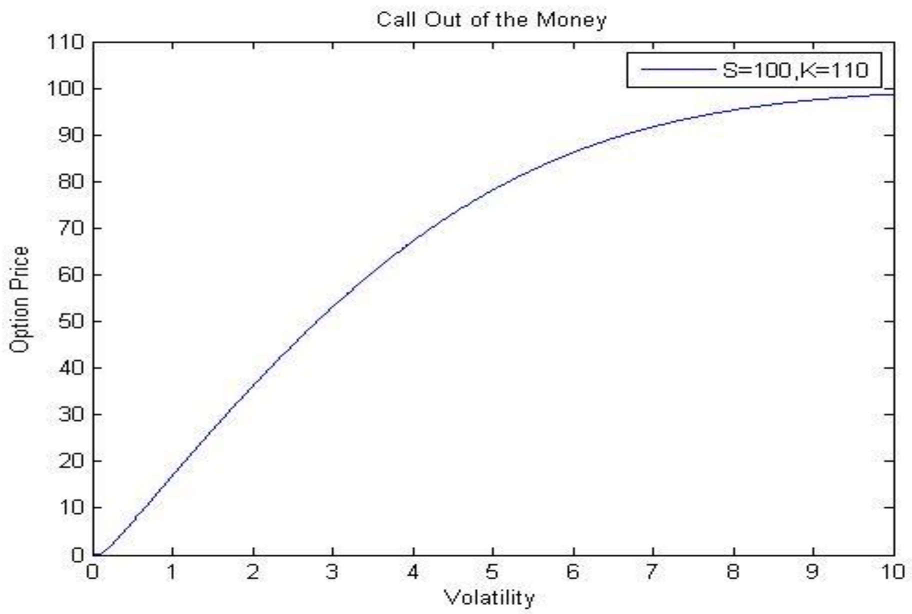

Figure 5.1.2 Out-of-the-money volatility and option price

For put options, there is a similar situation.

5.2 Correlation between implied volatility and asset returns

We arrange date, asset price and average risk in three columns as Table 5.2.1.

Date	Asset Price	Average Risk
10-01-2005	119	0.1822
11-01-2005	118.18	0.2069
12-01-2005	118.57	0.1417
13-01-2005	117.62	0.2294
14-01-2005	118.24	0.1998
18-01-2005	119.47	0.0618
19-01-2005	118.22	0.2812
20-01-2005	117.5	0.3574
21-01-2005	116.78	0.4419
24-01-2005	116.55	0.1888
25-01-2005	116.88	0.1806
26-01-2005	117.23	0.1586
27-01-2005	117.43	0.1473
28-01-2005	117.43	0.1370
31-01-2005	118.16	0.0618
01-02-2005	118.91	-0.0107
...	...	...
23-12-2013	182.53	0.1015
24-12-2013	182.93	0.0998
26-12-2013	183.855	0.0588
27-12-2013	183.845	0.0766
30-12-2013	183.82	0.0809
31-12-2013	184.69	0.0401

Table 5.2.1: Data structure of asset and risk

We use correlation coefficient to obtain the intuitive relationship between implied volatility and the underlying asset returns.

The correlation coefficient r between two vectors x and y is defined as

$$r = \frac{\sum_{t=1}^{n}(x_t - \bar{x})(y_t - \bar{y})}{\sqrt{\sum_{t=1}^{n}(x_t - \bar{x})^2 \cdot \sum_{t=1}^{n}(y_t - \bar{y})^2}}, \tag{5.1.2}$$

where x and y are the variables; n is the size of the variables; x_i is the ith element of variable x; $\bar{x}$ is the mean value of x; y_i is the ith element of variable y; $\bar{y}$ is the mean value of y.

We calculate the frequency percentage for the correlation of option risk and asset price by rolling the observed data over-time. Every observed data has 5 or 10 elements. For example, we let the asset price and the option risk between the dates 10-01-2005 and 14-01-2005 be variables x and y, respectively. We have the variable size $n = 5$. We calculate the correlation coefficient before moving the data forwards one working day, so that asset price and option risk are now between the dates 11-01-2005 and 17-01-2005 (weekend has been excluded) and repeat the same computation as in previous dates and count frequency correspondingly. This is the 'Delay 0' case in Table 5.2.1 below.

For the 'Delay -1' case in Table 5.2.1 below, we select asset prices between the dates 11-01-2005 and 17-01-2005; option risk between 10-01-2005 and 14-01-2005 as variables x and y respectively. We then carry out the same time rolling regression as in the case with 'Delay 0'.

After the process is finished, we calculate the frequency percentage for the sign of correlation coefficient. The results are summarised in Table 5.2.2 below.

Call option	Total frequency percentage for negative correlation		Total frequency percentage for positive correlation	
	Delay 0	Delay -1	Delay 0	Delay -1
2-months 5-day	92.99%	20.51%	7.01%	79.49%
2-months 10-day	99.29%	19.35%	0.71%	80.65%
3-months 5-day	94.19%	18.60%	5.81%	81.40%
3-months 10-day	99.29%	18.28%	0.71%	81.72%
4-months 5-day	94.19%	17.80%	5.81%	82.20%
4-months 10-day	99.51%	17.35%	0.49%	82.65%

Table 5.2.2 The correlation between implied volatility and asset returns for call options

Table 5.2.2 shows that, for call options, the instantaneous asset price has a negative correlation with the instantaneous option risk while the one-day delayed asset price has a positive correlation with the instantaneous option risk.

We conduct the same process for put options, the results of which are shown in Table 5.2.3 below:

Put options	Total frequency percentage for negative correlation		Total frequency percentage for positive correlation	
	Delay 0	Delay -1	Delay 0	Delay -1
2-months 5-day	15.14%	75.72%	84.86%	24.28%
2-months 10-day	7.25%	72.38%	92.75%	27.62%
3-months 5-day	13.49%	76.43%	86.51%	23.57%
3-months 10-day	6.14%	73.31%	93.86%	26.69%
4-months 5-day	13.23%	76.83%	86.77%	23.17%
4-months 10-day	5.96%	73.89%	94.04%	26.11%

Table 5.2.3 The correlation between implied volatility and asset returns for put options

Table 5.2.3 shows that, for put options, the instantaneous asset price has a positive correlation with the instantaneous option risk, while the one-day delayed asset price has a negative correlation with the instantaneous option risk

Remark 5.2.1 Here, X-month Y-day means that we use all options with time-to-maturity up to X months to calculate every day's algebraic average risk and take Y days of risk to correlate with Y days of returns of the underlying asset

Remark 5.2.2 '$Delay\ 0$' means risk and underlying asset returns are calculated simultaneously. '$Delay - 1$' means that risk is taken one day earlier than underlying asset returns. Therefore '$Delay - 1$' means to investigate predictive ability of risk on the underlying asset returns (Tang & Zhang, 2014).

5.3 Variance capture

Although correlation confirms the relationship between risk and underlying asset return, it is apparent that if we carry out regression analysis between risk and return, the R^2

statistic is very poor. Thus, correlation may not reflect the true relationship between risk and returns (Tang & Zhang, 2014).

In order to overcome this limitation, we carry out a regression of the simple average risk against various combinations of underlying prices or underlying returns, time to maturity and strike prices.

We arrange date, averaged implied volatility, averaged strike price, averaged time to maturity, asset price and asset return values for each day in six columns, see Table 5.3.1:

Date	Averaged Implied Volatility	Averaged Strike	Averaged T	Asset Price	Asset Return
11-02-2005	0.02838	120.50	127.00	120.77	0.00860
14-02-2005	0.04255	120.50	124.00	120.68	-0.00075
15-02-2005	0.06080	122.10	133.10	121.13	0.00373
16-02-2005	-0.02066	121.00	122.00	121.21	0.00066
17-02-2005	0.07913	120.50	121.00	120.23	-0.00809
18-02-2005	0.05307	120.50	120.00	120.39	0.00133
22-02-2005	0.13599	118.50	131.17	118.60	-0.01487
23-02-2005	0.09574	119.50	130.17	119.45	0.00717
...	...	...	...	...	...
26-11-2013	0.10377	173.18	124.59	180.68	0.00028
27-11-2013	0.09379	173.18	123.59	181.12	0.00244
29-11-2013	0.10337	173.18	121.59	181.00	-0.00066
02-12-2013	0.12519	173.18	118.59	180.53	-0.00260

Table 5.3.1: Data structure of averaged K, T, implied volatility, asset return and price

We obtain very good R^2 by using 5-day historical data rolling statistics. In particular, for call options the results are shown in Figure 5.3.1.

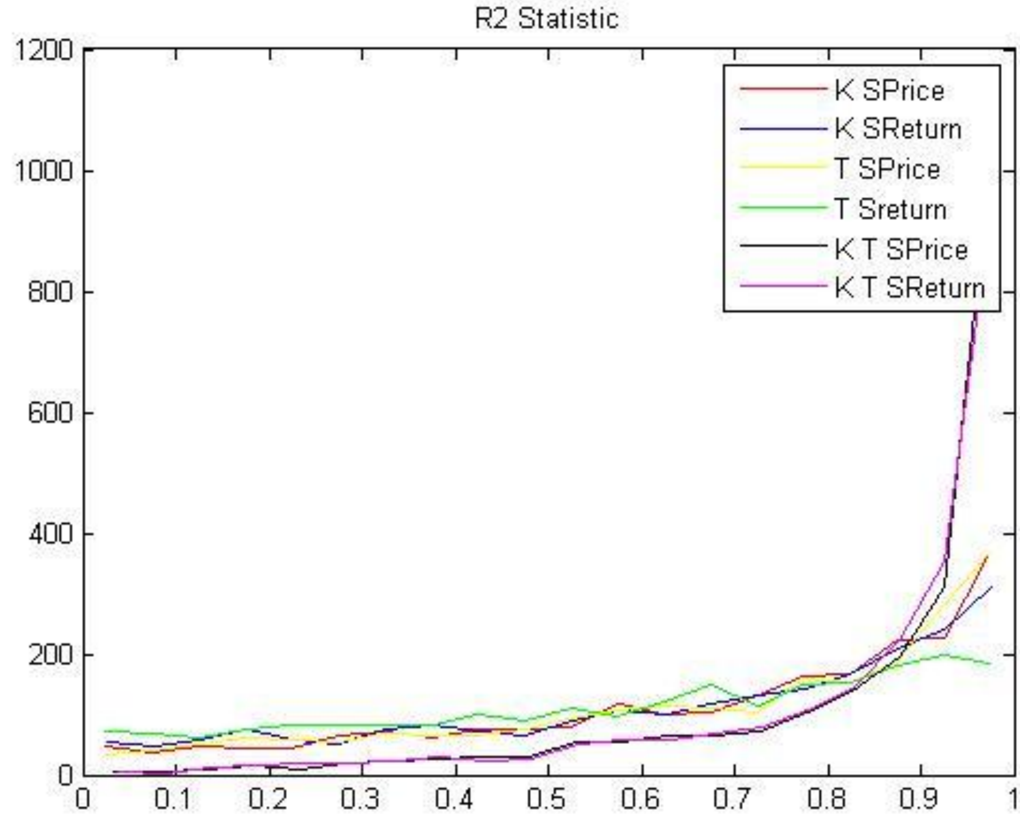

Figure 5.3.1 R^2 statistic for multi-regressed call options

For put options, the results are shown in Figure 5.3.2.

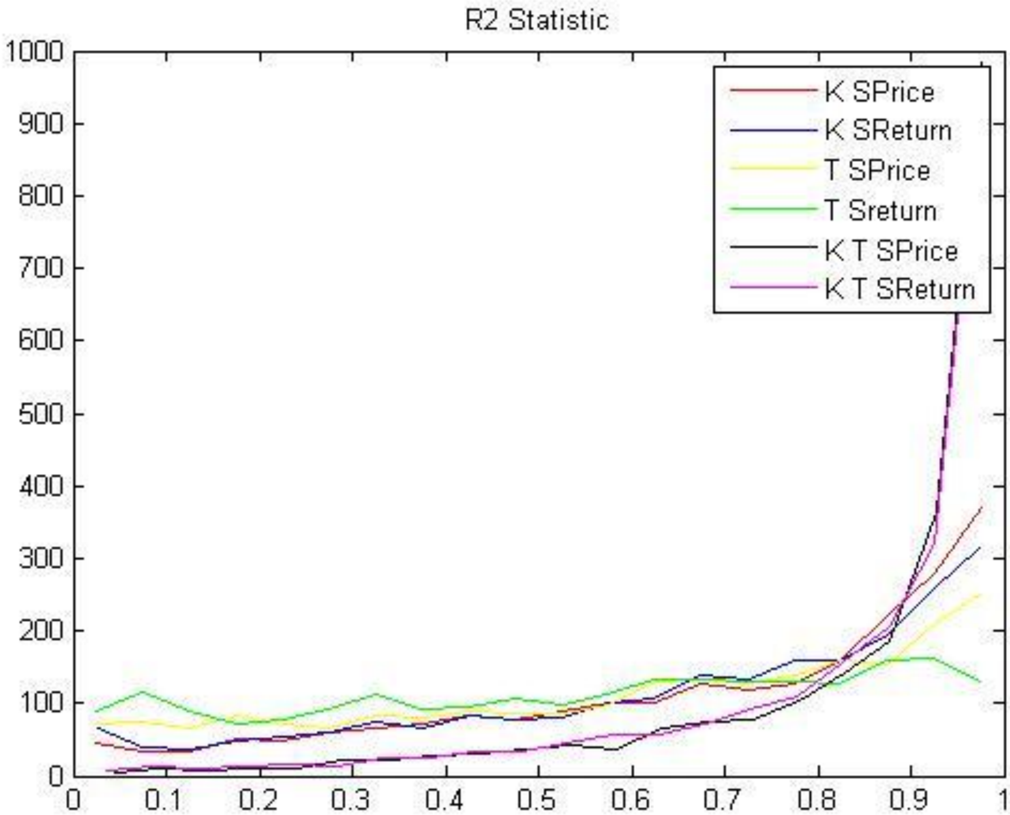

Figure 5.3.2 R^2 statistic for multi-regressed put options

Remark 5.3.1: Here, the horizontal axis represents the R^2 value of regression, the vertical axis represents the frequency that such R^2 has been achieved.

Remark 5.3.2: Other notations in the Figures 5.3.1 and 5.3.2:

- K SPrice —risk regresses against strike K and share price S
- K SReturn — risk regresses against strike K and share return
- T SPrice — risk regresses against time to maturity T and share price S
- T SReturn — risk regresses against time to maturity T and share return
- K T SPrice —risk regresses against strike K, maturity T and share price S
- K T SReturn —risk regresses against strike K, maturity T and share return

Since regression against ‘K T SPrice’ and ‘K T SReturn’ are particularly good, we summarise the statics on the signs of regression coefficients in Table 5.3.2 for call options and in Table 5.3.3 for put options.

K+,T+,SPrice+	0.026	K+,T+,SReturn+	0.008
K+,T+,SPrice+	0.017	K+,T+,SReturn+	0.035
K+,T-,SPrice+	0.025	K+,T-,SReturn+	0.008
K+,T-,SPrice-	0.017	K+,T-,SReturn-	0.034
K-,T+,SPrice+	0.001	K-,T+,SReturn+	0.023
K-,T+,SPrice-	**0.462**	K-,T+,SReturn-	**0.440**
K-,T-,SPrice+	0.008	K-,T-,SReturn+	0.015
K-,T-,SPrice-	**0.441**	K-,T-,SReturn-	**0.434**

Table 5.3.2 Frequency of signs for call options

K+,T+,SPrice+	**0.291**	K+,T+,SReturn+	**0.287**
K+,T+,SPrice+	0.001	K+,T+,SReturn+	0.005
K+,T-,SPrice+	**0.55**	K+,T-,SReturn+	**0.491**
K+,T-,SPrice-	0.016	K+,T-,SReturn-	0.076
K-,T+,SPrice+	0.027	K-,T+,SReturn+	0.045
K-,T+,SPrice-	0.027	K-,T+,SReturn-	0.009
K-,T-,SPrice+	0.022	K-,T-,SReturn+	0.057
K-,T-,SPrice-	0.064	K-,T-,SReturn-	0.029

Table 5.3.3 Frequency of signs for put options

Remark 5.3.3: The notation in the third row, first column for example,

K+,T-,SPrice+

means that the regression coefficients are positive regarding K (strike price), negative regarding T (time to maturity), and positive regarding SPrice (asset price). In the third column, SReturn stands for the underlying asset return. The numbers in the second and the fourth columns are the probability of the corresponding situations. (Tang and Zhang, 2014).

Remark 5.3.4: The highlighted areas in Tables 5.3.1 and 5.3.2 above are the dominating phenomena in terms of probability, as it can be clearly seen that they amount for more than 70 per cent of the possibilities.

Conclusion: Call option risk is negatively correlated to underlying price or asset returns. Put option risk is positively correlated to underlying price or returns just as was expected.

5.4 Negative risk study

As we mentioned in Section 5.1, when the option price is underpriced in the real market, the volatility can be counted as negative. We undertake a further study for the negatively priced options and find that they have different characteristics in call and put options. For call options, when the option price is undervalued, we discover that the stock price achieves a local maximum with dominant probability. When the call option prices are undervalued for several days continuously, the first local extreme value of stock price within that interval is most likely a maximum. On the contrary, for put options, when the option price is undervalued, the stock price achieves a local minimum with dominant probability. If the option prices are undervalued for several days continuously, the first local extreme value of stock price is most likely a local minimum.

We conducted real data testing with SPY option data from 2005 to 2013. The results of our testing are listed in Table 5.4.1 for call options and Table 5.4.2 for put options.

Year Call	Number of intervals where only maximum or maximum appears first	Number of intervals where minimum appears first followed by maximum	Number of intervals which cannot catch max	Possibility for catching maximum or maximum appears first
2005	2225	45	460	81.50%
2006	2370	108	659	75.55%
2007	2319	49	679	76.11%
2008	1597	56	621	70.23%
2009	1367	33	634	67.21%
2010	1321	12	461	73.63%
2011	1405	16	603	69.42%
2012	1102	10	521	67.48%
2013	1694	19	773	68.14%
Overall	15400	348	5411	72.78%

Table 5.4.1 Catching maximum for underpriced call options

Year Put	Number of intervals where only minimum or minimum appears first	Number of intervals where maximum appears first followed by minimum	Number of intervals which cannot catch min	Possibility for catching minimum or minimum appears first
2005	1220	24	213	83.73%
2006	469	9	82	83.75%
2007	617	5	106	84.75%
2008	1882	56	1145	61.04%
2009	1194	44	684	62.12%
2010	1550	82	650	67.92%
2011	2366	158	1096	65.36%
2012	1861	143	1061	60.72%
2013	1912	186	583	71.32%
Year	13071	707	5620	67.38%

Table 5.4.2 catching minimum for underpriced put options

5.5 Results and discussions

When we use the real world option prices to calculate B-S implied volatility, we find that the correlation between Black-Scholes implied volatility and spot price is high.

The Black-Scholes implied volatility has a negative linear relationship with the asset price for call options and a positive linear relationship with the asset price for put options instantaneously.

From the real world option data, we conclude that it is possible that the option trading price is less than the Black-Scholes zero risk option price. Consequently, we cannot obtain a possible value of volatility. Hence, we created an artificial formula to calculate the B-S implied volatility – called negative risk. We confirmed that when the B-S implied volatility of the corresponding real world option price falls below zero, the underlying asset price achieves a local extreme value with dominant probability: for call options, local maximum asset prices are caught with dominant probability, whereas for put option, local minimum asset prices are caught with dominant probability.

The study of implied volatility is of great interest in an effort to better understand financial options. Moreover, it enables researchers to explore the significant links with underlying asset price movements.

Bibliography

Alexander C., 2008, ‘Options’, in *Market Risk Analysis*, 1st edn, John Wiley & Sons Ltd, West Sussex, pp. 137- 226.

Bichteler K., 2002, *Stochastic Integration and Stochastic Differential Equations*, Texas: University of Texas.

Bakshi G., Cao C. and Chen Z., 1997, ‘Empirical performance of alternative option pricing models’, *Journal of Finance*, vol. 52, pp. 2003-2049.

Bates DS., 1996, ‘Testing option pricing models’, *Handbook of Statistics: Statistical methods in Finance*, vol. 14, pp. 567-611.

Baxter M., and Rennie A., 1996, *Financial calculus- An introduction to derivative pricing*, 4th edn, Cambridge university press, Cambridge.

Black F., and Scholes M., 1973, ‘The pricing of options and corporate liabilities’, *Journal of Political Economy*, vol. 81, pp. 637-654.

Boyle P., 1977, ‘Options: A Monte-Carlo approach’, *Journal of Financial Economics*, vol. 21, pp. 323-338.

Boyle P., Broadie M. and Glasserman P., 1997, ‘Monte Carlo methods for security pricing’, *Journal of Economic Dynamics and Control*, vol. 21, pp. 1267-1321.

Carr P, Geman H., Madan D.B. and Yor M., 2007, ‘Self-decomposability and option pricing’, *Mathematical finance*, vol. 17, no. 1, pp. 31-57.

Carr P. and Wu L., 2008, ‘Variance risk premiums’, *The review of financial studies*, vol. 22, no. 3, pp. 1311-1341.

Chesney M. and Scott L., 1989, ‘Pricing European currency options: A comparison of the modified Black-Scholes model and a random variable model’, *Journal of Financial and Quantitative Analysis,* vol. 24, pp. 267-284.

Cox J., Ingersoll J., and Ross S., 1985, ‘An intertemporal general equilibrium models of asset prices’, *Econometrica*, vol. 53, pp. 363-384.

Cox J., Ross S., and Rubinstein M., 1979, ‘Option Pricing: A Simplified Approach’, *Journal of Financial Economics*, no. 7, pp. 229-263.

'Derivatives: The beauty in the beast', The Economist, Vol. 331, No. 7863 (14 May 1994), pp. 21.

Derman E., Kani, L., 1994, 'Riding on the smile', *Risk magazine*, vol. 7, pp. 32-39.

Dumas B., Fleming J. and Whaley E., 1998, 'Implied volatility functions: empirical tests', *Journal of the American Statistical Association*, vol. 104, pp. 2059-2106.

Dupire B., 1994, 'Pricing with a smile', *Risk magazine*, pp. 126-129.

Etheridge A., 2004, *A course in financial calculus*, Cambridge: Cambridge university press.

Feller W., 1971, *An Introducion to Probability Theory and Its Applications*. Wiley, New York.

Garman M.., 1976, 'A general theory of asset valuation under diffusion state processes', *Working paper*, No. 50, University of California, Berkeley.

Gatheral Jim, 2006, *The volatility surface: a practitioner's guide*. John Wiley & Sons, Inc, New Jersey.

Geske R. and Shastri K., 1985, 'Valuation by Approximation: A Comparison of Alternative Option Valuation Techniques', *Journal of financial and quantitative analysis*, vol. 20, no. 1, pp. 45-71.

Hull C., White A., 1987, 'The pricing of options on assets with stochastic volatilities', *Journal of Finance*, vol. 42, pp. 281-300.

Hull, John, 2010, *Options, Futures and other Derivatives*, Pearson Education, Inc, Boston.

Heston, L.S., 1993, 'A closed-form solution for options with stochastic applications to bond and currency options', *The Review of Financial Studies*, vol. 6, pp. 327-343.

Itô, K., 1944, 'Stochastic Integral', *Procceedings of the Imperial Academy*, vol. 20, pp. 519-524.

Lin C., Zheng Z., Cai S. and Qiu W., 2012, *Financial Engineering: Theory and Practice*, Beijing: Peking university press.

Longstaff A., and Schwartz ES., 2001, 'Valuing American Option by Simulation: A simple Least-Square Approach', *The Review of Financial Studies*, vol. 14, no. 1, pp. 113-147.

Merton C., 1973, 'The theory of rational option pricing', *The Bell Journal of Economics and Management Science*, vol. 4, pp. 141-183.

Merton C., 1976, 'Option pricing when underlying stock returns are discontinuous', *Journal of Financial Economics*, vol. 3, pp. 125-144.

Michael C., Fu, Scott B., 1999, 'Pricing American options: A comparison of Monte Carlo simulation approaches', *Journal of Computational Finance*, vol. 4, pp. 39-88.

Rubinstein M., 1994, 'Implied binomial trees', *Journal of Finance*, vol. 49, pp. 771-818.

Schachermayer, Walter, 2002, 'No Arbitrage: On the Work of David Kreps', *Positivity*, vol. 6, no. 3, pp. 359-368.

Stein E. and Stein J., 1991, 'Stock price distribution with stochastic volatility: An analytic approach', *The Review of Financial Studies*, vol. 4, pp. 727-752.

Tang Q. and Yan D., 2010, 'Autoregressive trending risk function and exhaustion in random asset price movement', *Journal of Time Series Analysis*, Volume 6, Issue 6, pp. 465-470.

Tang Q. and Zhang L., 2014, 'Some risk measures and their applications in financial data analysis', in *The 2014 international conference on advances in big data analytics, Las Vegas, July 21-24, 2014,* CSREA Press, Las Vegas, pp. 173-179 (http://worldcomp-proceedings.com/proc/p2014/).

Wilmott P., Howison S. and Dewynne J. 1995, *The mathematics of financial derivatives*, New York: Cambridge university press.