

LOCAL VOLATILITY MODELLING

Roel van der Kamp

July 13, 2009

A DISSERTATION SUBMITTED FOR THE DEGREE OF
Master of Science in Applied Mathematics (Financial Engineering)



I have to understand the world, you see.

- Richard Philips Feynman

Foreword

This report serves as a dissertation for the completion of the Master programme in Applied Mathematics (Financial Engineering) from the University of Twente. The project was devised from the collaboration of the University of Twente with Saen Options BV (during the course of the project Saen Options BV was integrated into AllOptions BV) at whose facilities the project was performed over a period of six months.

This research project could not have been performed without the help of others. Most notably I would like to extend my gratitude towards my supervisors: Michel Vellekoop of the University of Twente, Julien Gosme of AllOptions BV and François Myburg of AllOptions BV. They provided me with the theoretical and practical knowledge necessary to perform this research. Their constant guidance, involvement and availability were an essential part of this project.

My thanks goes out to Irakli Khomasuridze, who worked beside me for six months on his own project for the same degree. The many discussions I had with him greatly facilitated my progress and made the whole experience much more enjoyable.

Finally I would like to thank AllOptions and their staff for making use of their facilities, getting access to their data and assisting me with all practical issues.

RvdK

Abstract

Many different models exist that describe the behaviour of stock prices and are used to value options on such an underlying asset. This report investigates the local volatility model in which the volatility of the underlying asset is assumed to be a deterministic function of both time and the underlying asset price.

First the report considers how the local volatility surface can be extracted from market data for option prices. Theoretically this can be achieved by Dupire's formula, but it appears that in practice it is better to transform this equation so that the local volatility surface can be extracted from the implied volatilities. To fit the implied volatility surface to market data smoothed thin plate splines are used.

Secondly a pricing mechanism has to be devised to value options using the local volatility surface. For this trinomial trees are used. The classical tree model is adjusted to make it work properly in the presence of local volatility, particularly to avoid the occurrence of negative transition probabilities. The method is quick and can easily incorporate discrete dividend payments.

The accuracy of this method is verified for European and American options. Prices generated for European options are compared to Black-Scholes prices and prices for American options are compared to prices generated by Monte Carlo simulations. It is shown that the model works accurately for both European and American options.

Finally the model is tested on real market data. The prices generated by the local volatility method are not always within the bid-ask spread of the market. Since the implied volatilities were extracted from the market data by inverting a different pricing mechanism, this shows there is non-negligible difference between these two methods. Also the stability of the local volatility surface and delta hedging are considered. On the basis of the analysis of the data used, no definitive statement can be made on the performance of the delta hedges suggested by the local volatility model compared to delta's suggested by the implied volatilities.

Contents

1	Introduction	1
1.1	Financial Terminology	1
1.2	Mathematical Framework	2
1.3	The Local Volatility Model	3
1.3.1	Volatility Skew	4
1.3.2	Explaining the Skew	4
1.3.3	Local Volatility	5
2	Obtaining the Local Volatility	7
2.1	Dupire's Equation	7
2.1.1	Derivation	7
2.1.2	Local Volatility as a Conditional Expectation	9
2.2	Problems Using Dupire's Formula	11
2.3	Local Volatility as a Function of Implied Volatility	11
2.4	Fitting the Implied Volatility Surface	14
2.4.1	Thin Plate Splines	14
3	Binomial Trees	17
3.1	The Standard Binomial Tree	17
3.1.1	Continuous Time Analysis	17
3.1.2	Discrete Time Analysis	18
3.1.3	Recombining the Binomial Tree	19
3.1.4	Binomial Tree with Local Volatility	20
3.1.5	Algorithm	21
3.1.6	Instability of the Binomial Tree	21
3.2	The Logarithmic Binomial Tree	22
3.2.1	Continuous Time	22
3.2.2	Discrete Time	22
3.2.3	Algorithm	24
3.2.4	Instability of the Logarithmic Binomial Tree	24
4	Trinomial Trees	25
4.1	Basic Equations	25
4.2	Fixing u and d	26
4.2.1	Negative Probabilities	27
4.2.2	Determining u_0 at the Origin of the Tree	28
4.2.3	Implementing a Cutoff for High Volatilities	29
4.2.4	Algorithm	30
4.3	Fixing Probabilities	31
4.3.1	Origin of the Tree	31
4.3.2	Rest of the Tree	32
4.3.3	Negative Probabilities and Cutoff	34
4.3.4	Algorithm	35

5	Dividends	36
5.1	Kinds of Dividends	36
5.2	Direct Implementation	37
5.3	Dividend Adjustments to Model	37
5.4	The Vellekoop-Nieuwenhuis Method	38
5.5	Interest Rates	39
6	Delta Hedging	40
6.1	Black-Scholes Framework	40
6.2	Minimum Variance Delta	40
6.3	Local Delta	41
6.3.1	Stickiness Assumptions	42
6.3.2	Performance of Local Delta	43
7	Monte Carlo Simulations	44
7.1	Creating Sample Paths	44
7.2	Antithetic Sampling	44
7.3	European Options	45
7.4	American Options	46
7.4.1	Bias and Convergence of LSM	46
8	Numerical Results	48
8.1	Finite Differences Approximation for Derivatives	48
8.2	SABR Model	49
8.2.1	Theoretical Framework	50
8.2.2	Volatility Surface	51
8.2.3	Binomial Tree Instability	52
8.2.4	Logarithmic Binomial Tree Instability	53
8.2.5	Comparison Between Fixing u and d and Fixing Probabilities	54
8.2.6	Valuing European Calls with p_f with Different Strikes	54
8.2.7	Monte Carlo Simulations for European Options	57
8.2.8	Monte Carlo Simulations for American Options	58
8.3	AEX	59
8.4	Royal Dutch/Shell	62
8.4.1	The Volatility Surfaces	63
8.4.2	Verifying Results	65
8.4.3	Stability of Local Volatility Surface	65
8.4.4	Delta Hedging	68
9	Conclusion	71
A	Appendix: Delta Hedging Results	72
	References	78

1 Introduction

For a full understanding of the contents of this report, some basic knowledge is needed. In 1.1 an overview of financial terms used in this report is given. Section 1.2 gives a short overview of some of the important mathematical concepts that are essential in the mathematical description of financial processes. For a more elaborate insight into these, and other, concepts the reader is referred to the literature [7, 39]. In section 1.3 an overview of research regarding local volatility is given and the model is introduced.

1.1 Financial Terminology

A derivative is a financial product whose value depends on some reference entity, which is commonly known as the underlying (asset). In theory the underlying can be anything as long as some sort of objective measurement of it can take place. This could be the value of a house, the temperature at Dam square in Amsterdam or the concentration of pollutants in the North Sea. The value of the derivative depends on the measured value of the underlying; it is contingent on the underlying and is therefore sometimes referred to as a contingent claim. In financial markets the underlying is usually a traded financial asset, such as an equity share, bond or commodity, or some observed economic variable such as an interest or currency exchange rate. Derivatives can be traded on an official exchange, such as NYSE Euronext, or through private transactions, usually referred to as over-the-counter (OTC).

A common derivative is an option. This gives the holder the right, but not the obligation, to perform a specific transaction under certain, agreed upon, conditions. The fact that the right to exercise an option is optional, means the option can never have a negative value.

The simplest options, known as plain vanilla options, are the standard call and put. A call option gives the holder the right to buy the underlying asset at, or before, a certain time for a certain price. This time is the option's time, or date, of maturity and the price is the option's strike price. Similarly, a put option gives the holder the right to sell the underlying asset at, or before the time of maturity for the strike price. An option that gives the right to perform this transaction, exercise the option, only on the exact time of maturity is said to be European. An option that can be exercised at any point in time before the maturity as well, is said to be American. The intermediary form, options that can be exercised at several specified points in time before maturity, are known as Bermudan options.

Any sort of option can be devised, as long as some other party is willing to pay for it. Options with non-standard pay-offs, strikes, exercise possibilities, that consist of a combination of other options or any other non-standard conceivable structure are known as exotic options.

An important concept in option theory is hedging. It is the replication of the contingent claims, by buying (longing) or selling (shorting) other financial products (usually the underlying). By ensuring that the replicating portfolio has the same payoff as the option does, the option trader eliminates the uncertainty of making a loss (or profit).

A market in which all contingent claims can be replicated is said to be complete. In a complete market an option has a unique price, which can be determined by finding the cost of setting up,

and maintaining, its replicating portfolio. In a complete market there are no possibilities of making a risk-free profit in excess of the risk-free interest rate, known as arbitrage opportunities.

1.2 Mathematical Framework

Some basic concepts in measurability and stochastic calculus are presented below. In the short overview given the mathematical details are intentionally kept to a minimum to enhance the readability.

All concepts in financial mathematics are defined within a certain probability space $(\Omega, \mathcal{F}, \mathbb{Q})$. Here Ω denotes the total space, \mathcal{F}_t denotes the σ -algebra of all the information that is known at time t and \mathbb{Q} denotes the risk neutral probability measure, which governs the probabilities of events occurring in this space. The domain of \mathbb{Q} is \mathcal{F} .

A random variable is a function that assigns values to outcomes of a probabilistic experiment. Its future value is uncertain. This uncertainty is known as stochasticity, explaining why random variables are also known as stochastic variables. If the value of a particular random variable, X_t , is known at time t it is said to be \mathcal{F}_t -measurable (notation $X_t \in \mathcal{F}_t$). For any time t_2 after t , the value of the random variable cannot be determined at time t .

The collection, X , of \mathcal{F}_t -measurable random variables, $\{X_t : 0 \leq t \leq T\}$, is a stochastic process. If a stochastic process Y behaves such that every realisation Y_t is \mathcal{F}_t -measurable, then it is said that it is adapted to the filtration $\{\mathcal{F}_t\}_{0 \leq t \leq T}$. Adapted processes are also known as non-anticipating processes, since their values do not depend on future events.

Conditional expectation is the expected value of a random variable given, conditional on, a certain amount of information. Let \mathcal{G} be a σ -algebra contained in \mathcal{F} . Then the conditional expectation of X given the information contained in \mathcal{G} is denoted by

$$\mathbb{E}[X|\mathcal{G}] \tag{1.1}$$

It then follows that when X is adapted to the filtration $\{\mathcal{F}_t\}_{0 \leq t \leq T}$

$$\mathbb{E}[X_t|\mathcal{F}_t] = X_t \tag{1.2}$$

If $X \notin \mathcal{G}$, then its value is unknown at time t and the expected value, under the probability measure \mathbb{Q} , is an objective prediction of the future value. Formally this means

$$\mathbb{E}[X|\mathcal{G}] = \int_{\mathcal{G}} X(\omega) d\mathbb{Q}(\omega) \tag{1.3}$$

In a complete market the value of an option at time t , V_t , with a payoff X_T at time T , is the expected value of the payoff, discounted by the risk free interest rate to time t , conditional on the information known at time t

$$V_t = \mathbb{E}[e^{-\int_t^T r_s ds} X_T | \mathcal{F}_t] \tag{1.4}$$

Most stochastic processes are described by stochastic differential equations(SDEs), usually of the form

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t \quad (1.5)$$

with $a(t, X_t)$ and $b(t, X_t)$ \mathcal{F}_t -measurable. W_t is standard Brownian motion (also known as a Wiener process). This is a random process that describes a motion beginning at $W_0 = 0$. In each time period $t_2 - t_1$ its increment, $W_{t_2} - W_{t_1}$, is independent of everything that happened before t_1 , and its values are normally distributed with mean 0 and variance $t_2 - t_1$.

Finally, the workhorse of financial mathematics can be introduced. If a stochastic process follows an SDE of the form (1.5) then a function f of this process and time is described by Itô's formula

$$\begin{aligned} df(t, X_t) &= \frac{\partial f(t, x)}{\partial t} dt + \frac{\partial f(t, x)}{\partial x} dX_t + \frac{1}{2} \frac{\partial^2 f(t, x)}{\partial x^2} dX_t \cdot dX_t \\ &= \left(\frac{\partial f(t, x)}{\partial t} + a(t, X_t) \frac{\partial f(t, x)}{\partial x} + \frac{1}{2} b(t, X_t)^2 \frac{\partial^2 f(t, x)}{\partial x^2} \right) dt + b(t, X_t) \frac{\partial f(t, x)}{\partial x} dW_t \end{aligned} \quad (1.6)$$

In general asset prices are assumed to be of the form (1.5). It is then obvious why Itô's formula plays such an important role in option theory. The price of an option, if it is a function of the underlying asset, satisfies this formula.

1.3 The Local Volatility Model

Modern option pricing theory came into existence with the advent of the influential paper by Black and Scholes [8]. In it the authors showed that, under certain model assumptions, there exists a unique price for European options since the payoff can be perfectly replicated by a portfolio consisting of the underlying asset and a risk free money account, which is the mathematical equivalent of buying bonds of an institution that is assumed to never default on its obligations. Usually the bonds issued by the sovereign governments, such as the United States (T-bills) and Germany (Bunds), are assumed to give risk free returns. Since this portfolio (which must be adjusted continuously) gives the same payoff as the option, the price of setting up this portfolio must be the price of the option.

In this model it is assumed that the asset which is underlying to the derivatives being considered (typically share prices), evolves according to geometric Brownian motion

$$\frac{dS_t}{S_t} = (r - q)dt + \sigma dW_t \quad (1.7)$$

Here it is assumed that the drift term, $r - q$, (the risk-free interest rate minus the dividend yield) and the diffusion term, σ , (commonly known as the asset's volatility) driving this process are constant. W_t is standard Brownian motion. If geometric Brownian motion accurately describes the dynamics of asset prices, then the asset prices have distributions which at all times are lognormal. Theoretically this is a nice result, since it gives rise to closed form analytical formulae for plain vanilla options. This Black-Scholes equation can then price plain vanilla options in a quick and

neat way, since only a few constant variables have to be considered.

Although satisfactory for European options, the Black-Scholes model comes up short for more complex options, such as Asian options (whose payoff depends on the average price of the underlying asset over time), barrier options (whose value depends on whether a specific boundary value has been attained by the underlying asset before its maturity) or even common American options. For these options no analytic solution can be given. Fortunately in the Black-Scholes framework, American and Asian options can be accurately priced by so-called binomial trees introduced by Cox, Ross and Rubinstein (CRR) [14, 41].

1.3.1 Volatility Skew

In reality things are more complicated than the model of Black and Scholes assumes. Market participants have long noted that using the same constant variables for all options result in prices not compatible with the market. It seems that different options on the same underlying asset are governed by different volatilities. The constant value for the volatility which, once plugged into the standard Black-Scholes equation, gives the market price for the option, commonly known as the implied volatility, seems to be dependent on the strike price, K , and the time to maturity, T , of the option under consideration. The dependence on T can easily be solved by introducing a time-dependency for the volatility as was shown shortly after the original Black-Scholes article [46].

The dependence on K is commonly known as the volatility ‘skew’, ‘smile’ or ‘frown’ (depending on the exact relation between volatility and the strike price). Before the market crash of 1987 no noticeable skew occurred, but ever since it has become a common feature of financial markets [39]. If the Black-Scholes model was an exact description of reality, it should be concluded that every option has an underlying asset with different dynamics, which is obviously not the case. This would also suggest that, because different options seem to have different volatilities, using binomial trees to value options, would mean having to build a different tree for the asset price process for each different option. This seems non-sensical. Why would the price process of the underlying asset depend on the value of the derivative on it? It also poses practical issues, since the valuing techniques that are used should be fast, providing market participants with up to date information. Furthermore, if the volatility is dependent on the maturity and strike price, how should an exotic option be priced whose parts are made up of options with different strikes and maturities? A better description of the volatility process such that one tree can be used to price different options is thus needed.

1.3.2 Explaining the Skew

Over the years different methods have been proposed to make adjustments to the Black-Scholes model to make it more accurately describe market reality, and in particular the volatility skew. The main proposed methods are stochastic volatility processes [19, 34, 38, 40], jump-diffusion processes [2, 45] and local volatility models [10, 15, 18, 20, 21, 25, 26, 27, 32, 49, 59].

Stochastic volatility processes were introduced by Hull and White [40]. In it the volatility itself is a process that satisfies a stochastic differential equation. The most famous stochastic volatility

models are the Heston model [38] and the SABR model [34] (section 8.2.1). How these models cause the volatility skew in the market is discussed in [19, 34].

Jump-diffusion models were first introduced by Merton [45]. These models incorporate discontinuous jumps in the underlying asset price. These resemble reality where events can have sudden impacts on asset prices. How this explains the volatility skew is described in [2].

The local volatility model assumes the volatility is a deterministic function of the asset price and time. It came into existence when Dupire [25, 26] showed that, in the presence of volatility skews, consistent models can be built if the asset price process is assumed to have the following dynamics under the risk neutral probability measure \mathbb{Q}

$$\frac{dS_t}{S_t} = (r_t - q_t)dt + \sigma(t, S_t)dW_t \quad (1.8)$$

where the volatility is now a deterministic function of time and the asset price and r_t and q_t denote the continuously compounded short rate and dividend respectively. In this case the diffusion process is usually referred to as local volatility.

1.3.3 Local Volatility

Local volatility models, which are widely used in the finance industry [27], are the subject of this report. Whereas stochastic volatility and jump-diffusion models introduce new features into the model resulting from sound economic arguments, local volatility models try to stay close to the Black-Scholes model by introducing more flexibility into the volatility. This is one of the main reasons of fierce criticism of local volatility models [4].

The drawback of stochastic volatility and jump-diffusion models is that in describing the asset dynamics, they introduce new sources of stochasticity. Since the stochastic volatility and jumps in asset prices cannot be traded, these models lose the completeness of the original Black-Scholes model.

In the local volatility model the only stochastic behaviour introduced into the volatility function is a result of it being a function of the underlying asset price (if r_t and q_t are deterministic). So there is still just one source of stochasticity, ensuring the completeness of the Black-Scholes model is preserved. Completeness is important, because it guarantees unique prices. This is the stated reason to develop the local volatility model in Dupire's original paper [25].

In [26] Dupire's method of working with local volatility is described. Around the same time other methods for local volatility methods were developed by Rubinstein [50] and Derman and Kani [18]. Both of these methods use so-called 'implied trees'. The idea of these is to price options with standard CRR trees with a constant volatility, and then adjust the volatility at different places in the tree to obtain the correct market prices for the options. This method is not clear-cut and numerous adjustments have to be made to make it work in practice. Furthermore, it is notoriously unstable [49]. Therefore Dupire's method is preferred over the implied tree method in this report.

This report is organised as follows. First it is described how the local volatility surface can be extracted from plain vanilla option prices in section 2. In section 3 and 4 it is described how the local volatility surface can be used to price options by use of binomial and trinomial trees, respectively. The incorporation of dividend payments in the tree is discussed in section 5. Delta hedging in the local volatility model follows in section 6 and section 7 is devoted to Monte Carlo simulations. Finally the model is tested on real world examples in section 8.

2 Obtaining the Local Volatility

Before the local volatility can be used to price derivatives, a procedure to obtain the local volatility function must be devised. In this section Dupire's formula is derived in 2.1, which allows the local volatility surface to be extracted from the prices of traded call options. Since in most practical situations implementing this formula proves not to be suitable for all options, as discussed in 2.2, a method is derived in 2.3 by which the local volatility surface can be extracted from the implied volatility surface. The process of fitting an implied volatility surface to real data is covered in 2.4.

2.1 Dupire's Equation

According to standard financial theory, the price at time t of a call option with strike price K , maturity time T is the discounted expectation of its payoff, under the risk-neutral measure. Letting $D_{0,T}$ denote the discount rate from the current time t_0 to maturity and $\phi(T, s)$ the risk neutral probability density of the underlying asset at maturity. More accurately the density should be written as $\phi(T, s; t_0, S_0)$, since it is the transition probability density function of going from state (t_0, S_0) to (T, s) . But since t_0 and S_0 are considered to be given constants, for brevity it is written as $\phi(T, s)$. This means

$$\begin{aligned} D_{0,T} &= e^{-\int_{t_0}^T r_s ds} \\ C &= \mathbb{E} [D_{0,T}(S_T - K)^+ | \mathcal{F}_0] \\ &= D_{0,T} \int_K^\infty (s - K) \phi(T, s) ds \end{aligned} \tag{2.1}$$

Here, as in the rest of this report, it is assumed that the term structure for the short rate r_t is a deterministic function known at the current time t_0 (\mathcal{F}_0 is the shorthand version of \mathcal{F}_{t_0}). Comparable equations in the presence of stochastic interest rates can be derived, but this is outside the scope of this report.

2.1.1 Derivation

The risk-neutral density distribution of the asset price at maturity is $\phi(T, s)$. Since this is a probability density function its time evolution will be described by the forward Kolmogorov (or Fokker-Planck) equation

$$0 = \frac{\partial \phi(t, s)}{\partial t} + (r_t - q_t) \frac{\partial}{\partial s} [s \phi(t, s)] - \frac{1}{2} \frac{\partial^2}{\partial s^2} [\sigma(t, s)^2 s^2 \phi(t, s)] \tag{2.2}$$

For further calculations the derivatives of (2.1) with respect to K are needed. They are given by

$$\begin{aligned}
\frac{\partial C}{\partial K} &= D_{0,T} \frac{\partial}{\partial K} \int_K^\infty (s - K) \phi(T, s) ds \\
&= D_{0,T} \left[-(K - K) \phi(T, K) - \int_K^\infty \phi(T, s) ds \right] \\
&= -D_{0,T} \int_K^\infty \phi(T, s) ds \\
\frac{\partial^2 C}{\partial K^2} &= -D_{0,T} \frac{\partial}{\partial K} \int_K^\infty \phi(T, s) ds \\
&= D_{0,T} \phi(T, K)
\end{aligned} \tag{2.3}$$

Combining equation (2.2) together with the definition in (2.1) gives

$$\begin{aligned}
\frac{\partial C}{\partial T} + r_T C &= D_{0,T} \int_K^\infty (s - K) \frac{\partial \phi(T, s)}{\partial T} ds \\
&= D_{0,T} \int_K^\infty (s - K) \left(\frac{1}{2} \frac{\partial^2}{\partial s^2} [\sigma(T, s)^2 s^2 \phi(T, s)] - (r_T - q_T) \frac{\partial}{\partial S_T} [S_T \phi(T, s)] \right) ds \\
&= \frac{1}{2} D_{0,T} \left(\left[(s - K) \frac{\partial}{\partial s} [\sigma(T, s)^2 s^2 \phi(T, s)] \right]_{s=K}^{s=\infty} - \int_K^\infty \frac{\partial}{\partial s} [\sigma(T, s)^2 s^2 \phi(T, s)] ds \right) \\
&\quad - D_{0,T} (r_T - q_T) \left([(s - K) s \phi(T, s)]_{s=K}^{s=\infty} - \int_K^\infty s \phi(T, s) ds \right) \\
&= -\frac{1}{2} D_{0,T} [\sigma(T, s)^2 S_T^2 \phi(T, s)]_{s=K}^{s=\infty} + (r_T - q_T) D_{0,T} \int_K^\infty s \phi(T, s) ds \\
&= \frac{1}{2} D_{0,T} \sigma(T, K)^2 K^2 \phi(T, K) + (r_T - q_T) \left(C + D_{0,T} K \int_K^\infty \phi(T, s) ds \right) \\
&= \frac{1}{2} \sigma(T, K)^2 K^2 \frac{\partial^2 C}{\partial K^2} + (r_T - q_T) \left(C + K \frac{\partial C}{\partial K} \right)
\end{aligned} \tag{2.4}$$

Here it is assumed that $\phi(T, S_T)$ behaves appropriately at the boundary condition of $S_T = \infty$ (for instance this is the case when ϕ decays exponentially fast for $S_T \rightarrow \infty$).

The final result is

$$\begin{aligned}
\frac{\partial C}{\partial T} &= \frac{1}{2} \sigma(T, K)^2 K^2 \frac{\partial^2 C}{\partial K^2} - q_T C - (r_T - q_T) K \frac{\partial C}{\partial K} \\
\Rightarrow \sigma(T, K)^2 &= 2 \frac{\frac{\partial C}{\partial T} + (r_T - q_T) K \frac{\partial C}{\partial K} + q_T C}{K^2 \frac{\partial^2 C}{\partial K^2}}
\end{aligned} \tag{2.5}$$

The latter of the above equations is commonly known as the Dupire formula, derived in this form by Derman and Kani [18] but the method was developed by Dupire [26]. Since at any point in time the value of call options with different strikes and times to maturity can be observed in the market, the local volatility is a deterministic function, even when the dynamics of the spot volatility are stochastic.

2.1.2 Local Volatility as a Conditional Expectation

A different approach dealing directly with (2.1) instead of considering the forward Kolmogorov equation, reveals an interesting property of the local volatility.

Reformulating (2.1) gives

$$\begin{aligned} C &= \mathbb{E} [D_{0,T}(S_T - K)^+ | \mathcal{F}_0] \\ &= \mathbb{E} [D_{0,T}(S_T - K) \mathbf{1}_{\{S_T > K\}} | \mathcal{F}_0] \end{aligned} \quad (2.6)$$

Where $\mathbf{1}$ is the indicator function having the following properties

$$\begin{aligned} \mathbf{1}_{\{s > K\}} &= \begin{cases} 1 & \text{if } s > K, \\ 0 & \text{if } s \leq K. \end{cases} \\ \frac{\partial}{\partial s} \mathbf{1}_{\{s > K\}} &= \delta(s - K) \\ \frac{\partial}{\partial K} \mathbf{1}_{\{s > K\}} &= \frac{\partial}{\partial K} (1 - \mathbf{1}_{\{K \geq s\}}) \\ &= -\delta(s - K) \end{aligned} \quad (2.7)$$

where $\delta(\cdot)$ is the Dirac-delta function.

Under normal integrability assumptions Fubini's theorem holds and the expectation and derivative operator can be interchanged. This leads to

$$\begin{aligned} \frac{\partial C}{\partial K} &= \frac{\partial}{\partial K} \mathbb{E} [D_{0,T}(S_T - K) \mathbf{1}_{\{S_T > K\}} | \mathcal{F}_0] \\ &= -\mathbb{E} [D_{0,T} \mathbf{1}_{\{S_T > K\}} | \mathcal{F}_0] - \mathbb{E} [D_{0,T}(S_T - K) \delta(S_T - K) | \mathcal{F}_0] \\ &= -\mathbb{E} [D_{0,T} \mathbf{1}_{\{S_T > K\}} | \mathcal{F}_0] \\ \frac{\partial^2 C}{\partial K^2} &= -\mathbb{E} \left[D_{0,T} \frac{\partial}{\partial K} \mathbf{1}_{\{S_T > K\}} | \mathcal{F}_0 \right] \\ &= \mathbb{E} [D_{0,T} \delta(S_T - K) | \mathcal{F}_0] \end{aligned} \quad (2.8)$$

From which it can be seen that the probability density function of the stock price at maturity being equal to K is the expected value of the Dirac-delta function

$$\phi(T, K) = \mathbb{E} [\delta(S_T - K) | \mathcal{F}_0] \quad (2.9)$$

Applying Itô to (2.6)

$$\begin{aligned} dC &= \mathbb{E} [d(D_{0,T}(S_T - K) \mathbf{1}_{\{S_T > K\}}) | \mathcal{F}_0] \\ &= \mathbb{E} \left[\frac{\partial D_{0,T}}{\partial T} (S_T - K) \mathbf{1}_{\{S_T > K\}} dT + D_{0,T} \frac{\partial}{\partial s} [(s - K) \mathbf{1}_{\{s > K\}}] \Big|_{s=S_T} dS_T \right. \\ &\quad \left. + D_{0,T} \frac{1}{2} \frac{\partial^2}{\partial s^2} [(s - K) \mathbf{1}_{\{s > K\}}] \Big|_{s=S_T} dS_T \cdot dS_T | \mathcal{F}_0 \right] \end{aligned} \quad (2.10)$$

and using the following identities

$$\begin{aligned}
\frac{\partial D_{0,T}}{\partial T} &= -r_T e^{-\int_{t_0}^T r_s ds} \\
&= -r_T D_{0,T} \\
\frac{\partial}{\partial s} [(s - K) \mathbf{1}_{\{s > K\}}] &= \mathbf{1}_{\{s > K\}} + (s - K) \delta(s - K) \\
&= \mathbf{1}_{\{s > K\}} \\
\frac{\partial^2}{\partial s^2} [(s - K) \mathbf{1}_{\{s > K\}}] &= \frac{\partial}{\partial s} \mathbf{1}_{\{s > K\}} \\
&= \delta(s - K)
\end{aligned} \tag{2.11}$$

The resulting expression for (2.10) is

$$\begin{aligned}
dC &= D_{0,T} \mathbb{E}[-r_T (S_T - K) \mathbf{1}_{\{S_T > K\}} dT + \mathbf{1}_{\{S_T > K\}} S_T [(r_T - q_T) dT + \sigma(T, S_T) dW_T] \\
&\quad + \frac{1}{2} \delta(S_T - K) S_T^2 \sigma^2(T, S_T) dT | \mathcal{F}_0] \\
&= D_{0,T} \mathbb{E}[r_T K \mathbf{1}_{\{S_T > K\}} - q_T S_T \mathbf{1}_{\{S_T > K\}} + \frac{1}{2} \delta(S_T - K) K^2 \sigma^2(T, S_T) | \mathcal{F}_0] dT \\
\Rightarrow \frac{\partial C}{\partial T} &= r_T D_{0,T} K \mathbb{E}[\mathbf{1}_{\{S_T > K\}} | \mathcal{F}_0] - q_T (C + D_{0,T} K \mathbb{E}[\mathbf{1}_{\{s > K\}} | \mathcal{F}_0]) \\
&\quad + \frac{1}{2} D_{0,T} K^2 \mathbb{E}[\delta(S_T - K) \sigma^2(T, S_T) | \mathcal{F}_0] \\
&= (r_T - q_T) D_{0,T} K \mathbb{E}[\mathbf{1}_{\{S_T > K\}} | \mathcal{F}_0] - q_T C + \frac{1}{2} D_{0,T} K^2 \mathbb{E}[\delta(S_T - K) \sigma^2(T, S_T) | \mathcal{F}_0]
\end{aligned} \tag{2.12}$$

The last term in this equation can be transformed by

$$\mathbb{E}[\delta(S_T - K) \sigma^2(T, S_T) | \mathcal{F}_0] = \mathbb{E}[\sigma^2(T, S_T) | S_T = K, \mathcal{F}_0] \mathbb{E}[\delta(S_T - K) | \mathcal{F}_0] \tag{2.13}$$

Now using (2.8), this results in

$$\frac{\partial C}{\partial T} = -(r_T - q_T) \frac{\partial C}{\partial K} - q_T C + \frac{1}{2} K^2 \mathbb{E}[\sigma^2(T, S_T) | S_T = K, \mathcal{F}_0] \frac{\partial^2 C}{\partial K^2} \tag{2.14}$$

$$\Rightarrow \mathbb{E}[\sigma^2(T, S_T) | S_T = K, \mathcal{F}_0] = 2 \frac{\frac{\partial C}{\partial T} + (r_T - q_T) \frac{\partial C}{\partial K} + q_T C}{K^2 \frac{\partial^2 C}{\partial K^2}} \tag{2.15}$$

Comparing equations (2.15) and (2.5) shows that the local variance can be seen as the expected variance at maturity given that the asset price at maturity is equal to the strike price. This result first appeared in [21]. It gives further insight into the nature of local volatility. An analogy can be made with interest rates. The local volatility surface is comparable to the yield curve for interest rates. It is an expectation of future instantaneous volatilities (future spot rates). This does not mean that this expected value will actually be realised but it is possible at the current time to lock in this value by trading different financial products. For interest rates this means buying and selling bonds of different maturities, for local volatility it means buying and selling options with differing strikes and maturities (the exact procedure for this is described in detail in [21]). Furthermore, the implied volatility is the constant value for the volatility which is consistent with option prices in the market, exactly like the yield is the constant value for the interest rate consistent with bond prices in the market.

2.2 Problems Using Dupire’s Formula

Given a certain local volatility function, the price of all sorts of contingent claims on the underlying can be priced. By the method described in 2.1 this process can be inverted, by extracting the local volatility surface from option prices given as a function of strike and maturity. The tacit assumption being that the option price is a continuous $C^{2,1}$ function, known over all possible strikes and maturities. Even if this assumption holds, problems arise in the implementation of (2.5). Since the option price function will never be known analytically, neither will its derivatives. Numerical approximations for the derivatives have to be made, which are by their very nature imperfect. Therefore problems can arise when the values to be approximated are very small and small absolute errors in the approximation can lead to big relative errors, perturbing the estimated quantity heavily. When the disturbed quantity is added to other values, the effect will be limited. This is not the case in Dupire’s formula where the second derivative with respect to the strike in the denominator stands by itself. This derivative will be very small for options that are far in- or out-of-the-money (the effect is particularly large for options with short maturities). Small errors in the approximation of this derivative will get multiplied by the strike value squared resulting in big errors at these values, sometimes even giving negative values, resulting in negative variances and complex local volatilities. This is, needless to say, unacceptable behaviour for a volatility function.

The continuity assumption of option prices is, of course, not very realistic. In practice option prices are known for certain discrete points. Usually option maturities correspond to the end of a certain fixed period, like the end of a month. So the number of different maturities is always limited. The same holds to a lesser degree for strikes. The result of this is that in practice the inversion problem is ill-posed: the solution is not unique and is unstable. This is an extra problem when dealing with Dupire’s formula in practice. One can smooth the option price data using Tikhonov regularisation [15, 35] or by minimising the function’s entropy [3, 51]. Both these methods try to estimate a stable option price function. These methods must, among other things, assure the resulting option price function is convex in the strike direction at every point to avoid negative local variance. This guarantees the positivity of the second derivative in the strike direction. This seems sensible, since the non-convexity of the option price leads to an arbitrage opportunity (a butterfly spread will have a negative price). It does, however, add a considerable amount of complexity to the model. An easier, and inherently more stable method to obtain the local volatility surface is to obtain it from the implied volatility surface.

2.3 Local Volatility as a Function of Implied Volatility

The local volatility can be described as a function of the implied volatility if a change of variables is made in (2.5) by using C as a function of some other variable. Instinctively this is not possible, because there is no closed form formula for C to be transformed. But use can be made of the Black-Scholes formula and the concept of implied volatility. The standard Black-Scholes environment with lognormal prices is a highly idealised world, which does not accurately describe reality. But as Rebonato [49] pointedly observed, the implied volatility is “the wrong number to put in the wrong formula to get the right price of plain-vanilla options”. Therefore the formula (2.5) can be expressed in terms of the implied volatility.

Using the following notation

$$\begin{aligned}\Sigma &= \sigma_{\text{imp}}(K, T) \\ \tau &= T - t_0\end{aligned}\tag{2.16}$$

with t_0 as the current time, and the following parametrisation (slightly adapting the method proposed in [32])

$$\begin{aligned}y &= \ln\left(\frac{K}{S_0}\right) + \int_{t_0}^T (q_s - r_s) ds \\ w &= \Sigma^2 \tau\end{aligned}\tag{2.17}$$

the option price has the expression

$$\begin{aligned}C_{\text{market}}(S_0, t_0, K, T, \sigma) &= C_{\text{BS}}(S_0, t_0, K, T, \Sigma) \\ &= S_0 e^{-\int_{t_0}^T q_s ds} [N(d_1) - e^y N(d_2)]\end{aligned}\tag{2.18}$$

where

$$\begin{aligned}d_1 &= -\frac{y}{\sqrt{w}} + \frac{\sqrt{w}}{2} \\ d_2 &= -\frac{y}{\sqrt{w}} - \frac{\sqrt{w}}{2}\end{aligned}\tag{2.19}$$

Now the partial derivatives of this expression of the call option with respect to T and K are needed, to plug into the Dupire formula.

$$\begin{aligned}\frac{\partial C}{\partial K} &= \frac{\partial C}{\partial y} \frac{\partial y}{\partial K} + \frac{\partial C}{\partial w} \frac{\partial w}{\partial K} \\ &= \frac{1}{K} \frac{\partial C}{\partial y} + \frac{\partial w}{\partial K} \frac{\partial C}{\partial w} \\ \frac{\partial^2 C}{\partial K^2} &= -\frac{1}{K^2} \frac{\partial C}{\partial y} + \frac{1}{K} \frac{\partial}{\partial K} \left(\frac{\partial C}{\partial y} \right) + \frac{\partial^2 w}{\partial K^2} \frac{\partial C}{\partial w} + \frac{\partial w}{\partial K} \frac{\partial}{\partial K} \left(\frac{\partial C}{\partial w} \right) \\ &= -\frac{1}{K^2} \frac{\partial C}{\partial y} + \frac{1}{K} \left(\frac{1}{K} \frac{\partial^2 C}{\partial y^2} + \frac{\partial w}{\partial K} \frac{\partial^2 C}{\partial w \partial y} \right) + \frac{\partial^2 w}{\partial K^2} \frac{\partial C}{\partial w} \\ &\quad + \frac{\partial w}{\partial K} \left(\frac{1}{K} \frac{\partial^2 C}{\partial y \partial w} + \frac{\partial w}{\partial K} \frac{\partial^2 C}{\partial w^2} \right) \\ &= \frac{1}{K^2} \left(\frac{\partial^2 C}{\partial y^2} - \frac{\partial C}{\partial y} \right) + \frac{2}{K} \frac{\partial w}{\partial K} \frac{\partial^2 C}{\partial w \partial y} + \frac{\partial^2 w}{\partial K^2} \frac{\partial C}{\partial w} + \left(\frac{\partial w}{\partial K} \right)^2 \frac{\partial^2 C}{\partial w^2} \\ \frac{\partial C}{\partial T} &= -q_T C + (q_T - r_T) \frac{\partial C}{\partial y} + \frac{\partial w}{\partial T} \frac{\partial C}{\partial w}\end{aligned}\tag{2.20}$$

Inserting these equations into (2.5) results in

$$\sigma_L^2 = 2 \frac{-q_T C + (q_T - r_T) \frac{\partial C}{\partial y} + \frac{\partial w}{\partial T} \frac{\partial C}{\partial w} + (r_T - q_T) \frac{\partial C}{\partial y} + (r_T - q_T) K \frac{\partial w}{\partial K} \frac{\partial C}{\partial w} + q_T C}{\left(\frac{\partial^2 C}{\partial y^2} - \frac{\partial C}{\partial y} \right) + 2K \frac{\partial w}{\partial K} \frac{\partial^2 C}{\partial w \partial y} + K^2 \frac{\partial^2 w}{\partial K^2} \frac{\partial C}{\partial w} + K^2 \left(\frac{\partial w}{\partial K} \right)^2 \frac{\partial^2 C}{\partial w^2}}\tag{2.21}$$

This equation can be simplified significantly by making use of the following identities

$$\begin{aligned}
\frac{\partial^2 C}{\partial w^2} &= \left(-\frac{1}{8} - \frac{1}{2w} + \frac{y^2}{2w^2} \right) \frac{\partial C}{\partial w} \\
\frac{\partial^2 C}{\partial w \partial y} &= \left(\frac{1}{2} - \frac{y}{w} \right) \frac{\partial C}{\partial w} \\
\frac{\partial^2 C}{\partial y^2} - \frac{\partial C}{\partial y} &= 2 \frac{\partial C}{\partial w}
\end{aligned} \tag{2.22}$$

which, after simplification, gives an expression of the local volatility in terms of the new variables y and w

$$\sigma_L^2 = \frac{\frac{\partial w}{\partial T} + (r_T - q_T)K \frac{\partial w}{\partial K}}{1 + K \frac{\partial w}{\partial K} \left(\frac{1}{2} - \frac{y}{w} \right) + \frac{1}{2} K^2 \frac{\partial^2 w}{\partial K^2} - \frac{1}{4} K^2 \left(\frac{\partial w}{\partial K} \right)^2 \left(\frac{1}{4} + \frac{1}{w} - \frac{y^2}{w^2} \right)} \tag{2.23}$$

Now the partial derivatives of w are given by

$$\begin{aligned}
\frac{\partial w}{\partial K} &= 2\Sigma\tau \frac{\partial \Sigma}{\partial K} \\
\frac{\partial^2 w}{\partial K^2} &= 2\tau \left(\frac{\partial \Sigma}{\partial K} \right)^2 + 2\Sigma\tau \frac{\partial^2 \Sigma}{\partial K^2} \\
\frac{\partial w}{\partial T} &= \Sigma^2 + 2\Sigma\tau \frac{\partial \Sigma}{\partial T}
\end{aligned} \tag{2.24}$$

and plugging these into (2.23) results in

$$\begin{aligned}
\sigma_L^2 &= \frac{\Sigma^2 + 2\Sigma\tau \left(\frac{\partial \Sigma}{\partial T} + (r_T - q_T)K \frac{\partial \Sigma}{\partial K} \right)}{1 + K\Sigma\tau \frac{\partial \Sigma}{\partial K} - 2\frac{Ky}{\Sigma} \frac{\partial \Sigma}{\partial K} + K^2\Sigma\tau \frac{\partial^2 \Sigma}{\partial K^2} - \frac{1}{4}K^2\Sigma^2\tau^2 \left(\frac{\partial \Sigma}{\partial K} \right)^2 + \frac{K^2y^2}{\Sigma^2} \left(\frac{\partial \Sigma}{\partial K} \right)^2} \\
&= \frac{\Sigma^2 + 2\Sigma\tau \left(\frac{\partial \Sigma}{\partial T} + (r_T - q_T)K \frac{\partial \Sigma}{\partial K} \right)}{\left(1 - \frac{Ky}{\Sigma} \frac{\partial \Sigma}{\partial K} \right)^2 + K\Sigma\tau \left(\frac{\partial \Sigma}{\partial K} - \frac{1}{4}K\Sigma\tau \left(\frac{\partial \Sigma}{\partial K} \right)^2 + K \frac{\partial^2 \Sigma}{\partial K^2} \right)}
\end{aligned} \tag{2.25}$$

which is consistent with other known versions of this formula such as the one found in [59]. This formula will be the main tool for extracting the local volatility surface from a given input. When the second derivative with respect to the strike price becomes very small it will not give rise to the same problems as would be experienced with the direct implementation of Dupire's formula (2.5).

From the formula above it can be seen that if the implied volatility does not exhibit any dependence on K

$$\begin{aligned}
\sigma_L^2 &= \Sigma^2 + 2\Sigma\tau \frac{\partial \Sigma}{\partial T} \\
&= \Sigma^2 + 2\Sigma\tau \frac{\partial \Sigma}{\partial \tau} \\
&= \frac{\partial(\Sigma^2\tau)}{\partial \tau} \\
\Rightarrow \Sigma^2 &= \frac{1}{\tau} \int_0^\tau \sigma_L^2(u) du
\end{aligned} \tag{2.26}$$

where the last integral is over the time variable. This means that, when the implied volatility surface does not have any skew, the implied total variance is the time average of the local variance.

2.4 Fitting the Implied Volatility Surface

When comparing equation (2.25) with (2.5) it is clear that the first problem described in 2.2 no longer exists. The transformation of Dupire's formula into one which depends on the implied volatility ensures that there no longer is a lone second derivative in the denominator as there was in (2.5). The second derivative of the implied volatility is now one term of a summation, so small errors in it will not necessarily lead to large errors in the local volatility function.

There is, of course, still the matter that the implied volatility is not a known continuous function of strike and maturity, but only known at certain points. To get the local volatility function from (2.25), some method has to be used to interpolate and extrapolate the given data points unto a surface. Since obtaining the local volatility out of the data involves taking derivatives, the extrapolated surface cannot be too rough, to avoid irregularities in the local volatility surface.

A good overview of different methods of fitting an implied volatility surface to data points can be found in [28]. All these methods can be subdivided in two parts. First a rough fit of a certain form is made to the data, which captures all the local information. Secondly some sort of smoothing is applied to this rough pre-smoothed surface, thereby ensuring the differentiability of the function and thus removing large spikes in the local volatility surface. Although this means the resulting surface does not necessarily go through all the data points, this can be justified given the inherent uncertainty of the data. In the market there is not one precise value for the option price, and thus implied volatility, because of the bid-ask spread. It seems reasonable to use the mid-market price for the modelling. Small deviations from this mid-market price are thus not a problem. It should of course be checked that the smoothed surface indeed matches all the bid-ask spreads.

Since the number of data points is always many times less than the number of grid points for the surface, there are many degrees of freedom in the fitting of the surface. The process of fitting the surface is therefore for a large part more art than science. It cannot be said with certainty which method of fitting the implied volatility surface is the best. Different methods have different advantages.

2.4.1 Thin Plate Splines

In this report the data is fitted to a surface by a thin plate spline (TPS), which is considered to be a natural candidate for this type of problem [10, 29]. The TPS is the two-dimensional equivalent of the cubic spline. First developed by Duchon [23], it gets its name from the physical process of bending a thin plate of metal. The TPS is constrained to go through all the data points and is the fit with the least amount of curvature. If the spline function is denoted by $f(x, y)$, and the bending energy function by

$$J = \iint_{\mathbb{R}^2} \left(\left(\frac{\partial^2 f}{\partial x^2} \right)^2 + 2 \left(\frac{\partial^2 f}{\partial x \partial y} \right)^2 + \left(\frac{\partial^2 f}{\partial y^2} \right)^2 \right) dx dy \quad (2.27)$$

the TPS is found by minimising the bending energy function

$$E = \frac{1}{n} \sum_{i=1}^n (f(x_i, y_i) - z_i)^2 + \lambda J \quad (2.28)$$

where z_i are the n data points at coordinates (x_i, y_i) and λ is the smoothing parameter. For $\lambda = 0$ the procedure simply finds the interpolating spline. When $\lambda > 0$ the resulting function is smoothed to reduce the function's curvature. By adjusting the value for λ the amount of smoothing can be controlled. This procedure ensures the TPS agrees with the original data as good as possible (when the TPS is smoothed it does not go through these points exactly), and minimal curvature. Although used in a slightly different setting, this method is similar to the Tikhonov regularisation used in [15, 35].

From the original TPS paper by Duchon [23] and the work by Meinguet [44] it follows that there is a unique solution to this problem and it can be written as

$$f(x_i, y_i) = \sum_{j=1}^n a_j A_{i,j} + \sum_{j=1}^3 b_j B_{i,j} \quad (2.29)$$

Here, A is an $[n \times n]$ matrix and B an $[n \times 3]$ matrix (I is the $[n \times n]$ identity matrix), where n denotes the number of data points. The entries of these matrices are given by

$$\begin{aligned} A_{i,j} &= \|(x_i, y_i) - (x_j, y_j)\|^2 \ln(\|(x_i, y_i) - (x_j, y_j)\|^2) \\ B_{i,(1:3)} &= (1, x_i, y_i) \end{aligned} \quad (2.30)$$

in which $\|\cdot\|$ is the Euclidean norm in \mathbb{R}^2 . Note that to get a correct fit it is usually necessary to scale the variables x and y before the actual fitting, due to the behaviour of the \ln function.

The subsequent work by Wahba and Wendelberg [58] shows, by inserting the expression above into (2.28) and some manipulation, that the vectors \vec{a} and \vec{b} that minimise the bending energy in (2.28) satisfy

$$\begin{aligned} \vec{f} &= (A + n\lambda I)\vec{a} + B\vec{b} \\ 0 &= B^T \vec{a} \end{aligned} \quad (2.31)$$

This set of equations is usually solved by making a QR-decomposition of the matrix B

$$B = QR = (Q_1|Q_2) \begin{pmatrix} R_1 \\ 0 \end{pmatrix} = Q_1 R_1 \quad (2.32)$$

Here Q is an $[n \times n]$ matrix with orthonormal vectors as columns and R is a $[n \times 3]$ upper triangular matrix. Since the bottom $n - 3$ rows of R are zero, it is useful to write Q and R in the form above. Here Q_1 is $[n \times 3]$ and Q_2 is $[n \times (n - 3)]$.

Since $B^T \vec{a} = 0$ it follows from the QR-decomposition that $\vec{a} = Q_2 \gamma$, for some $n - 3$ vector γ . By multiplying the first equation in (2.31) by Q_2^T (and using the orthonormality of Q such that

$$Q_2^T Q_1 = 0)$$

$$\begin{aligned}
Q_2^T f &= Q_2^T (A + n\lambda I) \vec{a} + Q_2^T B \vec{b} \\
Q_2^T f &= Q_2^T (A + n\lambda I) Q_2 \gamma + Q_2^T Q_1 R_1 \vec{b} \\
Q_2^T f &= Q_2^T (A + n\lambda I) Q_2 \gamma \\
\gamma &= (Q_2^T (A + n\lambda I) Q_2)^{-1} Q_2^T f \\
\Rightarrow \vec{a} &= Q_2^T (Q_2^T (A + n\lambda I) Q_2)^{-1} Q_2^T f \\
&= Q_2^T (Q_2^T A Q_2 + n\lambda I)^{-1} Q_2^T f
\end{aligned} \tag{2.33}$$

By multiplying the first equation in (2.31) by Q_1^T

$$\begin{aligned}
Q_1^T f &= Q_1^T (A + n\lambda I) \vec{a} + Q_1^T B \vec{b} \\
Q_1^T f &= Q_1^T A \vec{a} + n\lambda Q_1^T Q_2 \gamma + Q_1^T Q_1 R_1 \vec{b} \\
Q_1^T f &= Q_1^T A \vec{a} + R_1 \vec{b} \\
\Rightarrow \vec{b} &= R_1^{-1} Q_1^T (f - A \vec{a})
\end{aligned} \tag{2.34}$$

Finally the value of the smoothed TPS can be found by inserting these values into

$$f(x, y) = \sum_{i=1}^n a_i \|(x, y) - (x_i, y_i)\|^2 \ln(\|(x, y) - (x_i, y_i)\|^2) + b_0 + b_1 x + b_2 y \tag{2.35}$$

A complete overview of the TPS method is given in [57].

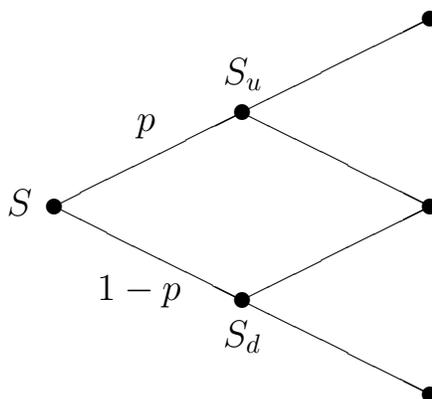
3 Binomial Trees

In this section the binomial tree as a method of pricing options is discussed. The standard version of the binomial tree and a version using the local volatility surface is described in 3.1. The logarithmic binomial tree is discussed in 3.2. To check the accuracy of these methods, the described methods are used for plain vanilla European options. Since the local volatility surface is obtained by using the implied volatility as an input, the right value of these options can be calculated from the Black Scholes equation. These results (which are illustrated in 8.2.3 and 8.2.4) give an indication of the accuracy of the methods used.

3.1 The Standard Binomial Tree

The first binomial tree used for pricing options, and the one that is described in this section, was developed by Cox, Ross and Rubinstein [14]. It models the possible paths of the price of the underlying asset in discrete time. Before considering an adaptation of this model for the case of local volatility, the original model with constant volatility is considered.

Assume that in any fixed time increment Δt the asset can either go up in value to S_u with probability p or down to S_d with probability $1 - p$ as depicted in the figure below



The unknown variables can be derived by equating the discrete time mean and variance of the asset to the values of mean and variance known from continuous time.

3.1.1 Continuous Time Analysis

In continuous time the asset price is assumed to follow geometric Brownian motion (under \mathbb{Q})

$$\frac{dS_t}{S_t} = (r_t - q_t)dt + \sigma dW_t \quad (3.1)$$

where r_t and q_t are the deterministic continuously compounded risk-free interest rate and dividend yield, respectively. From this it follows that the asset price is lognormally distributed since (using Itô)

$$\begin{aligned}
d \ln S_t &= \frac{1}{S_t} dS_t - \frac{1}{2S_t^2} dS_t \cdot dS_t \\
&= (r_t - q_t)dt + \sigma dW_t - \frac{1}{2}\sigma^2 dt \\
&= (r_t - q_t - \frac{1}{2}\sigma^2)dt + \sigma dW_t \\
S_t &= S_0 e^{\int_{t_0}^t (r_s - q_s) ds - \frac{1}{2}\sigma^2(t-t_0) + \sigma W_t}
\end{aligned} \tag{3.2}$$

From this the mean and variance at time node $j + 1$ (time = $t + \Delta t$) as seen from j (time = t) can be derived, where $S_j = S_{t_j}$. It is assumed that within the time increment Δt r_j (r_{t_j}) and q_j (q_{t_j}) remain constant

$$\begin{aligned}
\mathbb{E}[S_{j+1}] &= S_j e^{(r_j - q_j - \frac{1}{2}\sigma^2)\Delta t} \mathbb{E}[e^{\sigma \Delta W_t}] \\
&= S_j e^{(r_j - q_j - \frac{1}{2}\sigma^2)\Delta t} e^{\frac{1}{2}\sigma^2 \Delta t} \\
&= S_j e^{(r_t - q_t)\Delta t} \\
\text{Var}[S_{j+1}] &= S_j^2 e^{2(r_j - q_j - \sigma^2)\Delta t} \text{Var}[e^{\sigma W_t}] \\
&= S_j^2 e^{2(r_j - q_j - \sigma^2)\Delta t} \left(\mathbb{E}[e^{2\sigma \Delta W_t}] - (\mathbb{E}[e^{\sigma \Delta W_t}])^2 \right) \\
&= S_j^2 e^{2(r_j - q_j - \sigma^2)\Delta t} \left(e^{2\sigma^2 \Delta t} - (e^{\frac{1}{2}\sigma^2 \Delta t})^2 \right) \\
&= S_j^2 e^{2(r_j - q_j)\Delta t} (e^{\sigma^2 \Delta t} - 1)
\end{aligned} \tag{3.3}$$

with $\Delta W_t = W_{t+\Delta t} - W_t$. The mean of the asset gives, as could be expected, the regular forward price.

3.1.2 Discrete Time Analysis

In discrete time the mean and variance at time node $j + 1$ (time = $t + \Delta t$) as seen from j (time = t) are

$$\begin{aligned}
\mathbb{E}[S_{j+1}] &= pS_u + (1-p)S_d \\
\mathbb{E}[S_{j+1}^2] &= pS_u^2 + (1-p)S_d^2 \\
\text{Var}(S_{j+1}) &= \mathbb{E}[S_{j+1}^2] - (\mathbb{E}[S_{j+1}])^2 \\
&= pS_u^2 + (1-p)S_d^2 - p^2S_u^2 - 2p(1-p)S_uS_d - (1-p)^2S_d^2 \\
&= p(1-p)(S_u - S_d)^2
\end{aligned} \tag{3.4}$$

3.1.3 Recombining the Binomial Tree

Comparing the expressions for mean and variance in continuous and discrete time gives

$$\begin{aligned} S_j e^{(r_j - q_j)\Delta t} &= pS_u + (1 - p)S_d \\ \Rightarrow p &= \frac{S_j e^{(r_j - q_j)\Delta t} - S_d}{S_u - S_d} \end{aligned} \quad (3.5)$$

$$\begin{aligned} S_j^2 e^{2(r_j - q_j)\Delta t} (e^{\sigma^2 \Delta t} - 1) &= p(1 - p)(S_u - S_d)^2 \\ &= \frac{S_j e^{(r_j - q_j)\Delta t} - S_d}{S_u - S_d} \frac{S_u - S_j e^{(r_j - q_j)\Delta t}}{S_u - S_d} (S_u - S_d)^2 \\ &= (S_j e^{(r_j - q_j)\Delta t} - S_d) (S_u - S_j e^{(r_j - q_j)\Delta t}) \end{aligned} \quad (3.6)$$

A small change in notation is made

$$\begin{aligned} S_u &= uS_j \\ S_d &= dS_j \end{aligned} \quad (3.7)$$

To avoid confusion: dS_j always means d times S_j , an infinitesimal change in the price is always denoted by dS_t .

This transforms equation (3.6) into

$$\begin{aligned} e^{2(r_j - q_j)\Delta t} (e^{\sigma^2 \Delta t} - 1) &= (e^{(r_j - q_j)\Delta t} - d) (u - e^{(r_j - q_j)\Delta t}) \\ &= (u + d)e^{(r_j - q_j)\Delta t} - du - e^{2(r_j - q_j)\Delta t} \\ e^{(2(r_j - q_j) + \sigma^2)\Delta t} &= (u + d)e^{(r_j - q_j)\Delta t} - du \end{aligned} \quad (3.8)$$

The problem has been reduced to solving one equation with two unknowns. This leaves one degree of freedom which is used to make the tree recombining. This can be achieved in numerous ways, one of which is by imposing the following relation between u and d

$$d = u^{-1} \quad (3.9)$$

With this relation, equation (3.8) now becomes

$$e^{(2(r_j - q_j) + \sigma^2)\Delta t} = (u + u^{-1})e^{(r_j - q_j)\Delta t} - 1 \quad (3.10)$$

From which it becomes clear that d will satisfy the exact same equation because of relation (3.9). Therefore both u and d will be solutions of this equation, which is just a simple quadratic equation

$$u^2 - u (e^{-(r_j - q_j)\Delta t} + e^{((r_j - q_j) + \sigma^2)\Delta t}) + 1 = 0 \quad (3.11)$$

This equation has two solutions. By the manner of construction of the binomial tree it easy to see which solution corresponds with u and which one with d

$$\begin{aligned} u &= \beta + \sqrt{\beta^2 - 1} \\ d &= \beta - \sqrt{\beta^2 - 1} \\ \beta &= \frac{1}{2} (e^{-(r_j - q_j)\Delta t} + e^{((r_j - q_j) + \sigma^2)\Delta t}) \end{aligned} \quad (3.12)$$

Expanding the expressions for β and $\sqrt{\beta^2 - 1}$ in Puiseux series and comparing it with the same expressions for $e^{\sigma\sqrt{\Delta t}}$ and $e^{-\sigma\sqrt{\Delta t}}$ shows

$$\begin{aligned}
\beta &= 1 + \frac{1}{2}\sigma^2\Delta t + O(\Delta t^2) \\
\sqrt{\beta^2 - 1} &= \sigma\sqrt{\Delta t} + \frac{1}{2\sigma} \left((r_j - q_j)^2 + (r_j - q_j)\sigma^2 + \frac{3}{4}\sigma^4 \right) \Delta t^{\frac{3}{2}} + O(\Delta t^2) \\
e^{\sigma\sqrt{\Delta t}} &= 1 + \sigma\sqrt{\Delta t} + \frac{1}{2}\sigma^2\Delta t + \frac{1}{6}\sigma^3\Delta t^{\frac{3}{2}} + O(\Delta t^2) \\
e^{-\sigma\sqrt{\Delta t}} &= 1 - \sigma\sqrt{\Delta t} + \frac{1}{2}\sigma^2\Delta t - \frac{1}{6}\sigma^3\Delta t^{\frac{3}{2}} + O(\Delta t^2) \\
\Rightarrow u &= e^{\sigma\sqrt{\Delta t}} + O(\Delta t^{\frac{3}{2}}) \\
d &= e^{-\sigma\sqrt{\Delta t}} + O(\Delta t^{\frac{3}{2}})
\end{aligned} \tag{3.13}$$

So if the terms $O(\Delta t^{\frac{3}{2}})$ are neglected (which is justifiable for Δt small), this results in the approximations for u and d used in [14] that are widely used for the binomial tree.

Since it was assumed that the volatility was constant, u , d and therefore also p are constant at all nodes in the binomial tree. It is thus recombining at every point in time.

3.1.4 Binomial Tree with Local Volatility

When considering the local volatility model, where the volatility is a function of time and the asset price, the situation becomes more complicated. At different nodes the volatility will typically be different. Choosing condition (3.9) at each node will not make the tree recombining, since now going up at one time and then down is not equivalent anymore to going down at that time and then up. This presents a big problem, because for binomial trees to be used effectively it needs to be recombining. Without recombination the number of nodes will be 2^n at time step n . After a mere 100 steps, not enough to make accurate valuations, the number of nodes will be of the order 10^{30} . Implementing this for more time steps, not to mention actually using this tree multiple times, will be computationally impossible. For a recombining tree the number of nodes at time step n will simply be $n + 1$. This means making the binomial tree recombining is vital for the model to be used in practice.

Assuming it is possible to construct a recombining tree and standing at a fixed point in time j , there are $j + 1$ nodes and as many known asset values. At this point there are the following equations (i denotes the number of the node with the top assigned $i = 0$ counting downwards)

$$e^{(2(r_j - q_j) + \sigma_i^2)\Delta t} = (u_i + d_i)e^{(r_j - q_j)\Delta t} - d_i u_i \quad i = 0, \dots, j \tag{3.14}$$

To make sure the tree is recombining the following condition has to be satisfied

$$S_i d_i = S_{i+1} u_i \quad i = 0, \dots, j - 1 \tag{3.15}$$

This means that at each point in time there are $2j + 1$ equations for $2j + 2$ unknowns, which means that typically this system can be solved, leaving one degree of freedom. By the very construction of the tree it is ensured that the tree is recombining.

Using the degree of freedom to take

$$d_k = u_k^{-1} \quad (3.16)$$

at some node k there is a simple algorithm to construct the tree to value European options.

3.1.5 Algorithm

At time step j , there are $j + 1$ nodes, i denotes the number of the node, and σ_i is the local volatility at time j and node i .

1. Calculate u_k and d_k with the original formula for the standard recombining binomial tree (3.12) because of assumption (3.16)
2. Calculate d_{k-1} from the recombining condition (3.15)
3. Calculate u_{k-1} from (3.14)

$$u_{k-1} = \frac{e^{(2(r_j - q_j) + \sigma_i^2)\Delta t} - d_{k-1}e^{(r - q)\Delta t}}{e^{(r_j - q_j)\Delta t} - d_{k-1}} \quad (3.17)$$

4. Repeat steps 2 and 3 for $k - 2, \dots, 0$
5. Calculate u_{k+1} from the recombining condition (3.15)
6. Calculate d_{k+1} from (3.14)

$$d_{k+1} = \frac{u_{k+1}e^{(r_j - q_j)\Delta t} - e^{(2(r_j - q_j) + \sigma_i^2)\Delta t}}{u_{k+1} - e^{(r_j - q_j)\Delta t}} \quad (3.18)$$

7. Repeat steps 5 and 6 for $k + 2, \dots, j$
8. Put payoff values in tree at maturity, work backwards discounting the value of the option, where the transition probabilities are given by (3.5)

3.1.6 Instability of the Binomial Tree

When the binomial tree with local volatility, as described in the previous section, is implemented a major problem is encountered. For low values of the maximum amount of steps n , eg $n < 80$, in the tree the model is relatively stable and seems to work reasonably well. To improve the accuracy of the model however more time steps are needed. When this is done the model becomes very unstable. At some points in the tree the values for u and d start to oscillate (as is shown in 8.2.3).

This oscillation results in the fact that at some nodes u is smaller than d , resulting in irregular trees for the asset price. This behaviour depends on the choice of k in (3.16). When k is chosen to be 1 (the upper node), the tree becomes very unstable for $n > 80$. The optimal choice for k seems to be to take k in the middle of the tree branch $k = \lfloor \frac{n}{2} \rfloor$ (the square brackets denote rounding), where the tree becomes very unstable for around $n > 100$.

For some plain vanilla European options the model seems to give a reasonably well approximation for the option price. However, since the model is unstable, it is impossible to check the accuracy of the model sufficiently. Furthermore, given the irregular behaviour of the tree, it is impossible to predict if this model is in any way useful when dealing with American or exotic options.

3.2 The Logarithmic Binomial Tree

One possible solution to remedy the instability of the standard binomial tree with local volatility may be to use a logarithmic binomial tree where the logarithmic values of the asset price and u and d are used. Maybe the different numerical behaviour of these variables will lead to increased stability of the model.

3.2.1 Continuous Time

From (3.2) follow the equations for the mean and variance in continuous time

$$\begin{aligned} d \ln S_t &= (r_t - q_t - \frac{1}{2}\sigma^2)dt + \sigma dW_t \\ \ln S_t &= \ln S_0 + \int_{t_0}^t (r_s - q_s)ds - \frac{1}{2}\sigma^2(t - t_0) + \sigma W_t \end{aligned} \quad (3.19)$$

or in the language of the binomial tree with time nodes j and $j + 1$

$$\ln S_{j+1} = \ln S_j + (r_j - q_j - \frac{1}{2}\sigma^2)\Delta t + \sigma \Delta W_t \quad (3.20)$$

Giving the mean and variance of the logarithm of the asset price at time node $j + 1$ as seen from j

$$\begin{aligned} \mathbb{E}[\ln S_{j+1}] &= \ln S_j + (r_j - q_j - \frac{1}{2}\sigma^2)\Delta t \\ \text{Var}[\ln S_{j+1}] &= \sigma^2 \Delta t \end{aligned} \quad (3.21)$$

3.2.2 Discrete Time

In discrete time

$$\begin{aligned} \mathbb{E}[\ln S_{j+1}] &= p \ln(uS_j) + (1 - p) \ln(dS_j) \\ &= \ln S_j + p \ln u + (1 - p) \ln d \\ \mathbb{E}[(\ln S_{j+1})^2] &= p \ln^2(uS_j) + (1 - p) \ln^2(dS_j) \\ &= p (\ln u + \ln S_j)^2 + (1 - p) (\ln d + \ln S_j)^2 \\ &= \ln^2 S_j + p \ln^2 u + 2p \ln u \ln S_j + (1 - p) \ln^2 d + 2(1 - p) \ln d \ln S_j \end{aligned} \quad (3.22)$$

Leading to

$$\begin{aligned}
\text{Var}[\ln S_{j+1}] &= \mathbb{E}[(\ln S_{j+1})^2] - (\mathbb{E}[\ln S_{j+1}])^2 \\
&= \ln^2 S_j + p \ln^2 u + 2p \ln u \ln S_j + (1-p) \ln^2 d + 2(1-p) \ln d \ln S_j \\
&\quad - (\ln^2 S_j + 2p \ln S_j \ln u + 2(1-p) \ln S_j \ln d + p^2 \ln^2 u \\
&\quad + 2p(1-p) \ln u \ln d + (1-p)^2 \ln^2 d) \\
&= (p-p^2) \ln^2 u + [(1-p) - (1-p)^2] \ln^2 d - 2p(1-p) \ln u \ln d \\
&= p(1-p)(\ln^2 u - 2 \ln u \ln d + \ln^2 d) \\
&= p(1-p) \ln^2 \frac{u}{d}
\end{aligned} \tag{3.23}$$

So the result of mean matching is

$$\begin{aligned}
(r_j - q_j - \frac{1}{2}\sigma^2)\Delta t &= p \ln u + (1-p) \ln d \\
\Rightarrow p &= \frac{(r_j - q_j - \frac{1}{2}\sigma^2)\Delta t - \ln d}{\ln u - \ln d}
\end{aligned} \tag{3.24}$$

And for variance matching

$$\begin{aligned}
\sigma^2 \Delta t &= p(1-p) \ln^2 \frac{u}{d} \\
&= \left[\frac{(r_j - q_j - \frac{1}{2}\sigma^2)\Delta t - \ln d}{\ln u - \ln d} \right] \left[\frac{\ln u - (r_j - q_j - \frac{1}{2}\sigma^2)\Delta t}{\ln u - \ln d} \right] \ln^2 \frac{u}{d} \\
&= (r_j - q_j - \frac{1}{2}\sigma^2)\Delta t [\ln u + \ln d] - \ln u \ln d - (r_j - q_j - \frac{1}{2}\sigma^2)^2 (\Delta t)^2
\end{aligned} \tag{3.25}$$

When choosing $d = u^{-1} \Rightarrow \ln d = -\ln u$ at some node k , it yields

$$\begin{aligned}
\sigma^2 \Delta t &= \ln^2 u_k - (r_j - q_j - \frac{1}{2}\sigma^2)^2 (\Delta t)^2 \\
\Rightarrow \ln u_k &= \sqrt{\sigma^2 \Delta t + (r_j - q_j - \frac{1}{2}\sigma^2)^2 (\Delta t)^2}
\end{aligned} \tag{3.26}$$

Which gives the same result as before if expanded in a Puiseux series

$$\begin{aligned}
e^{\sqrt{\sigma^2 \Delta t + (r_j - q_j - \frac{1}{2}\sigma^2)^2 (\Delta t)^2}} &= 1 + \sigma \sqrt{\Delta t} + \frac{1}{2} \sigma^2 \Delta t + O\left(\Delta t^{\frac{3}{2}}\right) \\
\Rightarrow u_k &= e^{\sigma \sqrt{\Delta t}} + O\left(\Delta t^{\frac{3}{2}}\right)
\end{aligned} \tag{3.27}$$

For the logarithmic binomial tree with local volatility to be recombining, the following condition needs to be satisfied

$$\ln S_i + \ln d_i = \ln S_{i+1} + \ln u_{i+1} \tag{3.28}$$

3.2.3 Algorithm

The algorithm for the logarithmic binomial tree is not that different from the one for the standard binomial tree.

1. Calculate $\ln u_k$ and $\ln d_k$ for some fixed k from (3.26) and the condition that $\ln d_k = -\ln u_k$
2. Calculate $\ln d_{k-1}$ from the recombining condition (3.28)
3. Calculate $\ln u_{k-1}$ from (3.25)

$$\ln u_{k-1} = (r_j - q_j - \frac{1}{2}\sigma_{k-1}^2)\Delta t - \frac{\sigma_{k-1}^2 \Delta t}{\ln d_{k-1} - (r_j - q_j - \frac{1}{2}\sigma_{k-1}^2)\Delta t} \quad (3.29)$$

4. Repeat steps 2 and 3 for $k - 2, \dots, 0$
5. Calculate $\ln u_{k+1}$ from the recombining condition (3.28)
6. Calculate $\ln d_{k+1}$ from (3.25)

$$\ln d_{k+1} = \frac{\sigma_{k+1}^2 \Delta t}{(r_j - q_j - \frac{1}{2}\sigma_{k+1}^2) - \ln u_{k+1}} + (r_j - q_j - \frac{1}{2}\sigma_{k+1}^2)\Delta t \quad (3.30)$$

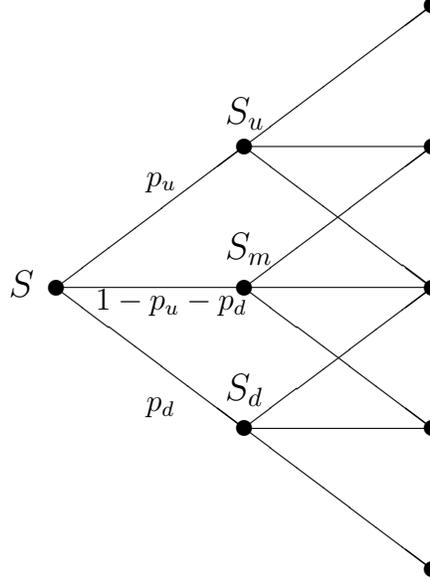
7. Repeat steps 5 and 6 for $k + 2, \dots, j$
8. Put payoff values in tree at maturity, work backwards discounting the value of the option, where the transition probabilities are known by (3.24)

3.2.4 Instability of the Logarithmic Binomial Tree

Unfortunately the logarithmic binomial tree exhibits the same unstable behaviour as the normal binomial tree (although it behaves slightly better, see 8.2.4). In both cases no adjustments can be made to counter the instability while keeping the tree recombining since there is only one degree of freedom at each time step, which needs to establish a relation between two variables at one node. It seems that more degrees of freedom are needed to get a stable model. A model with more degrees of freedom is the trinomial tree model, which is the subject of the next section.

4 Trinomial Trees

The principles of the trinomial tree are the same as those of the binomial tree, the only difference being that at each node there are three possibilities for the next time step, S_u , S_m and S_d , instead of two. The basic equations needed for the construction of the trinomial tree are derived in 4.1. Two different methods of constructing the trinomial tree are described in 4.2 and 4.3. Dividends?



4.1 Basic Equations

Again the main equations are derived by equating the mean and variance to the values obtained from the continuous time analysis. For the trinomial tree at each time the top node is denoted by 0 and, since there are $2j + 1$ nodes at time j , the bottom node is denoted by $2j$.

$$\begin{aligned}
 \mathbb{E}[S_{j+1}] &= p_u S_u + (1 - p_u - p_d) S_m + p_d S_d \\
 \mathbb{E}[S_{j+1}^2] &= p_u S_u^2 + (1 - p_u - p_d) S_m^2 + p_d S_d^2 \\
 \text{Var}[S_{j+1}] &= \mathbb{E}[S_{j+1}^2] - (\mathbb{E}[S_{j+1}])^2 \\
 &= p_u (S_u - S_m)^2 + p_d (S_m - S_d)^2 - (p_u (S_u - S_m) - p_d (S_m - S_d))^2
 \end{aligned} \tag{4.1}$$

Matching mean and variance gives (using the notation of the binomial tree)

$$\begin{aligned}
 \text{For } i &= 0, \dots, 2j \\
 e^{(r_j - q_j)\Delta t} &= p_{u,i} u_i + (1 - p_{u,i} - p_{d,i}) m_i + p_{d,i} d_i \\
 e^{2(r_j - q_j)\Delta t} (e^{\sigma_i^2 \Delta t} - 1) &= p_{u,i} (u_i - m_i)^2 + p_{d,i} (m_i - d_i)^2 \\
 &\quad - (p_{u,i} (u_i - m_i) - p_{d,i} (m_i - d_i))^2
 \end{aligned} \tag{4.2}$$

To make the tree recombining it must be ensured that the following conditions hold

$$\begin{aligned}
\text{For } i &= 0, \dots, 2j - 1 \\
m_i S_i &= u_{i+1} S_{i+1} \\
d_i S_i &= m_{i+1} S_{i+1}
\end{aligned} \tag{4.3}$$

Resulting in $4j + 2$ equations from mean and variance matching and $4j$ equations from the recombining conditions. Since there are $5(2j + 1)$ unknowns, there are $2j + 3$ degrees of freedom. These extra degrees of freedom with respect to the binomial tree will help solve the problems experienced with that model.

The main problem with the binomial tree used with the local volatility model is the instability of the stock tree. At some points the values for $u_{i,j}$ and $d_{i,j}$ began to oscillate, leading to situations where $u_{i,j} < 1$ such that $S_u < S_d$ mangling the shape of the tree (if many time steps are used it can even lead to $u_{i,j} < 0$ leading to negative asset prices). In the trinomial tree this can be remedied by choosing certain values for $m_{i,j}$ (using $2j + 1$ degrees of freedom). It is chosen to keep $m_{i,j}$ constant over all nodes i , simply denoting it by m_j . The time dependency of m_j will be discussed shortly.

Together with the recombining conditions (4.3) this leads to a robust tree, where $S_d < S_m < S_u$ is guaranteed (and negative assets prices are also avoided). Since this fixates almost all the points of the tree, the two remaining degrees of freedom are needed at the first and last node at each point in time.

4.2 Fixing u and d

Since m is already fixed it seems obvious to use the remaining degrees of freedom to fix u and d . Most of these values are already fixed by the recombining conditions and the choice for fixing m_j . Denoting the time step at one node by j , and remembering that by definition $S_{i+1,j+1}$ is at the same height as $S_{i,j}$ and i is counting down from the top to the bottom node, it follows from (4.3) that

$$\begin{aligned}
u_{i+1,j+1} &= \frac{S_{i+1,j+2}}{S_{i+1,j+1}} \\
&= \frac{S_{i,j+1} m_{i,j+1}}{S_{i,j} m_{i,j}} \\
&= u_{i,j} \frac{m_{j+1}}{m_j}
\end{aligned} \tag{4.4}$$

Also at the last node $u_{2j+2,j+1}$ is known by

$$\begin{aligned}
u_{2j+2,j+1} &= \frac{S_{2j+2,j+2}}{S_{2j+2,j+1}} \\
&= \frac{S_{2j,j} m_{2j,j} m_{2j+1,j+1}}{S_{2j,j} d_{2j,j}} \\
&= \frac{m_j m_{j+1}}{d_{2j,j}}
\end{aligned} \tag{4.5}$$

The equations for $d_{i+1,j+1}$ and $d_{0,j+1}$ are comparable.

$$\begin{aligned} d_{i+1,j+1} &= d_{i,j} \frac{m_{j+1}}{m_j} \\ d_{0,j+1} &= \frac{m_j m_{j+1}}{u_{0,j}} \end{aligned} \tag{4.6}$$

So the whole tree can be fixed by using the two remaining degrees of freedom to fix u_0 and d_{2j} at each point in time j .

The transition probabilities are derived from the first of (4.2)

$$\begin{aligned} p_{u,i} &= \frac{e^{(r_j - q_j)\Delta t} + p_{d,i}(m_j - d_i) - m_j}{u_i - m_j} \\ p_{d,i} &= \frac{e^{(2(r_j - q_j) + \sigma_i^2)\Delta t} - e^{(r_j - q_j)\Delta t}(u_i + m_j) + u_i m_j}{(u_i - d_i)(m_j - d_i)} \\ \Rightarrow p_{u,i} &= \frac{e^{(2(r_j - q_j) + \sigma_i^2)\Delta t} - e^{(r_j - q_j)\Delta t}(d_i + m_j) + d_i m_j}{(u_i - m_j)(u_i - d_i)} \\ \Rightarrow p_{m,i} &= 1 - p_{u,i} - p_{d,i} \\ &= \frac{e^{(r_j - q_j)\Delta t}(u_i + d_i) - u_i d_i - e^{(2(r_j - q_j) + \sigma_i^2)\Delta t}}{(u_i - m_j)(m_i - d_i)} \end{aligned} \tag{4.7}$$

4.2.1 Negative Probabilities

From the construction of the trinomial tree it is not guaranteed that the probabilities in (4.7) are between 0 and 1, and therefore may not be actual probabilities. Since all u 's and d 's are essentially fixed from the start, it is obvious they are chosen in a way such that $u_i > m_i > d_i$. This means that all the denominators in (4.7) are larger than zero. Temporarily denoting the numerator of $p_{d,i}$ by N , ensuring that $p_{d,i} \geq 0$ means $N \geq 0$. Using that $e^{\sigma_i^2 \Delta t} \geq 1$, because $\sigma_i \geq 0$, gives

$$\begin{aligned} N &= e^{(2(r_j - q_j) + \sigma_i^2)\Delta t} - e^{(r_j - q_j)\Delta t}(u_i + m_j) + u_i m_j \\ &\geq e^{2(r_j - q_j)\Delta t} - e^{(r_j - q_j)\Delta t}(u_i + m_j) + u_i m_j \\ &= \left(e^{(r_j - q_j)\Delta t} - u_i \right) \left(e^{(r_j - q_j)\Delta t} - m_j \right) \end{aligned} \tag{4.8}$$

So if it is chosen to let $m_j = e^{(r_j - q_j)\Delta t}$, this means $N \geq 0$ and thus $p_{d,i} \geq 0$. Expressions (4.7) can thus be written as

$$\begin{aligned} p_{u,i} &= \frac{e^{2(r_j - q_j)\Delta t} \left(e^{\sigma_i^2 \Delta t} - 1 \right)}{\left(u_i - e^{(r_j - q_j)\Delta t} \right) (u_i - d_i)} \\ p_{d,i} &= \frac{e^{2(r_j - q_j)\Delta t} \left(e^{\sigma_i^2 \Delta t} - 1 \right)}{(u_i - d_i) \left(e^{(r_j - q_j)\Delta t} - d_i \right)} \\ p_{m,i} &= \frac{e^{(r_j - q_j)\Delta t}(u_i + d_i) - u_i d_i - e^{(2(r_j - q_j) + \sigma_i^2)\Delta t}}{\left(u_i - e^{(r_j - q_j)\Delta t} \right) \left(e^{(r_j - q_j)\Delta t} - d_i \right)} \end{aligned} \tag{4.9}$$

From this it becomes clear that this choice of m_j also ensures that $p_{u,i} \geq 0$. However, no definitive statement can be made about the upper bound of these probabilities, since they depend on the level of the volatility. So some extra feature is needed in the model to ensure that $p_{m,i} > 0$.

Negative probabilities only occur for the transition to the middle node $p_{m,i}$, since $p_{u,i} \geq 0$ and $p_{d,i} \geq 0$. The trinomial tree is fixed from the start, and the values for u and d will be chosen to make sure no negative probabilities occur at the origin of the tree (see section 4.2.2). But as the tree progresses in time, the volatility changes, since it depends on the asset price and the time. When the volatility rises, the tree becomes too rigid to accommodate the increased volatility. The price process needs high values for u and low values for d to reflect the high volatility. But u and d are fixed, so the probabilities of going up and down grow, as can be seen in (4.9), giving rise to negative probabilities for the middle node.

To fix this problem u and d in the extreme nodes can be made time dependent. This is not sufficient though when the local volatility also rises with time (as it does in the SABR model, see 8.2.1) and due to the choice of m the asset rises for the middle nodes. It will still result in some negative transition probabilities. It also seems too much compensation to change u and d for those parts of the tree which behave normally.

A procedure is thus used to provide more flexibility at the points in the tree with high volatility. The tree is constructed according to the formulae in the previous section. At each point in time, before continuing to the next time step, the tree is checked for negative values of p_m . If such a negative value is encountered at the first or last node, the most extreme values of the asset, u_0 or d_{2j} is increased or decreased respectively, after which the tree is adjusted and probabilities are recalculated. If a negative probability is encountered at one of the inner nodes, the asset will not go up one step, stay in the middle or down one step, but instead will go up two steps, stay in the middle or down two steps. u and d are adjusted accordingly and the probabilities are recalculated. This effectively changes the tree (for all inner nodes) into a recombining quintinomial tree (five possible transitions at each node) where the probabilities of going up two steps and down two steps, p_{uu} and p_{dd} , are zero, but when negative probabilities are encountered, those probabilities are adjusted and the probabilities of going one step up or down are set to zero.

4.2.2 Determining u_0 at the Origin of the Tree

As mentioned in the previous section, at the origin of the tree the values for u and d will be chosen to make sure no negative probabilities occur. p_m in (4.9) needs to be larger than or equal to zero. This is guaranteed when

$$0 \leq e^{(r_0 - q_0)\Delta t}(u_0 + d_0) - u_0 d_0 - e^{(2(r_0 - q_0) + \sigma_0^2)\Delta t} \quad (4.10)$$

Since the values of both u_0 and d_0 at this point can be chosen freely, it is ensured that the

above inequality is satisfied. For instance if $d_0 = \frac{m_0^2}{u_0} = \frac{1}{u_0} e^{2(r_0 - q_0)\Delta t}$ then

$$\begin{aligned}
0 &\leq u_0, e^{(r_0 - q_0)\Delta t} + \frac{1}{u_0} e^{3(r_0 - q_0)\Delta t} - e^{2(r_0 - q_0)\Delta t} - e^{(2(r_0 - q_0) + \sigma_0^2)\Delta t} \\
0 &\leq u_0^2 - u_0 e^{(r_0 - q_0)\Delta t} \left(e^{\sigma_0^2 \Delta t} + 1 \right) + e^{2(r_0 - q_0)\Delta t} \\
\Rightarrow u_0 &\geq u_{\min} := \frac{1}{2} e^{(r_0 - q_0)\Delta t} \left(e^{\sigma_0^2 \Delta t} + 1 + \sqrt{\left(e^{\sigma_0^2 \Delta t} + 1 \right)^2 - 4} \right)
\end{aligned} \tag{4.11}$$

The other solution for u_0 would require it to be less than a value which is typically less than 1, and is therefore discarded.

Different choices for u_0 can be made as long as (4.11) is satisfied. A large value will make the tree wider, ensuring that negative probabilities will occur as little as possible. However, it gives rise to more inaccuracy. A possible reason for this can be that when the tree is wider there will be fewer distinct values at maturity in that part that matters most: the middle part where the probability of ending up in that part of the tree is largest. Since these values are the most likely to occur and therefore contribute the most to the option value, having fewer of these value should lead to more imprecision. Therefore, to get the best result, the choice is made to put $u_0 = u_{\min}$.

4.2.3 Implementing a Cutoff for High Volatilities

Another aspect that has to be considered is what happens in the tree when the asset price becomes very large. In some models, including the SABR model depending on the specific parameters, the local volatility is increasing in the asset price. In the trinomial tree, as indeed is also the case in the binomial tree, the values at the top of the tree become very large. If the local volatility surface is extrapolated to these prices, large volatilities can occur. This leads to negative probabilities for the middle nodes even after adjusting u by large amounts. This causes serious delay in the computational time and leads to a general breakdown of the model and should thus be avoided.

The choice is made to implement a cutoff value for the asset dependency of the local volatility. Above some value for the asset price, for the calculation of the local volatility it is assumed the asset price is equal to this cutoff value. This removes the asset dependency above the cutoff value, while keeping the time dependency. A natural question is if this can be done and if it will compromise the accuracy of the model. To answer this question the effect of this cutoff value is determined by implementing the trinomial tree for a large variety of cutoff values and strike prices (with maturity of 1 year and 1000 time steps). This results in two observations:

- Placing the cutoff value below $3S_0$ results in significant variations of the order of a few cents.
- For a cutoff value above $3S_0$, the cutoff value does not change the modeled price by more than a tenth of a cent in the most extreme case.

These observations suggest that as long as the cutoff value is implemented at a relatively high value ($\geq 3S_0$) the cutoff does not affect the option price significantly and therefore will not cause any significant problems.

4.2.4 Algorithm

At the origin of the tree:

1. Fix $u_0 = u_{\min}$ such that (4.11) is satisfied. Then fix $m_0 = e^{(r_0 - q_0)\Delta t}$ and $d_0 = \frac{m_0^2}{u_0}$
2. Calculate the transition probabilities from (4.9)
3. Calculate the tree at time step 1

For all time steps $j \geq 1$:

4. Fix $u_{0,j}$ and $d_{2j,j}$. The other u 's and d 's are known from (4.4), (4.5) and (4.6)
5. Calculate tree values for time $j + 1$
6. Calculate the transition probabilities from (4.9)
7. Check for negative probabilities at each node $i = 0, \dots, 2j$
8. If $p_{m,0,j} < 0$
 - Increase $u_{0,j}$
 - Start over from step 5
9. If $p_{m,2j,j} < 0$
 - Decrease $d_{2j,j}$
 - Start over from step 5
10. If $p_{m,i,j} < 0$ for $i \notin \{0, 2j\}$
 - $uu_{i,j} = \frac{S_{i-1,j+1}}{S_{i,j}}$
 - $dd_{i,j} = \frac{S_{i+3,j+1}}{S_{i,j}}$
 - Set $p_{u,i,j} = p_{d,i,j} = 0$
 - Calculate $p_{uu,i,j}$, $p_{m,i,j}$ and $p_{dd,i,j}$ from (4.9) with $uu_{i,j}$ and $dd_{i,j}$ instead of $u_{i,j}$ and $d_{i,j}$
 - Start over from step 5
11. Increase time step and repeat steps 4-10 as long as the time of maturity is not reached
12. Put payoff values in the tree at maturity, work backwards discounting the value of the option (let V denote the value of the option)

$$V_{i,j} = e^{-r_j \Delta t} (p_{uu,i,j} V_{i-1,j+1} + p_{u,i,j} V_{i,j+1} + p_{m,i,j} V_{i+1,j+1} + p_{d,i,j} V_{i+2,j+1} + p_{dd,i,j} V_{i+3,j+1}) \quad (4.12)$$

4.3 Fixing Probabilities

The choice to be made at the end of 4.1 was what to do with the two remaining degrees of freedom, to evaluate the most extreme nodes. In the previous section these degrees were used to fix u and d . Although this is the most logical choice since the other degrees of freedom were used to fix the values for m , it is not perfect, since it sometimes results in negative transition probabilities for the middle nodes. Another possibility for using the two remaining degrees of freedom is to fix the transition probabilities at these extreme nodes. This will lead to slightly more complicated formulas but since it puts restrictions on the probabilities and lets u or d adjust in these points, this will sometimes lead to fewer problems with negative probabilities.

These probabilities are fixed by (omitting the time subscript)

$$p_{u,0} = p_{d,2j} = p_f \quad (4.13)$$

4.3.1 Origin of the Tree

At the first node of the tree there are no recombining conditions. Equations (4.2) need to be solved for u_0 and d_0 . Both probabilities are fixed since $2j + 1 = 1$. The first of these equations gives

$$d_0 = \frac{e^{(r_j - q_j)\Delta t} - p_{u,0}u_0 - (1 - p_{u,0} - p_{d,0})m_j}{p_{d,0}} \quad (4.14)$$

Plugging this into the second equation

$$\begin{aligned} & e^{2(r_j - q_j)\Delta t} \left(e^{\sigma_0^2 \Delta t} - 1 \right) \\ &= p_{u,0}(u_0 - m_j)^2 + p_{d,0} \left(m_j - \frac{e^{(r_j - q_j)\Delta t} - p_{u,0}u_0 - (1 - p_{u,0} - p_{d,0})m_j}{p_{d,0}} \right)^2 \\ & \quad - \left[p_{u,0}(u_0 - m_j) - \left(m_j p_{d,0} - (e^{(r_j - q_j)\Delta t} - p_{u,0}u_0 - (1 - p_{u,0} - p_{d,0})m_j) \right) \right]^2 \\ &= p_{u,0}(u_0 - m_j)^2 + \frac{1}{p_{d,0}} \left(p_{u,0}(u_0 - m_j) + \left(m_j - e^{(r_j - q_j)\Delta t} \right) \right)^2 - \left(e^{(r_j - q_j)\Delta t} - m_j \right)^2 \\ &= p_{u,0}(u_0 - m_j)^2 - \left(e^{(r_j - q_j)\Delta t} - m_j \right)^2 \\ & \quad + \frac{1}{p_{d,0}} \left[p_{u,0}^2 (u_0 - m_j)^2 + 2p_{u,0}(u_0 - m_j)(m_j - e^{(r_j - q_j)\Delta t}) + (m_j - e^{(r_j - q_j)\Delta t})^2 \right] \\ &= \left(p_{u,0} + \frac{p_{u,0}^2}{p_{d,0}} \right) (u_0 - m_j)^2 + 2\frac{p_{u,0}}{p_{d,0}} \left(u_0 m_j - u_0 e^{(r_j - q_j)\Delta t} - m_j^2 + m_j e^{(r_j - q_j)\Delta t} \right) \\ & \quad + \frac{1 - p_{d,0}}{p_{d,0}} (e^{(r_j - q_j)\Delta t} - m_j)^2 \end{aligned} \quad (4.15)$$

For the choice $m_j = e^{(r_j - q_j)\Delta t}$ this gives

$$m_j^2 \left(e^{\sigma_0^2 \Delta t} - 1 \right) = \left(p_{u,0} + \frac{p_{u,0}^2}{p_{d,0}} \right) (u_0 - m_j)^2 \quad (4.16)$$

Or

$$\frac{p_{d,0}}{p_{u,0}} m_j^2 \left(e^{\sigma_0^2 \Delta t} - 1 \right) = (p_{u,0} + p_{d,0})(u_0^2 - 2u_0 m_j + m_j^2) \quad (4.17)$$

which leads to the following quadratic equation for u_0

$$\begin{aligned} 0 &= au_0^2 + bu_0 + c \\ a &= p_{u,0} + p_{d,0} \\ b &= -2m_j(p_{u,0} + p_{d,0}) \\ c &= m_j^2(p_{u,0} + p_{d,0}) - \frac{p_{d,0}}{p_{u,0}} m_j^2 \left(e^{\sigma_0^2 \Delta t} - 1 \right) \end{aligned} \quad (4.18)$$

This equation has two solutions. Since one of these solution is less than one, it is discarded, leading to the final equation's for u_0 and d_0 (using that at the origin of the tree $p_{u,0} = p_{d,0} = p_f$)

$$\begin{aligned} u_0 &= \frac{-b + \sqrt{b^2 - 4ac}}{2a} \\ &= m_j + m_j \sqrt{\frac{\left(e^{\sigma_0^2 \Delta t} - 1 \right)}{2p_f}} \\ d_0 &= \frac{e^{(r_j - q_j)\Delta t} - p_{u,0}u_0 - (1 - p_{u,0} - p_{d,0})m_j}{p_{d,0}} \\ &= m_j - m_j \sqrt{\frac{\left(e^{\sigma_0^2 \Delta t} - 1 \right)}{2p_f}} \end{aligned} \quad (4.19)$$

4.3.2 Rest of the Tree

When the tree is at the time step denoted by j most of the u 's and d 's are fixed because of the choice of fixing all values of m_j and the recombining conditions (4.3). The equations for $u_{i+1,j+1}$, $u_{2j,j+1}$, $d_{i+1,j+1}$ and $d_{0,j+1}$ are equal to those in section 4.2. So once u and d are calculated at one timestep in the tree, the values for u and d are all known except for u_0 and d_{2j} (omitting the time indices).

At the first node at time j there are two equations for the two unknowns u_0 and $p_{d,0}$ (since $p_{u,0} = p_f$)

$$\begin{aligned} e^{(r_j - q_j)\Delta t} &= p_f u_0 + (1 - p_f - p_{d,0})m_j + p_{d,0}d_0 \\ \Rightarrow p_{d,0} &= \frac{m_j - e^{(r_j - q_j)\Delta t} + p_f(u_0 - m_j)}{m_j - d_0} \end{aligned} \quad (4.20)$$

and

$$\begin{aligned}
e^{2(r_j - q_j)\Delta t} \left(e^{\sigma_0^2 \Delta t} - 1 \right) &= p_f(u_0 - m_j)^2 + p_{d,0}(m_j - d_0)^2 - (p_f(u_0 - m_j) - p_{d,0}(m_j - d_0))^2 \\
&= p_f(u_0^2 - 2u_0m_j + m_j^2) + (m_j - e^{(r_j - q_j)\Delta t} + p_f(u_0 - m_j))(m_j - d_0) \\
&\quad - \left(p_f(u_0 - m_j) - [m_j - e^{(r_j - q_j)\Delta t} + p_f(u_0 - m_j)] \right)^2 \\
&= p_f(u_0^2 - u_0m_j - u_0d_0 + m_jd_0) \\
&\quad + \left(m_j - e^{(r_j - q_j)\Delta t} \right) (m_j - d_0) - \left(e^{(r_j - q_j)\Delta t} - m_j \right)^2
\end{aligned} \tag{4.21}$$

Or for $m_j = e^{(r_j - q_j)\Delta t}$

$$\frac{m_j^2}{p_f} \left(e^{\sigma_0^2 \Delta t} - 1 \right) = u_0^2 - u_0(m_j + d_0) + m_jd_0 \tag{4.22}$$

Resulting again in a quadratic equation for u_0 , where only the value bigger than 1 is considered a valid solution

$$\begin{aligned}
0 &= au_0^2 + bu_0 + c \\
a &= 1 \\
b &= -(m_j + d_0) \\
c &= m_jd_0 - \frac{m_j^2}{p_f} \left(e^{\sigma_0^2 \Delta t} - 1 \right) \\
\Rightarrow u_0 &= \frac{-b + \sqrt{b^2 - 4ac}}{2a}
\end{aligned} \tag{4.23}$$

The same analysis holds for d_{2j} where $p_{d,2j} = p_f$ and u_{2j} are known

$$\begin{aligned}
e^{(r_j - q_j)\Delta t} &= p_{u,2j}u_{2j} + (1 - p_{u,2j} - p_f)m_j + p_{d,2j}d_{2j} \\
\Rightarrow p_{u,2j} &= \frac{e^{(r_j - q_j)\Delta t} - m_j + p_f(m_j - d_{2j})}{u_{2j} - m_j}
\end{aligned} \tag{4.24}$$

$$\begin{aligned}
e^{2(r_j - q_j)\Delta t} \left(e^{\sigma_{2j}^2 \Delta t} - 1 \right) &= p_{u,2j}(u_{2j} - m_j)^2 + p_f(m_j - d_{2j})^2 \\
&\quad - (p_{u,2j}(u_{2j} - m_j) - p_f(m_j - d_{2j}))^2 \\
&= \left(e^{(r_j - q_j)\Delta t} + p_f(m_j - d_{2j}) - m_j \right) (u_{2j} - m_j) \\
&\quad + p_f(m_j^2 - 2m_jd_{2j} + d_{2j}^2) - (e^{(r_j - q_j)\Delta t} - m_j)^2 \\
&= p_f(d_{2j}^2 - d_{2j}m_j + m_ju_{2j} - d_{2j}u_{2j}) \\
&\quad + \left(e^{(r_j - q_j)\Delta t} - m_j \right) (u_{2j} - m_j) - (e^{(r_j - q_j)\Delta t} - m_j)^2
\end{aligned} \tag{4.25}$$

Or for $m_j = e^{(r_j - q_j)\Delta t}$

$$\frac{m_j^2}{p_f} \left(e^{\sigma_{2j}^2 \Delta t} - 1 \right) = d_{2j}^2 - d_{2j}(m_j + u_{2j}) + m_ju_{2j} \tag{4.26}$$

The only viable solution for this equation is the one that is less than one

$$\begin{aligned}
0 &= ad_{2j}^2 + bd_{2j} + c \\
a &= 1 \\
b &= -(m_j + u_{2j}) \\
c &= m_j u_{2j} - \frac{m_j^2}{p_f} \left(e^{\sigma_{2j}^2 \Delta t} - 1 \right) \\
\Rightarrow d_{2j} &= \frac{-b - \sqrt{b^2 - 4ac}}{2a}
\end{aligned} \tag{4.27}$$

The transition probabilities can then be found by the equations (4.9), giving all the necessary equations for building the trinomial tree and valuing options on it.

4.3.3 Negative Probabilities and Cutoff

When the last two degrees of freedom for the trinomial tree were used to fix the values for u and d and the extreme nodes, at some nodes negative transition probabilities appeared for the middle transition. The suspicion was that this was caused by a lack of flexibility in the tree. The tree with fixed probabilities at the extreme nodes is more flexible and can therefore better handle the changes in the local volatility.

One consequence is that the values for the asset at the top of the tree increase significantly. This does not seem to result in bigger inaccuracies of the option price and is therefore not deemed to be important. Another consequence is that when the tree is used with the SABR implied volatility surface as an input 8.2.1, it results in negative probabilities far fewer times than when u and d were fixed. To be more precise: only for relatively high values for p_{fixed} do negative probabilities arise. For plain vanilla options with strike price $K = S_0$ and $T = 1$ this maximum acceptable value for p_{fixed} is 45.7% for 100 time steps and declines almost linearly to 44.9% for 1000 time steps. Below these values no negative probabilities arise, resulting in a ‘cleaner’ model. However, this behaviour depends on the shape of the implied volatility surface, so no definite preference for one of both versions of the tree can be given.

The exact choice of p_{fixed} does not affect the price significantly. Only at very low values (less than 10%) and few time steps in the tree (less than 200) do significant deviations occur. The choice is made to let $p_{\text{fixed}} = 30\%$. There is no significant difference in computational time between the model where u and d are fixed and when the probabilities are fixed.

When implementing this model for the trinomial tree again a cutoff value for the local volatility is used as described in 4.2.3.

For a numerical comparison between the two versions of the tree see 8.2.5. Numerical results describing the accuracy of the tree developed in this section for valuing European options can be found in 8.2.6. American options are considered in 8.2.8.

4.3.4 Algorithm

1. At the first node of the tree p_u and p_d are fixed in advance. u_0 and d_0 can be calculated according to (4.19)
2. At the next time steps the u 's and d 's are given by (4.4), (4.5) and (4.6)
3. The two remaining values $u_{0,j+1}$ and $d_{2j+2,j+1}$ can be calculated by (4.23) and (4.27)
4. Since the tree at time step $j + 1$ is now completely fixed the transition probabilities can be calculated from (4.9)
5. Repeat steps 2.-4. until the tree is complete
6. Put the payoff of the option to be priced in a separate tree and use the transition probabilities to calculate the discounted expected payoff of the tree

5 Dividends

Up until now the effect of dividends on option pricing has been ignored. Since most assets on which options are written do pay out dividends it is important to incorporate this into the pricing model. Dividends are payments by the issuer of a financial asset (the asset underlying the option) to the holder of the asset. In financial modelling it is usually assumed that the payment of a dividend amount leads to a decline in the asset's price by the same amount. This relies on the assumption of a frictionless market in which the transfer of dividends does not encounter obstacles such as taxes or transaction costs. It also assumes that, when considering equity shares, a certain amount of cash on the balance sheet of a company is valued at exactly the same value as when the amount is on the shareholder's bank account.

5.1 Kinds of Dividends

Although the word 'dividend' is only used for shares, equivalent principles hold for other asset classes. The regular interest payments, known as coupons, to bondholders, act in exactly the same way. For currency exchange options, the interest rate of the foreign currency behaves like a dividend payment [39].

For practical reasons there is usually a difference between the date when for all trading purposes the asset is assumed to pay out the dividend, the ex-dividend date, and the actual time of payment. After the ex-dividend date the holder of the underlying asset does not have a claim to the declared dividend. In this report the ex-dividend date is used as the payment date.

Three different forms of dividends can be distinguished: continuous, proportional and discrete.

Continuous dividends are the easiest to deal with. As can be seen from the formulae in this section, the presence of a continuous dividend yield usually amounts in a simple reduction of the interest rate. For options on equity shares this situation never occurs. There is no company in the world that pays out continuous dividends. However it does occur for options on currency exchange rates. Since foreign interest rates can be continuous, continuous dividend yields will occur for currency options .

Proportional dividends assume that paid out dividends are proportional to the underlying asset's price at the ex-dividend date. This is easy to implement, since it results in a reduction in the value of the asset price at each node, by a fixed percentage. A radically different share prices could change the decision by a listed company concerning its dividend policy, providing some logical basis for the assumption of proportional dividends. However, in most cases dividend payments are announced in advance and subsequent changes in the share price do not affect the payments. Therefore in most models, discrete dividends are used.

Discrete dividends describe the real-world situation for options on equity shares. The timing and size of dividends are usually well known in advance. The size of the dividend is not a given deterministic function of the share price, such as with proportional dividends. But due to the assumptions stated above, some dependency between the two does exist. Given the assumption

that at the ex-dividend date the share price declines by the dividend size some assumption has to be made for the occasion when the dividend size is equal or larger than the share price. Negative share prices do not occur in reality, due to the limited liability provisions applicable to publicly held companies. Therefore the assumption is made, as it is in [56], that the dividend size is equal to the minimum of the stated dividend and the share price. When the share price falls below the dividend size at the ex-dividend date, it is assumed the company only pays out the amount equal to the share price, thereby reducing the share price to zero.

5.2 Direct Implementation

Discrete dividends add complexity to the pricing of options, that do not arise for continuous and proportional dividends. First consider the traditional Black-Scholes framework with a constant volatility term. When working with binomial trees the u and d are fixed at the beginning and are, because the volatility is constant, constant over the whole tree. Incorporating a fall in the share price at a specific point in time, reflecting the dividend payment, destroys the recombining property of the tree. This is, for reasons given in 3.1.4, unacceptable. The problem arises because from the dividend date forward, different values for u and d should be used from to make the tree recombining, which is not possible in this framework.

Arguably, more flexibility is present when using trinomial trees with local volatility. Some small adjustments in (4.4) and the equivalent equation for d can take care of the impact discrete dividends have on the recombining property of the tree. Since recombination in the trinomial tree is imposed, this should not be a problem. Still, discrete dividends cannot be implemented in this way. The reason is that for low values of the share price, which are always present in the lower region of the tree, the discrete dividend will cause the share price to go to zero. Since the share price increases to a higher node by multiplying by u , this means that when the share price hits zero, it will stay zero (a straightforward consequence of the adjusted geometric Brownian motion framework). This means that this part of the tree cannot possibly recombine with the non-zero part of the tree. Therefore another method must be used to implement the impact of discrete dividends on the option price.

5.3 Dividend Adjustments to Model

In the literature different methods are proposed on how to deal with discrete dividends.

One approach, known as the escrowed model, is to separate the dividends from the asset price [39]. Since the discrete dividends are deterministic, they can be valued simply by discounting their value by the risk-free interest rate. Then the tree is built with the spot price of the asset minus the present value of the future dividends at the start of the tree. As the escrowed model brings the dividends back to the current time, the forward model pushes the dividends forwards to the maturity date. In the forward model the value of the dividends are compounded by the risk-free interest rate and added to the strike value of the option [47]. A combination of the escrowed and the forward model is proposed in [9], where the dividends closest to the current date are brought

backwards and the other ones are pushed forwards.

It seems these models work relatively well in the standard Black-Scholes environment for European options, but they suffer from some logical flaws as is pointed out in [36, 56]. In the escrowed model the spot price of the underlying asset is adjusted by the dividends during the life of the option. This results in a different price process than the one used in the forward model. It is therefore not surprising that they lead to different values for path-dependent options [31]. Furthermore it follows that different maturities imply different price processes, but it is clear that the price dynamics of the asset do not depend on the particular options depending on it. A remedy for this would be to include all future dividends, but, as pointed out in [56], this is also unsatisfactory from a logical point of view since changes in dividends after the option's maturity should not affect the option price.

In a local volatility model modifying the spot price cannot be justified, since the local volatility surface is calculated with the current spot price. By adjusting the spot price it is not clear if the same local volatility surface can be used. In making the decision if it can be used, an assumption has to be made on how the surface changes under spot price moves. There is no reason to assume the surface depends on the asset's spot price, but there is no guarantee that it doesn't.

Since both the escrowed and the forward model compress all the dividends in one point they do not sufficiently incorporate the effects that the dividends will have on early exercise decisions. It is well known that early exercise for American call options is only profitable just before the dividend dates [39]. By compressing all the dividends into one point, it removes the possible moments of early exercise from the tree. So American calls, and by similar reasoning puts, cannot be accurately priced using the escrowed or forward model.

5.4 The Vellekoop-Nieuwenhuis Method

The Vellekoop-Nieuwenhuis model [56], does not adjust the price process of the asset, but instead adjusts the discounting process of the option on the tree. The tree is built without any dividends. The payoff can then be discounted in the normal way up until the last dividend date (which is the first that is encountered when working backwards). At these nodes, whose asset values are written collectively as S , the price of the option is a function of the asset price, $f(S)$. But this is without dividends. At these nodes the asset price declines by the value of the dividend, d . So the correct prices at these nodes are actually $f(S - d)$. Assuming that f is a continuous function, these values can be approximated by interpolating the values $f(S)$ to $f(S - d)$. Here it is assumed that the asset price never becomes negative, so actually $f(S)$ is interpolated onto $f(\max(S - d, 0))$. Cubic splines are used as the method of interpolation, in contrast to [56] where linear interpolation is used. To avoid extrapolation the points $f(S)$ are augmented by the point $f(0)$. Since an asset price with value zero will always remain zero, this value can be easily determined for both the call and the put, both European and American. Then the option value can be discounted in the regular way until a new dividend is encountered, at which point the same procedure is performed.

In the original article binomial trees are used, but, as is explicitly mentioned there, the method can also be applied to trinomial trees. A property that does not carry over unto the local volatility

model, is the convergence proof. One of the assumptions used in proving convergence, is that in the absence of dividends uniform convergence exists. This is the case for trees with constant volatility, for which uniform convergence is proven [37]. A convergence result does not exist for trees with local volatility (yet), so convergence of the price obtained by the tree, to the true option price, is not guaranteed.

5.5 Interest Rates

At all points in the preceding sections the interest rate was said to be a known deterministic function of time. However this is not the way in which interest rates are observed in the market. Although it is easy to derive the interest rate function from market data, for completeness it is included in this report.

Interest rates are observed indirectly in the market, they are derived from the price of risk-free bonds. A zero-coupon bond is a claim on a specific amount of cash in the future, without the payment of intermediary amounts. Since it is risk free, the price is the amount discounted by the risk-free interest rate

$$p(t_0, T) = e^{-R(t_0, T)(T-t_0)} \quad (5.1)$$

where $p(t_0, T)$ is the bond price at time t_0 for a claim of 1 unit of cash at time T . $R(t_0, T)$ is the bond's yield, which can be easily calculated from bond prices. When this rate is used for other purposes than describing a bond, it is referred to as the spot interest rate for a certain maturity. Equal bonds with different maturities have different prices, and therefore different yields. Because the yield is a function of the time to maturity, it is usually called the yield curve. The time evolution of the short interest rate can be derived from this yield curve.

The price of a risk-free bond can be written as the expected value of the discounted value of the 1 unit of cash at maturity. So when interest rates are deterministic

$$\begin{aligned} p(t_0, T) &= \mathbb{E}[e^{-\int_{t_0}^T r_s ds} | \mathcal{F}_0] \\ &= e^{-\int_{t_0}^T r_s ds} \end{aligned} \quad (5.2)$$

By combining this equation with (5.1)

$$\begin{aligned} \int_{t_0}^T r_s ds &= R(t_0, T)(T - t_0) \\ r_T &= \frac{\partial R(t_0, T)}{\partial T} (T - t_0) + R(t_0, T) \\ r_t &= \frac{\partial R(t_0, y)}{\partial y} \Big|_{y=t} (t - t_0) + R(t_0, t) \end{aligned} \quad (5.3)$$

6 Delta Hedging

An essential practice for the option trading industry is hedging. The simplest version, delta hedging, is the subject of this section. In 6.1 delta hedging in the classical Black-Scholes framework is presented. The minimum variance hedge is given in 6.2 and the hedge given by the local volatility model, the local delta, is considered in 6.3.

6.1 Black-Scholes Framework

The most commonly known technique to ensure a trader with a long or short position of an option against fluctuations in the price of the underlying asset is delta hedging. The total portfolio is made delta-neutral by adding a certain amount of the asset. Thus if Π is the value of the portfolio and Black-Scholes is assumed

$$\Pi = V_{BS} - S\Delta_{BS} \quad (6.1)$$

Then, the portfolio is hedged only at the point when the option position is taken, the delta of the option can be calculated by equating the total delta of the portfolio to zero

$$\Delta_{BS} = \frac{\partial V_{BS}}{\partial S} \quad (6.2)$$

This is the classic result, which holds in the Black-Scholes environment and would hold in the real markets if all prices were described by geometric Brownian motion. But when other models are used such as the local volatility model things change. When trying to fit specific model variables α the following equation must be satisfied

$$V(S, t, K, T) = V_{BS}(S, t, K, T, \alpha) \quad (6.3)$$

Calculating the derivative of this equation results in

$$\begin{aligned} \frac{\partial V}{\partial S} &= \frac{\partial V_{BS}}{\partial S} + \frac{\partial V_{BS}}{\partial \alpha} \frac{\partial \alpha}{\partial S} \\ &= \Delta_{BS} + \frac{\partial V_{BS}}{\partial \alpha} \frac{\partial \alpha}{\partial S} \end{aligned} \quad (6.4)$$

Which is just a generalisation of the Black-Scholes model since in that model the second term in the last equation would equal zero. The point this equation makes, is that different models lead to different delta hedges.

6.2 Minimum Variance Delta

Observing the statements above, [1] suggests defining a minimum variance (MV) delta that minimises the quadratic covariation of the changes in the delta hedged portfolio Π with the changes in the underlying asset. Setting this equal to zero

$$\begin{aligned} 0 &= d\Pi \cdot dS \\ &= (dV - dS\Delta_{MV}) \cdot dS \\ \Rightarrow \Delta_{MV} &= \frac{dV \cdot dS}{dS \cdot dS} \end{aligned} \quad (6.5)$$

Considering the local volatility model and using Itô gives

$$\begin{aligned}
dV &= \frac{\partial V}{\partial S} dS + \frac{\partial V}{\partial t} dt + \frac{\partial V}{\partial \sigma} d\sigma + \frac{1}{2} \frac{\partial^2 V}{\partial S^2} dS \cdot dS \\
&\quad + \frac{1}{2} \frac{\partial^2 V}{\partial \sigma^2} d\sigma \cdot d\sigma + \frac{\partial^2 V}{\partial S \partial \sigma} dS \cdot d\sigma \\
d\sigma &= \frac{\partial \sigma}{\partial S} dS + \frac{\partial \sigma}{\partial t} dt + \frac{1}{2} \frac{\partial^2 \sigma}{\partial S^2} dS \cdot dS \\
d\sigma \cdot d\sigma &= \frac{\partial \sigma}{\partial S} dS \cdot dS \\
dS \cdot d\sigma &= \frac{\partial \sigma}{\partial S} dS \cdot dS \\
\Rightarrow dV \cdot dS &= \frac{\partial V}{\partial S} dS \cdot dS + \frac{\partial V}{\partial \sigma} d\sigma \cdot dS \\
&= \left(\frac{\partial V}{\partial S} + \frac{\partial V}{\partial \sigma} \frac{\partial \sigma}{\partial S} \right) dS \cdot dS
\end{aligned} \tag{6.6}$$

this results in

$$\Delta_{MV} = \frac{\partial V}{\partial S} + \frac{\partial V}{\partial \sigma} \frac{\partial \sigma}{\partial S} \tag{6.7}$$

This is an important adjustment to the normal delta, which is the first term on the right-hand side of the equation above. It has been noticed that applying the normal delta for delta hedging purposes leads to inaccurate values of the delta [5, 12, 22]. It seems that to get good delta hedges in the local volatility model, some sort of vega hedge (the sensitivity with respect to the volatility) has to be incorporated to get good results [27]. These ad-hoc alterations are messy and seem to come from nowhere. But (6.7) suggests that this vega adjustment should have been incorporated all along.

6.3 Local Delta

As remarked in [16, 24, 55], volatility is usually negatively correlated to the asset price. Since plain vanilla options have positive vega's, this would suggest that in general $\Delta_L \leq \Delta_{BS}$. Although the MV delta is presented in a nice analytic form, for the local volatility model it cannot be used in this way. The partial derivatives of the option value used in (6.7) are not known. Therefore the delta for the local volatility delta will always be calculated numerically.

One of the criticisms of local volatility models is that it implies dynamics of the implied volatility surface that are not compatible with market practices [4, 34]. This behaviour is illustrated when the local volatility function is time independent and solely a function of the asset price. In [34] Hagan derived, by perturbation techniques, that the implied volatility at time t_0 with asset price S_0 can then be approximated by

$$\Sigma(K, T; S_0) = \sigma \left(\frac{1}{2} [S_0 + K] \right) \tag{6.8}$$

which implies that as the spot of the asset moves up by ΔS

$$\Sigma(K, T; S_0 + \Delta S) = \sigma \left(\frac{1}{2} [(S_0 + \Delta S) + K] \right) = \sigma \left(\frac{1}{2} [S_0 + (K + \Delta S)] \right) = \Sigma(K + \Delta S, T; S_0) \tag{6.9}$$

Suggesting that as the spot price moves up the implied volatility curve moves to the left, which, according to Hagan, is contrary to the experience of market participants. For this reason it is assumed that local volatility models give wrong greeks and therefore provide wrong hedges.

The reasoning of the critique above is, however, only justified if certain ‘stickiness’ assumptions hold, highlighting the importance of being aware of the different possible assumptions that can be made.

6.3.1 Stickiness Assumptions

The term ‘sticky’ to describe the dynamics of the implied volatility curve (the cross section of the implied volatility surface for a given maturity) was introduced in [17]. In general three different stickiness assumptions can be made

- If the implied volatility curve does not change when the spot price of the underlying asset changes the curve is said to be sticky-strike

$$\Sigma(K, T; S_0) = \Sigma(K, T; S_0 + \Delta S) \quad (6.10)$$

- If the implied volatility curve does not change when the spot price of the underlying asset changes but the moneyness ($\frac{S}{K}$) does not change the curve is said to be sticky-moneyness.

$$\Sigma(K, T; S_0) = \Sigma(K + \frac{K}{S_0}\Delta S, T; S_0 + \Delta S) \quad (6.11)$$

- If the local volatility surface is assumed to be fixed, then the model is said to be model consistent

$$\begin{aligned} \sigma(S, t; S_0) &= \sigma(S, t; S_0 + \Delta S_0) \\ \Rightarrow \Sigma(K, T; S_0) &\approx \Sigma(K + \Delta S, T; S_0 - \Delta S) \end{aligned} \quad (6.12)$$

From this it can be seen that Hagan’s critique only holds when one assumes that in reality the implied volatility is sticky moneyness. However as pointed out in [27] and [49] this may not be the best assumption in every market.

Sections 6.1 and 6.2 gave a theoretical description of calculating delta’s. In practice delta’s are usually calculated numerically. The simplest approach is by taking the central difference approximation

$$\Delta(S_0) = \frac{\partial V(S_0)}{\partial S} \approx \frac{V(S_0 + \Delta S) - V(S_0 - \Delta S)}{2\Delta S} \quad (6.13)$$

where ΔS is some small change in S_0 . This approach, however, does not recognise the observations made about different stickiness assumptions. If a change in S_0 changes the local volatility surface, the new perturbed option values ($V(S_0 + \Delta S)$ and $V(S_0 - \Delta S)$) should be calculated using the new local volatility surface. A more accurate numerical calculation of $\Delta(S_0)$ would thus be (borrowing notation from [27])

$$\begin{aligned} \Delta(S_0) \approx \frac{1}{2\Delta S} [&V(S_0 + \Delta S, \sigma(\Sigma(K + \alpha\Delta S, T; S_0 + \beta\Delta S); S_0 + \beta\Delta S) \\ &- V(S_0 - \Delta S, \sigma(\Sigma(K - \alpha\Delta S, T; S_0 - \beta\Delta S); S_0 - \beta\Delta S))] \end{aligned} \quad (6.14)$$

The sticky-strike assumption corresponds to $\alpha = 0$, $\beta = 1$, sticky-moneyness to $\alpha = \frac{K}{S_0}$, $\beta = 1$ and model consistent to $\alpha = 0$, $\beta = 0$ (no change in the local volatility surface). Which one of these assumptions should be made is not clear a priori. If the local volatility function captures the actual diffusion process for the asset price, then it should remain constant, indicating the model consistent stickiness assumption is justified. There are, however, many indications that different markets justify different stickiness assumptions [49]. [16] suggests that for equity markets and equity indices the model consistent assumption is the most natural.

6.3.2 Performance of Local Delta

The earliest comparison of the performance of local delta hedging versus Black-Scholes (or implied) delta hedging, suggests that the traditional BS delta performs best [24]. Most subsequent research [12, 13, 16, 55] concludes the opposite, although [16] provides a more nuanced picture. Since the local delta captures local information, the local delta should provide a better hedge. Given the multiple adjustments that are necessary to make the BS delta work accurately [5, 12, 22], the BS delta does not inspire much confidence as a good hedge. The numerical results presented in section 8.4.4 present a mixed picture.

7 Monte Carlo Simulations

The principle method of pricing options presented in this report is the trinomial tree. The reason why Monte Carlo simulations are included is to compare the outcomes of the trinomial tree to a reference value. For standard European options this reference value can be both the Black-Scholes formula, which provides a closed form formula for the option prices, and Monte Carlo simulations. For American options only the latter is available to generate reference values.

In Monte Carlo simulations the dynamics of the asset price under the risk neutral measure \mathbb{Q} as given by (1.8) are simulated. By sampling a great number of paths an approximation is made for the actual dynamics of the underlying asset. By calculating what the value of a particular option would be under many sample paths and then averaging the result, an approximation is made for the expected discounted value of the payoff, thus giving an estimate for the option price under consideration. The standard method is described in 7.1-7.3. Although the standard Monte Carlo method can only be used for European options, it can be modified to accommodate other options (such as American and exotic options). To do this the method developed by Longstaff and Schwartz [43] is used, which is described in 7.4.

7.1 Creating Sample Paths

The dynamics as given in (1.8) are made by fixing S_0 and discretising the equation into

$$\begin{aligned} S_{t+\Delta t} &= S_t (1 + (r_t - q_t)\Delta t + \sigma(t, S_t)\Delta W_t) \\ \Delta W_t &= W_{t+\Delta t} - W_t \sim N(0, \Delta t) = \sqrt{\Delta t} N(0, 1) \\ S_{t+\Delta t} &= S_t \left(1 + (r_t - q_t)\Delta t + \sigma(t, S_t) \epsilon_t \sqrt{\Delta t} \right) \end{aligned} \tag{7.1}$$

where $\epsilon_t \sim N(0, 1)$ is distributed according to the standard normal distribution and independent samples are used at different times.

7.2 Antithetic Sampling

The simulation of the function ϵ_t in (7.1) is the essential part in Monte Carlo simulation, since it is the only source of randomness, all the other values are deterministic. Thus making sure the sampling of this function as accurate as possible is of the utmost importance. A method that is used often to create better estimates, and which reduces the variance of the estimate, is antithetic samples.

The standard normal distribution is symmetric. If ϕ is the probability density function of the standard normal distribution this means that $\phi(x) = \phi(-x)$. Thus when ϵ_t is simulated, it is equally likely to generate a sample value X_i as it is to get $-X_i$. To make ensure this property holds exactly during the simulation, everytime a sample X_i is generated, the value $-X_i$ (the antithetic sample) is used in another sample path, obtaining twice the number of sample paths.

Calculating the estimate for the option value and the confidence interval, relies on the Strong Law of Large Numbers and the Central Limit Theorem. Let Y_i be an option value calculated using

one sample path, and let \tilde{Y}_i be the value calculated by the antithetic path. Then Y_i and \tilde{Y}_i are not independent from each other, but each pair (Y_i, \tilde{Y}_i) is independent and identically distributed. This means that the average can be defined as

$$\bar{Y} = \frac{1}{2n} \left(\sum_{i=1}^n Y_i + \sum_{i=1}^n \tilde{Y}_i \right) = \frac{1}{n} \sum_{i=1}^n \left(\frac{Y_i + \tilde{Y}_i}{2} \right) \quad (7.2)$$

where n is the number of original samples that are generated and $\frac{Y_i + \tilde{Y}_i}{2}$ are iid. So by the Law of Large Numbers

$$\bar{Y} \rightarrow \mathbb{E}[Y] \quad \text{as } n \rightarrow \infty \quad (7.3)$$

This ensures that eventually the average will converge to the actual option value.

By the Central Limit Theorem

$$\frac{\bar{Y} - \mathbb{E}[Y]}{\sigma/\sqrt{n}} \rightarrow N(0,1) \quad \text{as } n \rightarrow \infty \quad (7.4)$$

$$\sigma^2 = \text{Var}(\bar{Y})$$

Denoting the cumulative density function of the normal distribution by Φ , the 95% confidence interval will thus be given by

$$\bar{Y} \pm \Phi^{-1}(0.975) \frac{\sigma}{\sqrt{n}} \quad (7.5)$$

Since Y_i and \tilde{Y}_i are negatively correlated, the variance of \bar{Y} is smaller than if $2n$ samples of Y_i would have been considered. As such antithetic sampling is known as a variance reduction technique. Other variance reduction techniques include control variates, stratified sampling and importance sampling [33].

7.3 European Options

By calculating the payoff at maturity and discounting, the value of a standard European option can be determined. This is a simple procedure, requiring few programming skills. The drawback is that a large amount of simulations (in the order of millions) have to be performed to get accurate results. The reason for this is that to get an accurate estimate for the price of an option, the probability density distribution of the asset price at maturity has to be determined. Both in the tree and in Monte Carlo simulations this is the essence of the method. In the tree however, many different possible outcomes are determined, each of which is assigned a probability value. In Monte Carlo simulations this probability is determined by looking at many different simulations. To accurately approximate the probability density distribution at maturity, the number of simulations has to be of such a high order that highly unlikely outcomes are also represented in the sample. The result of which is that although the concept of Monte Carlo simulations is simple, the execution can be time consuming.

Numerical results for valuing European options by Monte Carlo simulations can be found in 8.2.7.

7.4 American Options

Since a European option can only be exercised at maturity, calculating the payoff is easy. For an American option the situation is more complicated, since it can be exercised at each moment in time before maturity. Simulations for each infinitesimal time period is not practical, so the time points at which the American option can be exercised is limited to the discrete points in time used in the simulation. In effect a Bermudan option is simulated, which is used as an approximation to the American option.

A further complication is that it is usually not known when it is optimal to exercise the option. To determine whether it is optimal to exercise the option at a particular point in time the immediate exercise value should be compared to the expected value of continuing with the option, conditional on the current state. So if j denotes the time step and V_j is the option value at that time

$$V_j = \max(e^{r_j \Delta t} \mathbb{E}[V_{j+1} | \mathcal{F}_j], E_j) \quad (7.6)$$

where E_j is the immediate exercise value. For the tree this was easy to calculate, since the value $\mathbb{E}[V_{j+1} | \mathcal{F}_j]$ was known when discounting backwards from the transition probabilities. During a Monte Carlo simulation this is not the case. An approximation has to be made. The method used here is the commonly used Longstaff-Schwartz method [43]. At maturity the payoff of the option is known. Working backwards the expected value of continuing with the option (the continuation value) is estimated from the discounted cash flows at later times. It is assumed that the continuation value can be estimated to be a linear combination of functions on the information at time j

$$\mathbb{E}[V_{j+1} | \mathcal{F}_j] = \sum_{i=1}^N a_i \psi_i(S_j) \quad (7.7)$$

where ψ_i are a set of deterministic functions, referred to as basis functions.

By taking a set of sample paths and regressing the discounted cash flows of future times onto the basis functions the coefficients a_i can be estimated. This is usually done in a least squares sense (the method is therefore usually denoted as the Least Squares Monte Carlo Method or LSM). This way an estimate of the continuation value can be made for each sample path, giving a criterion for early exercise by comparing this value to the immediate exercise value.

7.4.1 Bias and Convergence of LSM

By approximating the American option with a Bermudan option a bias is introduced into the estimation of the option value. Since a Bermudan option has less possibilities of early exercise, its value is always less or equal to an American option with the same specifications. This low bias from sub-optimal exercise is mentioned in the original Longstaff-Schwartz article and it is suggested that the estimated option value can never be larger than the true option value because of this. This, however, is not all there is to it.

As noted in [30, 33] there is also a high bias. The LSM approximates the expected value of the option at a later time conditional on the current information, by regressing the continuation value

on certain basis functions. These continuation values however, are obtained from information on the future paths of the simulations. This seems like a clear violation of the measurability of the conditional expectation (V_j should be \mathcal{F}_j -measurable, but now information from t_{j+1} is used). By considering the future cash flows a superoptimal exercise strategy is employed. As a result the bias of the method is mixed. One could hope that the positive and negative bias somehow cancel each other out, but there is no clear indication that this will indeed happen.

The LSM is similar to the method proposed by Tsitsiklis and Van Roy [53, 54], where the value function is approximated instead of the continuation value, but Longstaff and Schwartz decided to use only those sample paths that were in the money for the regression, improving the efficiency of the algorithm. Convergence of LSM is studied in [11, 52, 60] where almost sure convergence of the algorithm is shown as the number of basis functions and sample paths used go to infinity. This is proven for the algorithm given in the original article, which assumes constant volatility. It is thus not clear whether this method will converge to the real value of the option for a non-flat local volatility surface. Increasing the amount of samples will usually reduce the variance and thus shows some convergence, but this does not necessarily improve the accuracy of the estimate.

Numerical results for valuing American options by Monte Carlo simulations can be found in 8.2.8.

8 Numerical Results

8.1 Finite Differences Approximation for Derivatives

If an analytic expression exists for a differentiable function, the derivative can be calculated analytically. But in most cases either there exists no such analytic expression or the expression is so complicated that taking derivatives renders it impossible to get an analytic expression (as is the case in the SABR model, see section 8.2.1). The finite differences approach is a numerical method often used to get approximating values for derivative values of a function for which values are known on a fixed grid. The essence of this method is to approximate the given function with a ordinary polynomial of order n . The factors of this polynomial can then be determined by considering $n + 1$ values of the given function. The higher the order of the approximation, the higher the accuracy, but the more complicated the formulae. In this report the approximations are of the order 4.

Assume an equidistantly spaced grid with interpoint distances of Δx , on which the function $u(x)$ is known at the grid points. For 4th order approximations, 5 points need to be considered, u_0, u_1, u_2, u_3 and u_4 , placed at $x_0 = 0, x_1 = \Delta x, x_2 = 2\Delta x, x_3 = 3\Delta x$ and $x_4 = 4\Delta x$ respectively. The polynomial approximation for the function and the first two derivatives is

$$\begin{aligned} u(x) &\approx c_0 + c_1x + c_2x^2 + c_3x^3 + c_4x^4 \\ u'(x) &\approx c_1 + 2c_2x + 3c_3x^2 + 4c_4x^3 \\ u''(x) &\approx 2c_2 + 6c_3x + 12c_4x^2 \end{aligned} \tag{8.1}$$

the first of which gives the following expressions for the known values of u

$$\begin{aligned} u_0 &\approx c_0 \\ u_1 &\approx c_0 + c_1\Delta x + c_2(\Delta x)^2 + c_3(\Delta x)^3 + c_4(\Delta x)^4 \\ u_2 &\approx c_0 + 2c_1\Delta x + 4c_2(\Delta x)^2 + 8c_3(\Delta x)^3 + 16c_4(\Delta x)^4 \\ u_3 &\approx c_0 + 3c_1\Delta x + 9c_2(\Delta x)^2 + 27c_3(\Delta x)^3 + 81c_4(\Delta x)^4 \\ u_4 &\approx c_0 + 4c_1\Delta x + 16c_2(\Delta x)^2 + 64c_3(\Delta x)^3 + 256c_4(\Delta x)^4 \end{aligned} \tag{8.2}$$

This system of five equations with five unknowns (the u 's and Δx are known) can be solved to yield expressions for the polynomial constants

$$\begin{aligned} c_0 &\approx u_0 \\ c_1 &\approx \frac{-25u_0 + 48u_1 - 36u_2 + 16u_3 - 3u_4}{12\Delta x} \\ c_2 &\approx \frac{35u_0 - 104u_1 + 114u_2 - 56u_3 + 11u_4}{24(\Delta x)^2} \\ c_3 &\approx \frac{-5u_0 + 18u_1 - 24u_2 + 14u_3 - 3u_4}{12(\Delta x)^3} \\ c_4 &\approx \frac{u_0 - 4u_1 + 6u_2 - 4u_3 + u_4}{24(\Delta x)^4} \end{aligned} \tag{8.3}$$

Once inserted into (8.1), these equations result in the approximations for the derivatives of u with

respect to the variable x in the points x_0, x_1, x_2, x_3 and x_4 .

$$\begin{aligned}
u'(x_0) &\approx \frac{-25u_0 + 48u_1 - 36u_2 + 16u_3 - 3u_4}{12\Delta x} \\
u''(x_0) &\approx \frac{35u_0 - 104u_1 + 114u_2 - 56u_3 + 11u_4}{12(\Delta x)^2} \\
u'(x_1) &\approx \frac{-3u_0 - 10u_1 + 18u_2 - 6u_3 + u_4}{12\Delta x} \\
u''(x_1) &\approx \frac{11u_0 - 20u_1 + 6u_2 + 4u_3 - u_4}{12(\Delta x)^2} \\
u'(x_2) &\approx \frac{u_0 - 8u_1 + 8u_3 - u_4}{12\Delta x} \\
u''(x_2) &\approx \frac{-u_0 + 16u_1 - 30u_2 + 16u_3 - u_4}{12(\Delta x)^2} \\
u'(x_3) &\approx \frac{-u_0 + 6u_1 - 18u_2 + 10u_3 + 3u_4}{12\Delta x} \\
u''(x_3) &\approx \frac{-u_0 + 4u_1 + 6u_2 - 20u_3 + 11u_4}{12(\Delta x)^2} \\
u'(x_4) &\approx \frac{3u_0 - 16u_1 + 36u_2 - 48u_3 + 25u_4}{12\Delta x} \\
u''(x_4) &\approx \frac{11u_0 - 56u_1 + 114u_2 - 104u_3 + 35u_4}{12(\Delta x)^2}
\end{aligned} \tag{8.4}$$

When approximating the values of derivatives in a gridpoint, it should be the most accurate when an equal amount of points on both sides are used for this procedure (i.e. the values for x_2). These approximations are usually referred to as central differences. Since the derivation of equations above only depends on the relative position of the x values with respect to each other and not on their absolute position, the formulas hold on the entire grid. Therefore in all points on the grid the central differences are used, except in the cases where there are not two gridpoints on both sides of the point in which we are making approximations. In these boundary points the formulas with x_1 and x_3 are used when there is one gridpoint on one side of the point and the those with x_0 and x_4 when it is a true boundary point (i.e. on the edge of the grid).

Note that the variable x in these derivations is arbitrary, such that the formulae hold for derivatives with respect any variable, where the derivatives are made in a any direction on the grid. For higher order derivative approximations the expressions can be easily deduced from the expressions above. Mixed derivatives are slightly more complicated and deserve some more work. Since higher order and mixed derivatives are not used in this report the formulae for these are not derived here.

8.2 SABR Model

As was seen in 2.3, to derive the local volatility surface implied volatilities are needed as an input. To check how well the methods presented in this report work, at first it is checked in an environment where the implied volatility surface is known.

8.2.1 Theoretical Framework

In [34] the Stochastic- $\alpha\beta\rho$, or shortly SABR, model was derived in which the volatility function was assumed to be stochastic. The dynamics of the SABR model are given by

$$\begin{aligned} dF_t &= \alpha_t F_t^\beta dW_t^1 \\ d\alpha_t &= \nu \alpha_t dW_t^2 \\ dW_t^1 \cdot dW_t^2 &= \rho dt \end{aligned} \quad (8.5)$$

Here F_t is the forward price, α_t the volatility function with starting value $\alpha_0 = \alpha$, ρ is the correlation between the two Brownian motions and ν is the volatility of the volatility.

The SABR model has the advantage that there exists an analytic expression for the implied volatility as a function of the strike price K and the forward price $f = F_0 = Se^{(r_0 - q_0)T}$. The best result would be a precise analytic expression for the local volatility function. In the SABR model the implied volatility is closely approximated by

$$\begin{aligned} \Sigma(f, K) &\approx \alpha(fK)^{\frac{\beta-1}{2}} \left[1 + \frac{(1-\beta)^2}{24} \ln^2\left(\frac{f}{K}\right) + \frac{(1-\beta)^4}{1920} \ln^4\left(\frac{f}{K}\right) \right]^{-1} \left(\frac{z}{x(z)} \right) \\ &\left[1 + \left(\frac{(1-\beta)^2}{24} \frac{\alpha^2}{(fK)^{1-\beta}} + \frac{1}{4} \frac{\rho\beta\nu\alpha}{(fK)^{\frac{(1-\beta)}{2}}} + \frac{2-3\rho^2}{24} \nu^2 \right) \tau \right] \end{aligned} \quad (8.6)$$

where

$$\begin{aligned} z &= \frac{\nu}{\alpha} (fK)^{\frac{(1-\beta)}{2}} \ln\left(\frac{f}{K}\right) \\ x(z) &= \ln\left(\frac{\sqrt{1-2\rho z + z^2} + z - \rho}{1-\rho}\right) \end{aligned} \quad (8.7)$$

which is derived in [34].

Since an analytic, closed form formula for the implied volatility exists, plugging it into (2.25) will result in an expression for the local volatility. However, it is easy to see that this formula would be huge. Particularly the derivative with respect to T is large, rendering it impossible to work with. Programs such as Maple cannot handle it. Even the simpler approximating formula

$$\Sigma(f, K) \approx \frac{\alpha}{f^{1-\beta}} \left(1 - \frac{1}{2} [1 - \beta - \rho\lambda] \ln\left(\frac{K}{f}\right) + \frac{1}{12} [(1-\beta)^2 + (2-3\rho^2)\lambda^2] \ln^2\left(\frac{K}{f}\right) \right) \quad (8.8)$$

proves to be unworkable for the same reasons. Therefore it is necessary to focus on numerical approximations of (2.25).

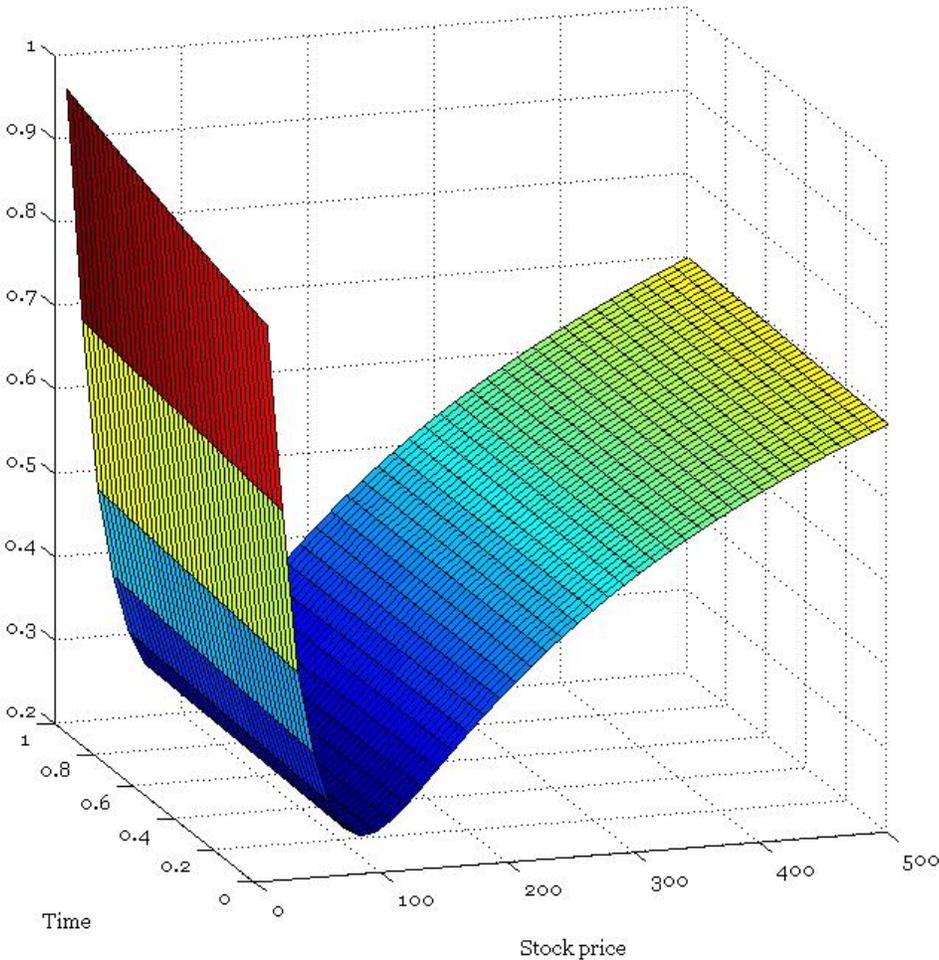
A different approximation for the implied volatility surface is made in [6]. Just like (8.6) this is an asymptotic expansion of the implied volatility surface with an error of $O(\Delta\tau^2)$. From [48] it becomes clear that the SABR model is better approximated by the function derived in [6] and

specified in [42]. Although this gives a slight improvement of the model, it makes things significantly more complicated. Formula (8.6) is quite compact. The formula derived in [42] fills two pages if written in an efficient manner. Implementing this would be time-consuming and increases the chance of making errors. Since in this report an analytic input for the implied volatility surface is only used for stylised examples, in which the actual input is quite arbitrary (the goal is to work with market data as the input), the slight improvement is deemed to not to be worth the effort so (8.6) will be used.

8.2.2 Volatility Surface

The local volatility surface derived from (2.25) if (8.6) is used as an input, with $\alpha = 0.4$, $\beta = 0.9$, $\rho = 0.3$, $\nu = 0.4$ for an option with $T = 1$, $S_0 = K = 100$, $r = 0.05$, $q = 0$

Local Volatility

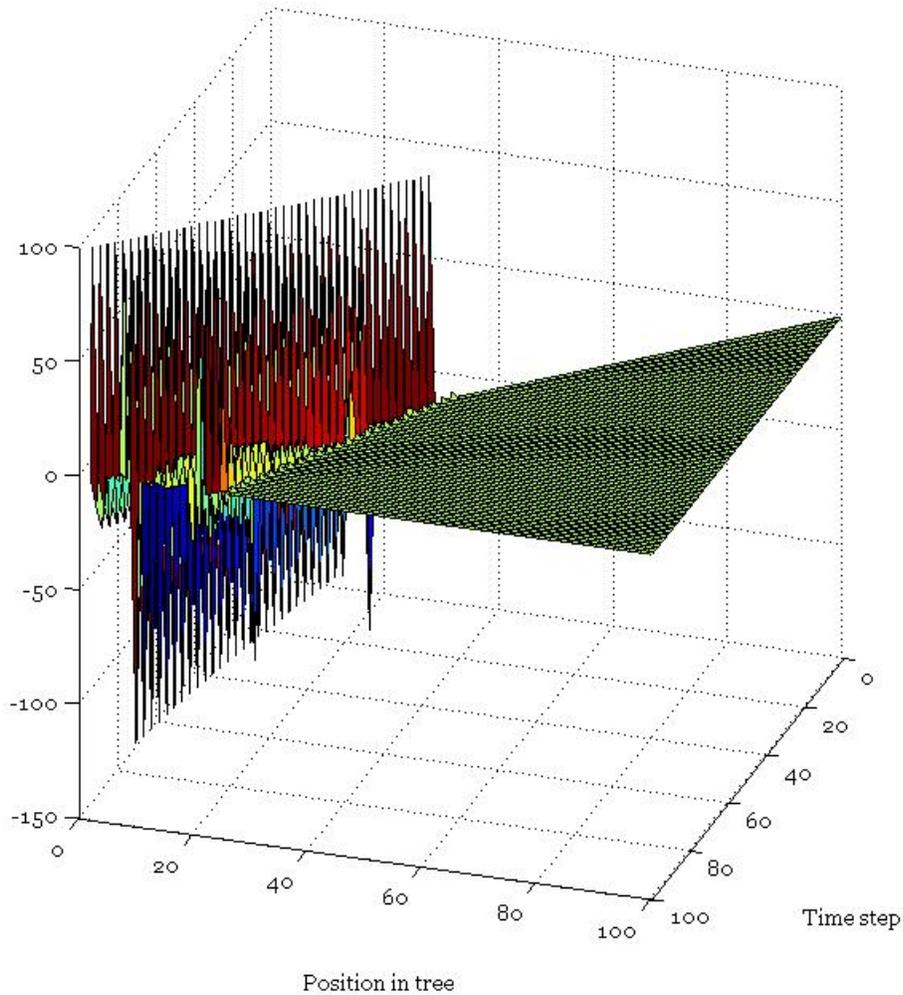


The local volatility exhibits a strong dependency on the underlying asset price. It goes to infinity as the asset price approaches zero as can be seen from (8.6). The time dependency is less pronounced, but the volatility is increasing in time, especially for small values of S .

8.2.3 Binomial Tree Instability

In section 3.1.6 the inherent instability of the binomial tree, as used with local volatility, was described. The main problem is that the value for u , d and p become highly irregular. As an example the different values of u as they appear in a binomial tree with 100 time steps are plotted below as a function of the position in the tree. The underlying asset has initial value $S_0 = 100$ and the local volatility surface used is the one shown in 8.2.2. Since the values of u vary wildly the values are capped at 100 on the positive side (actual values attained were of the order 10^{20}).

Values of u on tree

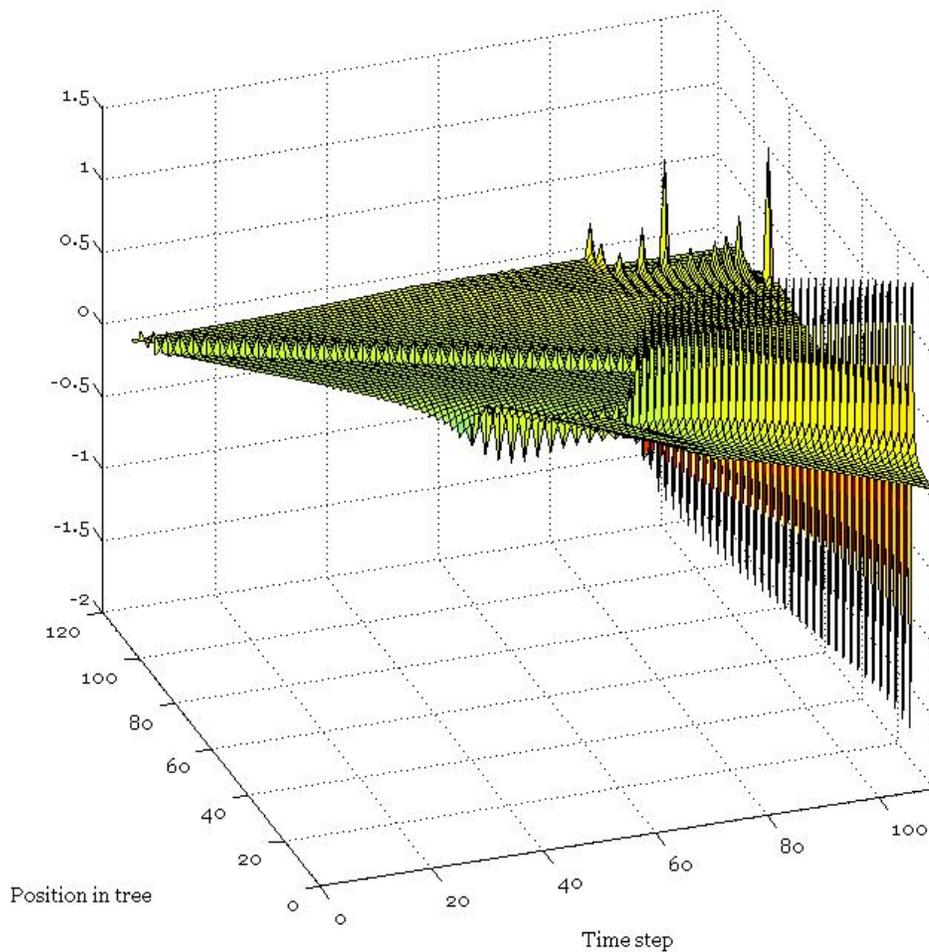


It should be noted that although this tree seems highly unstable (negative value for u are obviously unrealistic), for European options it did result in prices closely resembling the Black-Scholes prices (differences of a few cents). For more time steps, needed to increase the accuracy, this was not the case.

8.2.4 Logarithmic Binomial Tree Instability

When the exact same local volatility as given in 8.2.2 is used for the logarithmic binomial tree, the same instability arises (as described in 3.2.4). It does behave slightly better than the regular binomial tree, as can be seen in the graph below, which did not need to be capped for large values as was needed for the regular binomial tree. However, the large oscillations in the values for u and the occurrence of negative values are clearly unacceptable in any model that tries to describe reality.

Values of u on tree



8.2.5 Comparison Between Fixing u and d and Fixing Probabilities

When constructing the trinomial tree in section 4 a choice had to be made what to do with the degrees of freedom. Most were used to fix the increment of making the middle transition, for the remaining degrees two possible choices were presented, fixing u and d in 4.2, and fixing probabilities in 4.3. Both versions of the tree were said to have similar accuracies. The numerical comparison is given in the table below.

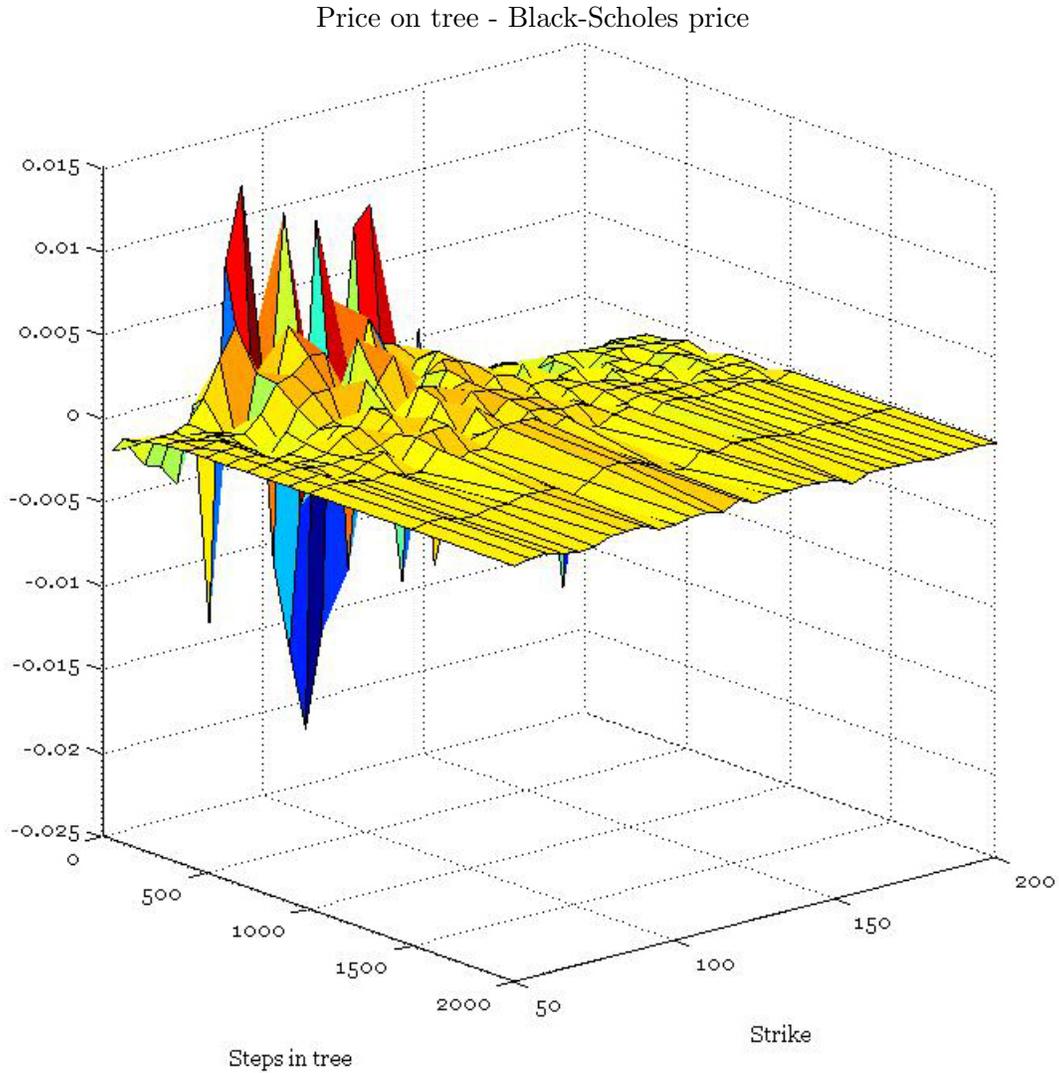
For different values of the number of steps in the tree a European call and put options with $S_0 = K = 100$ are priced using both trees and compared to the Black-Scholes price. The differences between the trees are equal for both call and put and are therefore only mentioned once per number of time steps.

Black-Scholes: Call = 12.4707, Put = 7.5936						
Time Steps	Call/Put	Tree p_f	Tree u and d	Diff. Trees	Diff./BS	
50	C	12.4597	12.4692	$-9.46 \cdot 10^{-3}$.0758	%
	P	7.5826	7.5921		.1245	%
100	C	12.4809	12.4824	$-1.55 \cdot 10^{-3}$.0124	%
	P	7.6038	7.6054		.0204	%
200	C	12.4711	12.4710	$9.94 \cdot 10^{-5}$.0008	%
	P	7.5941	7.5940		.0013	%
300	C	12.4734	12.4650	$8.32 \cdot 10^{-3}$.0667	%
	P	7.5963	7.5880		.1096	%
400	C	12.4694	12.4707	$-1.21 \cdot 10^{-3}$.0097	%
	P	7.5924	7.5936		.0160	%
500	C	12.4726	12.4724	$1.88 \cdot 10^{-4}$.0015	%
	P	7.5956	7.5954		.0025	%
600	C	12.4713	12.4727	$-1.33 \cdot 10^{-3}$.0107	%
	P	7.5943	7.5956		.0176	%
700	C	12.4700	12.4722	$-2.26 \cdot 10^{-3}$.0181	%
	P	7.5929	7.5952		.0297	%
800	C	12.4717	12.4715	$2.44 \cdot 10^{-4}$.0020	%
	P	7.5947	7.5944		.0032	%
900	C	12.4717	12.4706	$1.08 \cdot 10^{-3}$.0086	%
	P	7.5946	7.5935		.0142	%
1000	C	12.4705	12.4696	$9.00 \cdot 10^{-4}$.0072	%
	P	7.5935	7.5926		.0119	%
1500	C	12.4701	12.4670	$3.09 \cdot 10^{-3}$.0247	%
	P	7.5931	7.5900		.0406	%
2000	C	12.4705	12.4697	$8.24 \cdot 10^{-4}$.0066	%
	P	7.5935	7.5926		.0109	%

8.2.6 Valuing European Calls with p_f with Different Strikes

To assess the accuracy of the trinomial tree with p_f , European call options are valued with the tree with the local volatility surface from 8.2.2 as an input. Since this surface is derived from an

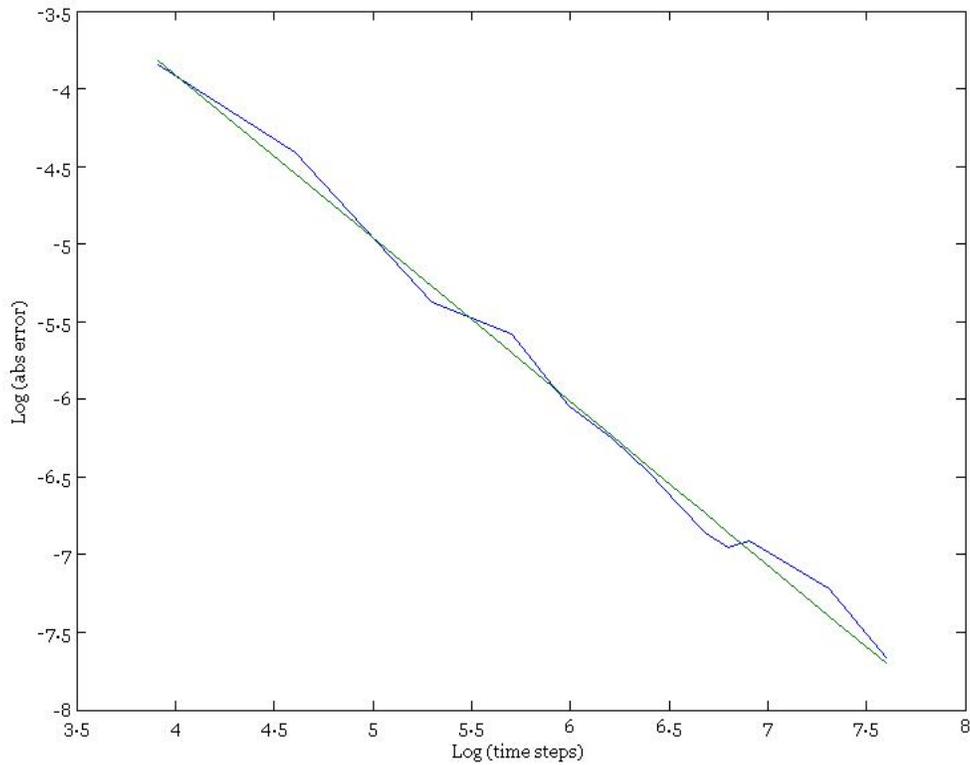
implied volatility surface, the Black-Scholes price is known. Valuing the options over a range of strikes (from $\frac{1}{2}S_0$ to $2S_0$ with intervals of 5) and deducting the Black-Scholes price, shows how the tree performs for different European options. The time steps considered are the same ones as given in the table in 8.2.5.



The minimum and maximum differences between the tree and the Black-Scholes formula, over the different strikes per given time step are given in the table below. The difference goes from a few cents for 50 time steps to a few hundredths of a cent for 2000 time steps

Time Steps	Min Diff	Max Diff
50	$-2.14 \cdot 10^{-2}$	$1.21 \cdot 10^{-2}$
100	$-1.22 \cdot 10^{-2}$	$1.02 \cdot 10^{-2}$
200	$-4.63 \cdot 10^{-3}$	$4.13 \cdot 10^{-3}$
300	$-3.76 \cdot 10^{-3}$	$3.04 \cdot 10^{-3}$
400	$-2.38 \cdot 10^{-3}$	$2.21 \cdot 10^{-3}$
500	$-1.88 \cdot 10^{-3}$	$1.92 \cdot 10^{-3}$
600	$-1.56 \cdot 10^{-3}$	$1.49 \cdot 10^{-3}$
700	$-1.25 \cdot 10^{-3}$	$1.15 \cdot 10^{-3}$
800	$-1.05 \cdot 10^{-3}$	$1.02 \cdot 10^{-3}$
900	$-7.85 \cdot 10^{-4}$	$9.51 \cdot 10^{-4}$
1000	$-9.98 \cdot 10^{-4}$	$8.73 \cdot 10^{-4}$
1500	$-7.34 \cdot 10^{-4}$	$6.24 \cdot 10^{-4}$
2000	$-4.70 \cdot 10^{-4}$	$3.12 \cdot 10^{-4}$

To gain insight into how the values on the trinomial tree converges to Black-Scholes prices, the logarithm of the maximum absolute difference in the table above is plotted against the logarithm of the time steps in the tree.



The dependency seems to be linear. This suggests a dependence of the form $y = bx^a$ (with y denoting the absolute error and x the number of time steps in the tree). The fitted green line has

a slope of -1.05 . This means that for European options the tree converges to Black-Scholes prices approximately like $\frac{1}{x}$.

8.2.7 Monte Carlo Simulations for European Options

European at-the-money ($S_0 = K = 100$) options are simulated for different sample paths, with 100 time steps and a flat volatility surface of $\sigma = 0.4$. The results with their respective 95% confidence intervals are given below. The computational time is given in seconds and denotes the time for the call and put calculations combined.

Black-Scholes: Call = 18.0230, Put = 13.1459

Samples	Call/Put	MC	95% CI	MC - BS	Comp. Time ¹
10^3	C	17.3246	1.9710	-.6984	< 1
	P	12.7012	1.0273	-.4447	
$5 \cdot 10^3$	C	18.0382	.8546	.0153	< 1
	P	13.2132	.4657	.0673	
$5 \cdot 10^4$	C	18.0944	.2745	.0714	< 1
	P	13.1868	.1479	.0409	
$5 \cdot 10^5$	C	17.9755	.0862	-.0474	8
	P	13.1310	.0466	-.0149	
$5 \cdot 10^6$	C	18.0338	.0274	.0108	90
	P	13.1530	.0148	.0071	
$5 \cdot 10^7$	C	18.0214	.0086	-.0016	919
	P	13.1505	.0047	.0046	

The same simulations are performed with the local volatility surface of 8.2.2 as an input.

Black-Scholes: Call = 12.4707, Put = 7.5936

Samples	Call/Put	MC	95% CI	MC - BS	Comp. Time ¹
10^3	C	12.3582	1.2981	-.1125	< 1
	P	7.4788	.6807	-.1148	
$5 \cdot 10^3$	C	12.2987	.5761	-.1720	< 1
	P	7.5114	.3017	-.0823	
$5 \cdot 10^4$	C	12.5358	.1855	.0651	3
	P	7.6333	.0961	.0397	
$5 \cdot 10^5$	C	12.4435	.0581	-.0272	29
	P	7.5918	.0303	-.0019	
$5 \cdot 10^6$	C	12.4673	.0184	-.0034	305
	P	7.6007	.0096	.0071	
$5 \cdot 10^7$	C	12.4667	.0058	-.0040	2993
	P	7.6015	.0030	.0079	

In all but one case the Black-Scholes approximation lies within the 95% confidence interval. In the only case it doesn't ($5 \cdot 10^7$ samples for the non-flat volatility) the estimate is already quite accurate with an error less than one cent.

¹Computational time (s) in MATLAB[®] version R2008b on an Intel 3.06 GHz CPU computer with 2 GB of RAM

8.2.8 Monte Carlo Simulations for American Options

The options priced are at-the-money American options with $T = 1$, $S_0 = K = 100$ with flat volatility surface $\sigma = 0.4$, with 10^5 sample paths used for each regression, repeated 10 times (giving one million samples), 100 time steps. The regressions are done upon either normal polynomials

$$p_n(x) = \sum_{i=0}^n a_i x^i \quad (8.9)$$

so the basis functions are x^i or the basis functions are weighted Laguerre polynomials given by

$$L_n(x) = \frac{e^{x/2}}{n!} \frac{d^n}{dx^n} e^{-x} x^n \quad (8.10)$$

Tree (500 steps): Call = 18.0243, Put = 13.6689

Basis Functions	Call/Put	MC	95% CI	MC - Tree	Comp. Time ²
5 normal	C	17.8841	.0444	-.1402	222
	P	13.6776	.0138	.0086	
5 weighted Lag	C	17.9593	.0464	-.0650	311
	P	13.6838	.0137	.0149	
7 normal	C	18.0097	.0477	-.0146	363
	P	13.6901	.0138	.0211	
7 weighted Lag	C	18.0435	.0484	.0192	674
	P	13.6925	.0138	.0086	
9 normal	C	18.0491	.0473	.0248	514
	P	13.6952	.0137	.0262	
11 normal	C	18.0716	.0481	.0473	670
	P	13.6985	.0138	.0296	

And the same options priced with the volatility surface as in 8.2.2

Tree (500 steps): Call = 12.4726, Put = 8.1206

Basis Functions	Call/Put	MC	95% CI	MC - Tree	Comp. Time ²
5 normal	C	12.3750	.0300	-.0976	322
	P	8.1178	.0096	-.0039	
5 weighted Lag	C	12.4356	.0313	-.0370	401
	P	8.1272	.0096	.0065	
7 normal	C	12.4745	.0318	.0019	475
	P	8.1285	.0096	.0079	
7 weighted Lag	C	12.4921	.0319	.0195	791
	P	8.1291	.0096	.0084	
9 normal	C	12.4839	.0318	.0113	623
	P	8.1397	.0096	.0190	
11 normal	C	12.4764	.0323	.0038	768
	P	8.1178	.0096	-.0029	

²Computational time (s) in MATLAB[®] version R2008b on an Intel 3.06 GHz CPU computer with 2 GB of RAM

The results presented in the table above, indicate that the tree model and Monte-Carlo simulations give comparable values for American options. This suggests the tree model works accurately for both European and American options.

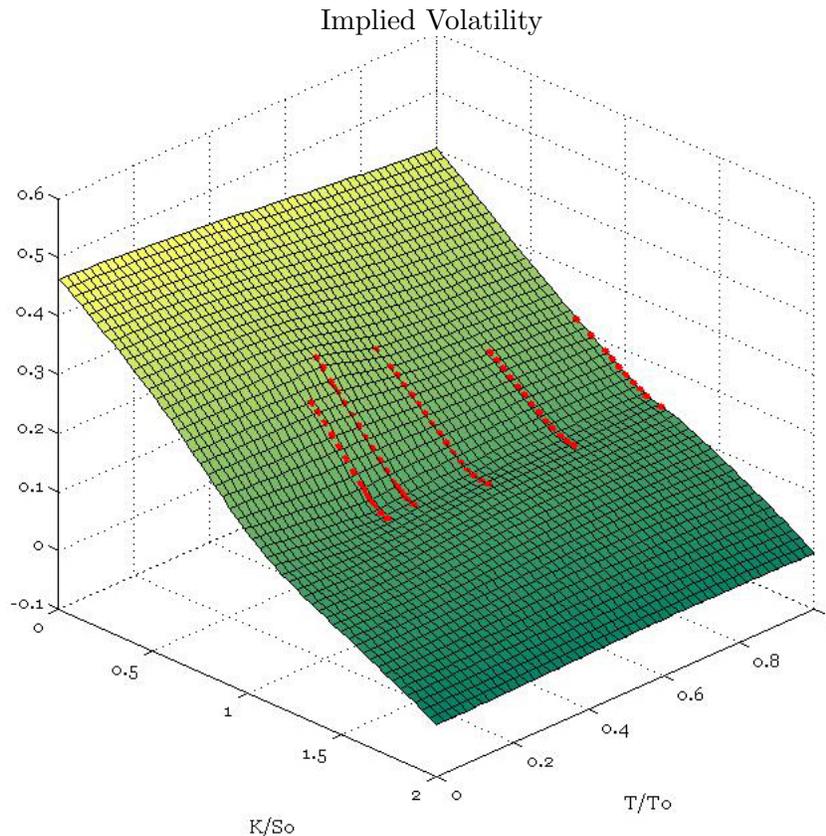
8.3 AEX

In section 2.4 it was said that fitting an implied volatility surface to market data was for a large part more art than science. This can be well illustrated by the case of fitting options on the AEX index. The market data was observed on May 21, 2007, at a given point in time during the day when the index was at 536.92. All options are European.

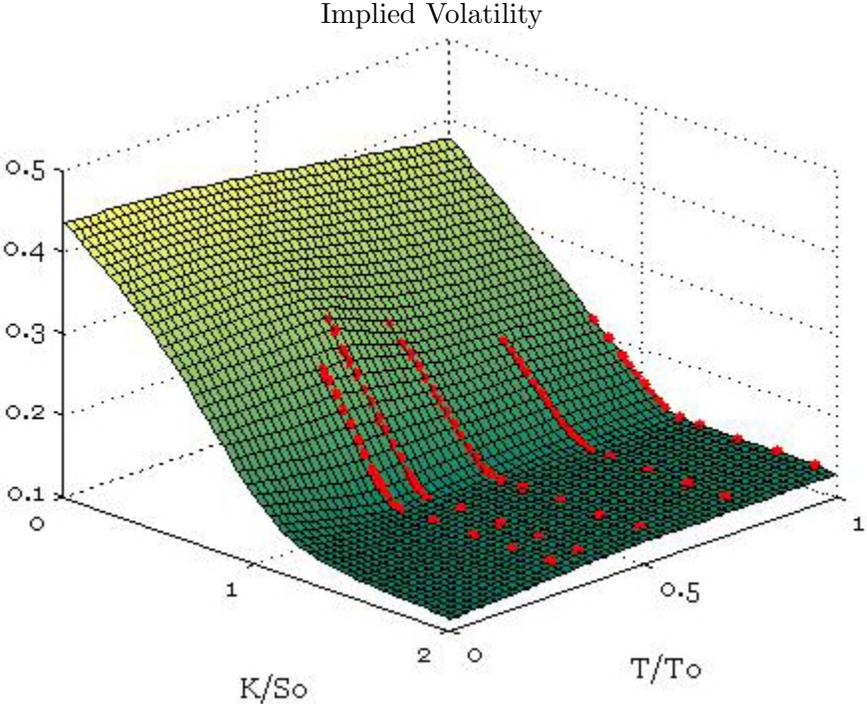
The spot rate of interest $R(t_0, T)$ for different times of maturity T (given in days until the maturity date) are given by

T (days)	26	61	89	124	152	215	305	397
$R(t_0, T)$ (%)	3.964	4.023	4.077	4.141	4.173	4.244	4.318	4.265

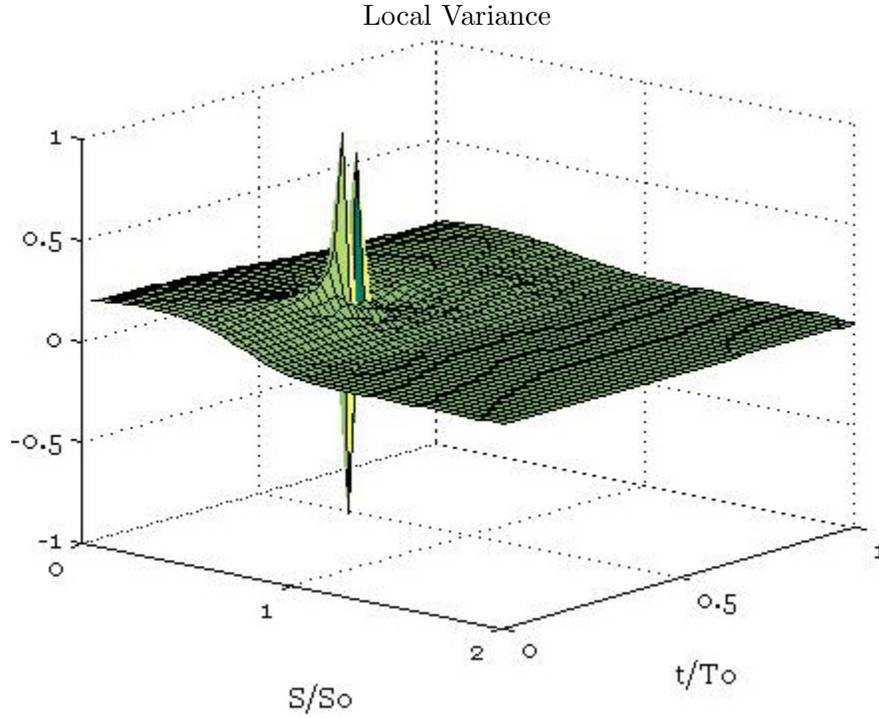
Options are available for the last five maturities in the table above. An initial fitting of the TPS surface (as described in 2.4) gives the following result



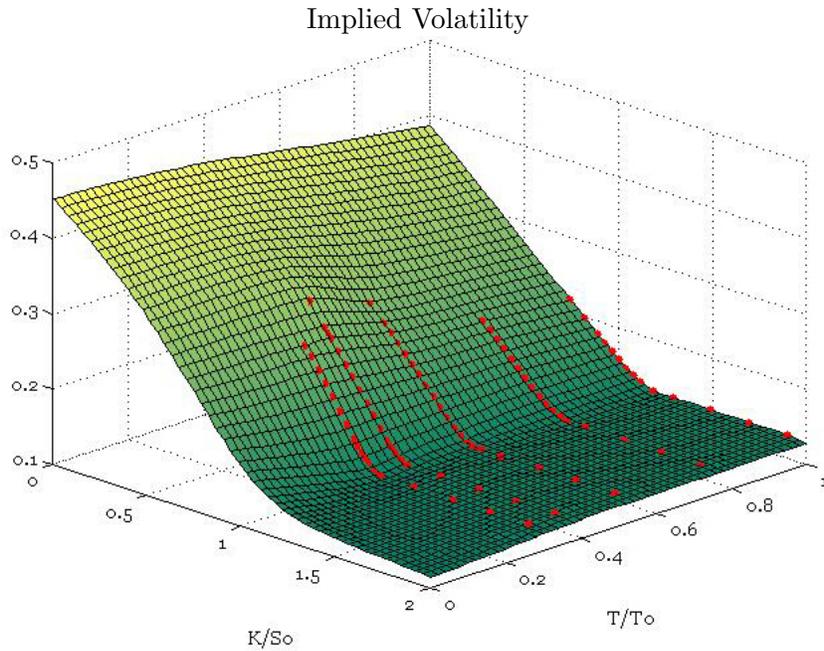
Here the strike and maturity axes are rescaled by the spot price of the index S_0 and the largest maturity T_0 . The red dots are the mid market prices. From the figure above, it is clear that this TPS surface will not suffice. The absence of data for some regions means that in these regions the TPS makes an extrapolation. While this is not a problem for low strike values, this leads to negative values for large strikes. This is of course not realistic, and thus not a valid extrapolation. Some assumption thus has to be made for the shape of the surface for large strike values. Here the assumption is made that for strikes larger than the largest strike of the available data for that maturity, the implied volatility remains constant. This leads to



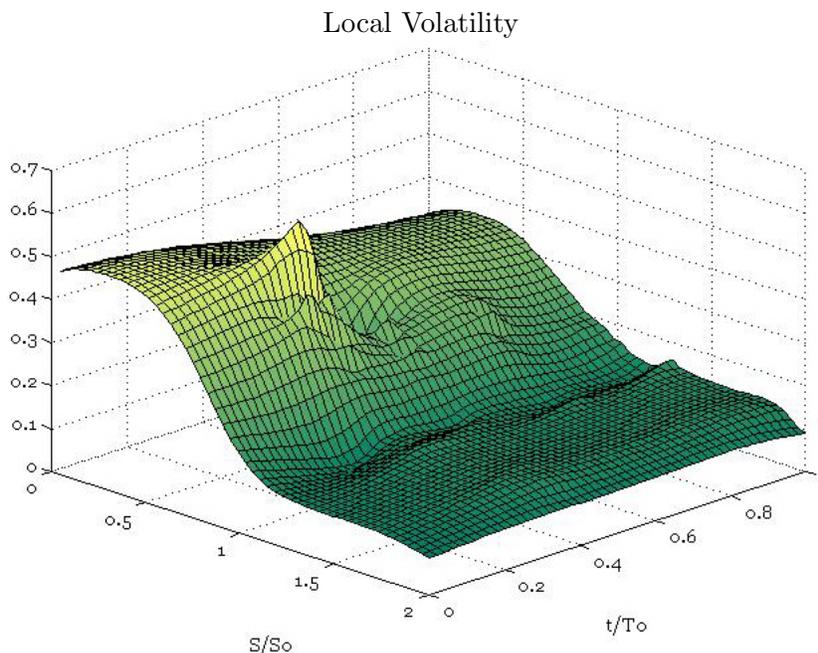
This is clearly more acceptable. The resulting local variance surface is



Negative local variances make no sense and would lead to a local volatility surface with complex values. But this can be resolved by making a small adjustment to the TPS surface. When fitting the implied volatility surface, it is not necessary that the surface goes exactly through all the mid market prices. The only criterion should be that it falls within the bid ask spread. By removing one data point for fitting purposes (the second lowest strike for the second maturity), the fitted surface becomes



This is a small adjustment from the previous fitted surface. At the site of the point that was removed the surface is still between the bid ask spread, so this surface is equally justifiable. The resulting local volatility surface is



So small adjustments in the implied volatility can lead to large changes in the local volatility. Care should thus be taken when the surface is fitted to the market data. A possible criterion for the inclusion of data points could be the size of the bid ask spread. Removing a point with a large bid ask spread (as was done here), will not likely lead to problems since the resulting surface will usually still go through the bid ask spread. This should of course always be checked.

8.4 Royal Dutch/Shell

European options are relatively easy to value. To test the model thoroughly it should be used on American options. Here options on Royal Dutch/Shell Class A (RDSA) shares, as traded on NYSE Euronext Amsterdam are considered. The data is from trading on January 2, 2006. The spot price of the underlying is 26.035.

The spot rate of interest $R(t_0, T)$ for the different times of maturity T (given in days until the maturity date) are given by

T (days)	19.6	47.6	75.6	166.6	257.6	292.6	348.6	530.6
$R(t_0, T)$ (%)	2.400	2.459	2.518	2.668	2.786	2.831	2.866	2.978
T (days)	656.6	719.6	901.6	1020.6	1083.6	1447.6	1811.6	
$R(t_0, T)$ (%)	3.035	3.051	3.099	3.130	3.146	3.217	3.265	

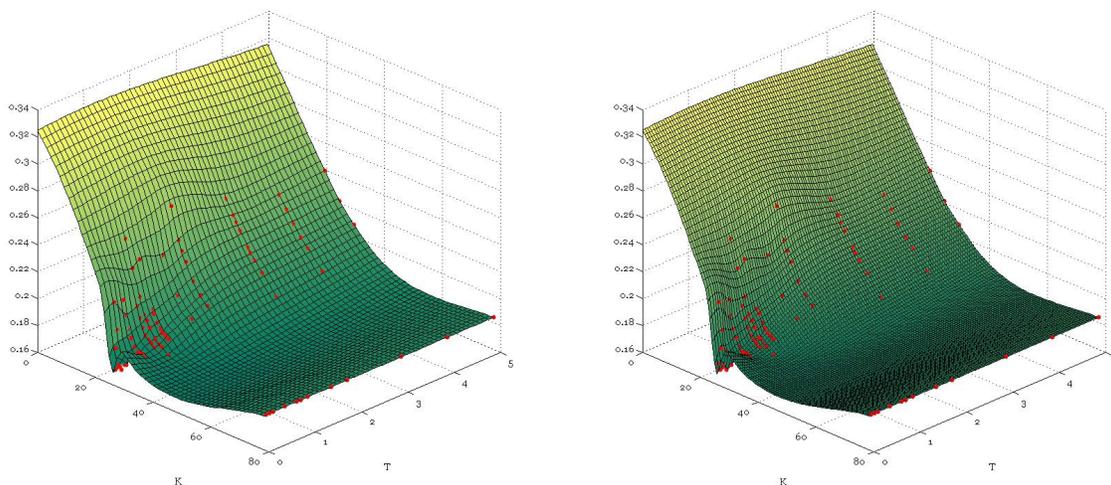
And the date, t (given in days from January 2, 2006), and size of future dividends, in euros, are

t (days)	37.8	128.8	212.8	303.8	401.8	492.8	576.8
Dividend	0.22	0.22	0.22	0.22	0.22	0.235	0.235
t (days)	667.8	765.8	856.8	940.8	1038.8	1129.8	1220.8
Dividend	0.235	0.235	0.235	0.235	0.235	0.235	0.235
t (days)	1311.8	1402.8	1493.8	1584.8	1675.8	1766.8	
Dividend	0.235	0.235	0.235	0.235	0.235	0.235	

8.4.1 The Volatility Surfaces

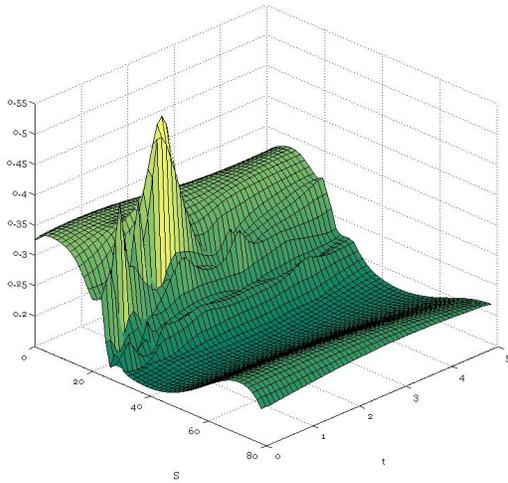
As described in 8.3, some assumption has to be made for the implied volatility at large strikes. Here the implied volatility at $3S_0$ is fixed at 0.185 for all strikes. How much the surface should be smoothed cannot be determined up front. It usually depends on the grid on which it is calculated. In theory this shouldn't make a difference, but since the derivatives are determined numerically (see 8.1), it does. Consider for instance the implied volatility surface for RDSA with $\lambda = 0$ for both a 50x51 grid (T at calculated at 50 points, K at 51) and a 100x101 grid

Implied Volatility

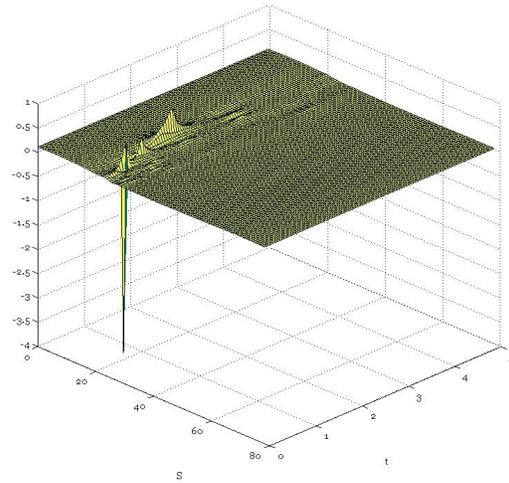


These are of course identical except for the grid on which they are defined. This results however in vastly different local volatility/variance surfaces

Local Volatility

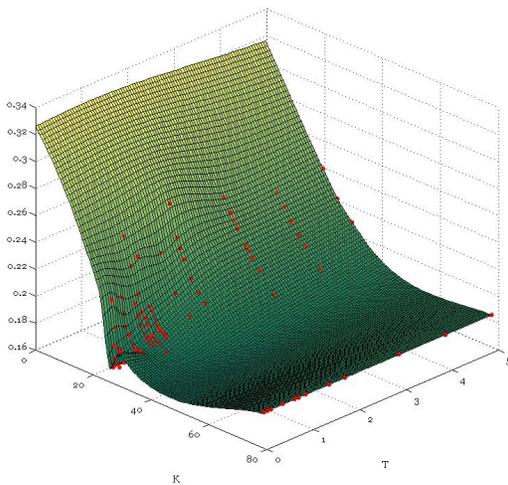


Local Variance

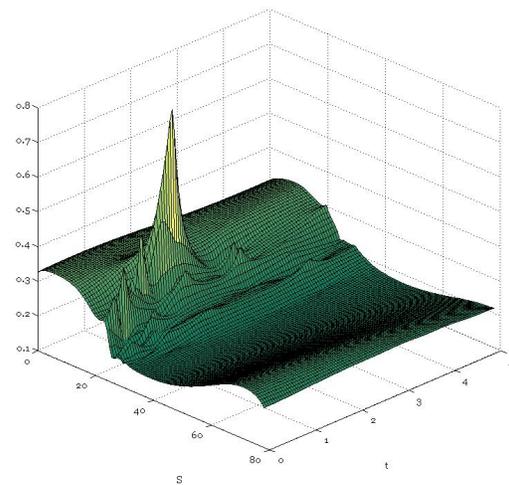


where the left one is local volatility and the right one local variance. This means the second implied volatility surface needs more smoothing. To determine which smoothing parameter λ should be used, it is first put at zero, and if the local variance is negative at some point it is increased by 0.001 until the local variance is positive everywhere. Following this procedure the required smoothing for the 100x101 grid is $\lambda = 0.001$. The resulting implied volatility and local volatility surfaces for the 100x101 grid are then

Implied Volatility



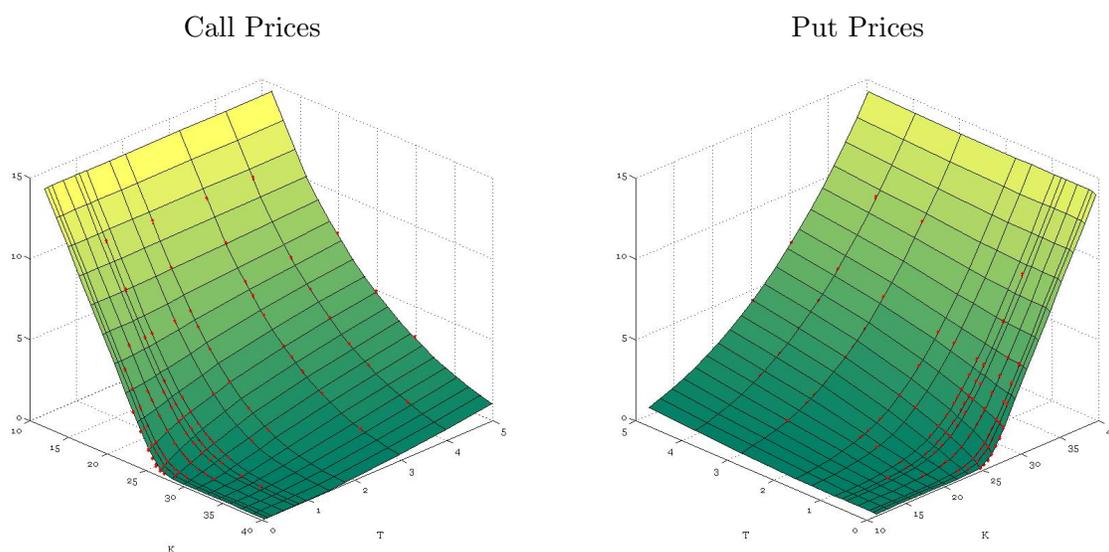
Local Volatility



Of course this gives the minimum smoothing necessary to get a viable local volatility surface. In some cases it may be advantageous to use a larger smoothing parameter to avoid large spikes in the local volatility surface, which usually lead to negative transition probabilities in the trinomial tree.

8.4.2 Verifying Results

The purpose of the local volatility surface is to price options on RDSA. These resulting prices can be compared to the market data, to assess how the model performs. This may seem a bit redundant, since the local volatility surface was in the end extracted from the market data. This would mean it gives back the input, not an impressive result. In this case, however, the input are implied volatilities, which are obtained from the market prices by inverting a pricing process for American options, used by AllOptions. So comparing the prices given by the model with market prices is in effect comparing the local volatility pricing process to the one used to obtain the implied volatilities. In the figures below different American calls (left) and puts (right) are valued with a trinomial tree with 400 time steps, and the market prices given by the red dots

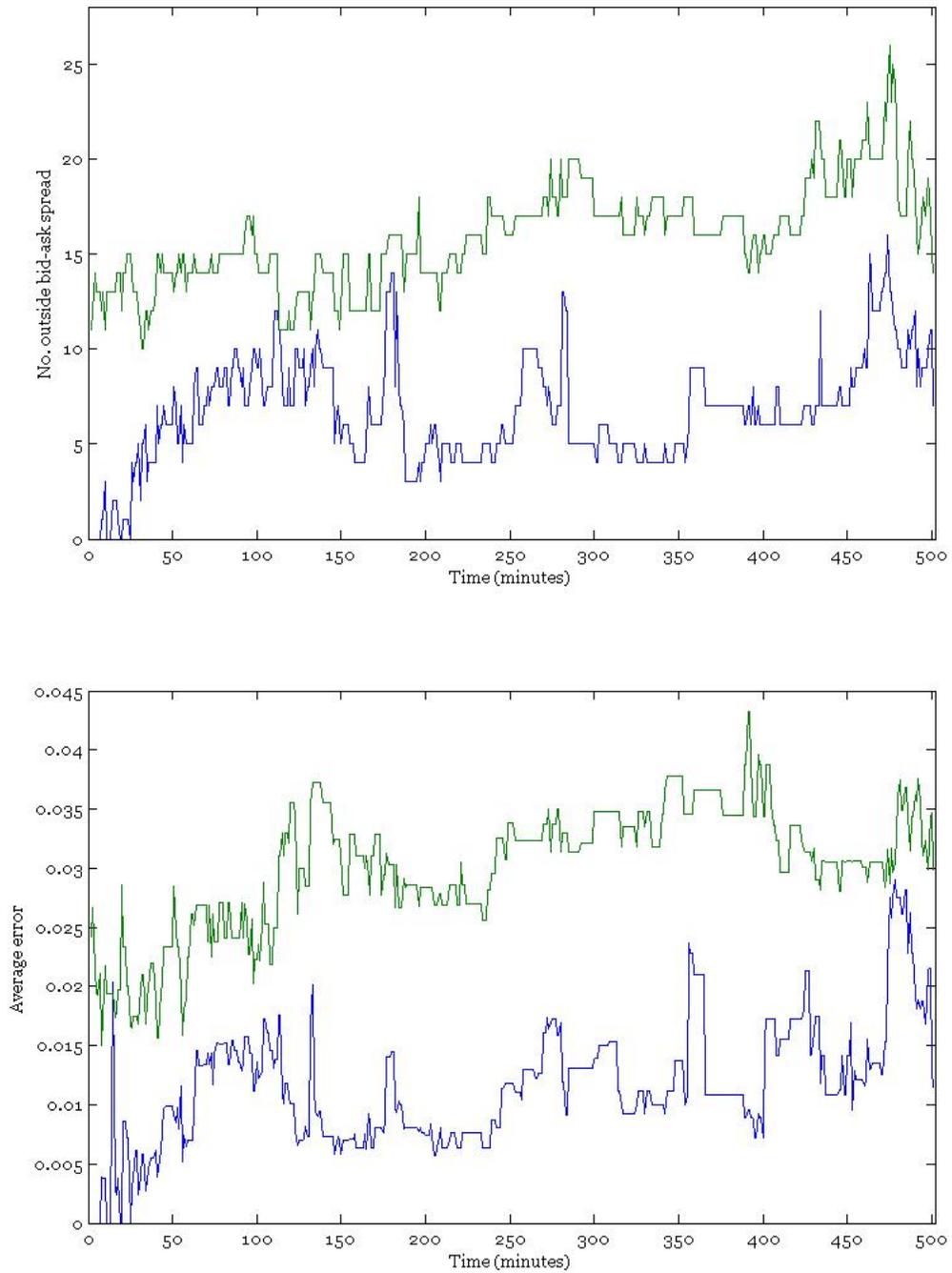


Although this is a nice illustration, it would be better to value the exact same options for which there are market prices, and compare these. How well these match is given by two numbers, the number of modelled prices that fall outside the bid ask spread, and the average absolute difference or absolute ‘error’ for these values between the modelled price and the nearest bid or ask. For the data presented above, the 99 values for the call options all fall between bid and ask. For the put options 11 out of 95 options fall outside the bid ask spread, with an average absolute error of 0.0242. It seems the models give the same results for call prices, but different values for put prices.

8.4.3 Stability of Local Volatility Surface

If the local volatility model is used in practice, it would not be practical to have to recalculate the local volatility surface every minute. It would be easier to obtain the local volatility surface at some point in time and then keep using this same surface to value options at other times. To see what this means for RDSA options the local volatility surface is calculated from implied volatility data at a fixed point. Then at every following minute, the spot price is observed, and the values for the same options are calculated with a trinomial tree with 400 time steps and compared to the market data. This is done for 500 points (minutes), and the number of modelled prices that fall

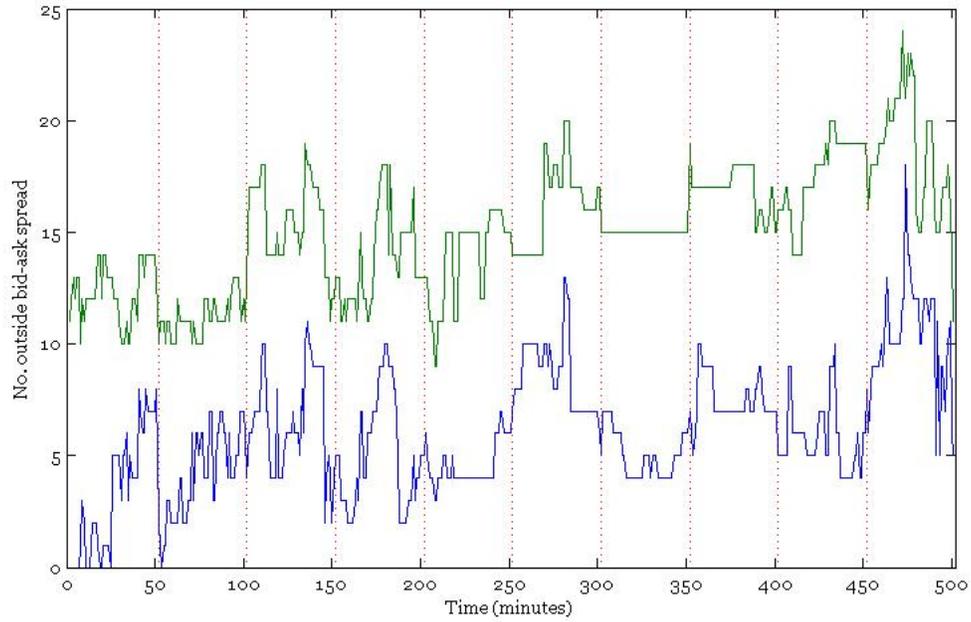
outside the bid ask spread, and the average absolute error for these options is recorded. The result is as follows

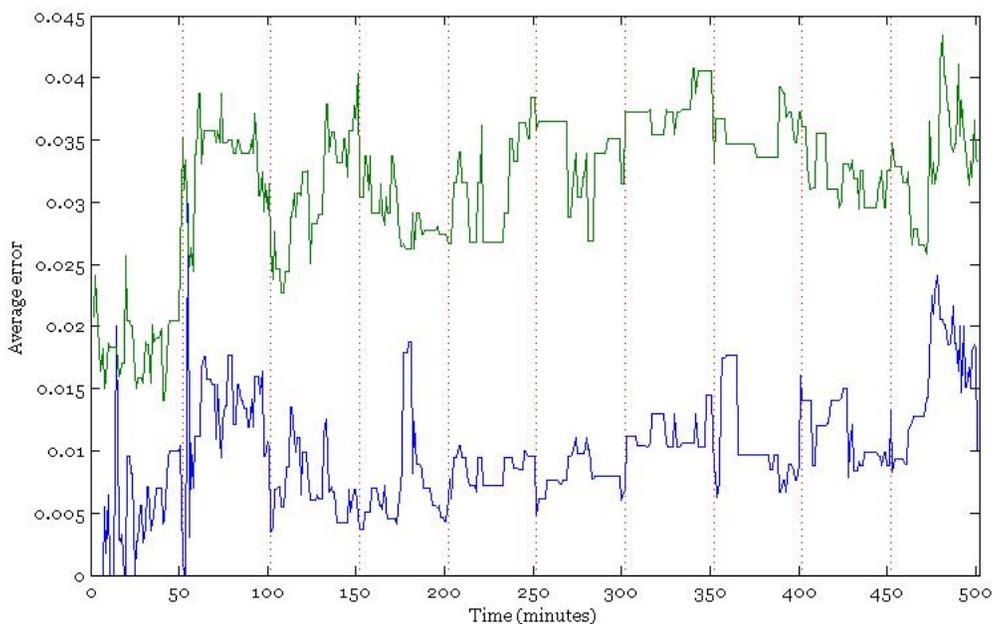


In these graphs the blue line denotes the values for call options (99 in total) and the green line denotes put options (95 in total). These lines are fairly flat, although a slight upwards trend can be detected in all. It would suggest that the local volatility surface obtained at one point can be

used equally well at other points up to 500 minutes, since the number of points outside the bid ask spread and their average absolute error is about the same at minute 1 as it is at minute 500. The slight upward trend of the lines, does suggest that this cannot be said for any point in time in the future. The models seem to agree quite well for call options, while for put options, significant differences are observed.

An obvious improvement would be to recalibrate the local volatility surface at constant intervals. When the local volatility surface is recalculated every 50 minutes the result is





In these graphs the dotted red line indicates the time when the volatility surface is recalibrated to the current market data. It seems like the recalibration does not make that much of a difference. It would thus seem that the local volatility surface implied by the data, does not change significantly over the day or that possible changes in the local volatility surface do not lead to significantly different prices. This suggests that recalibration is not worth the extra effort. Just as easily the same volatility surface can be used during the entire day.

8.4.4 Delta Hedging

The local delta of options can be calculated according to the formulae given in 6.3. If the delta is correct it should predict the change in the option price (δV) if the underlying changes by a certain amount (δS). At two points in time the value of an option and its underlying is observed and the delta is calculated. A measure of the accuracy of the delta is then given by

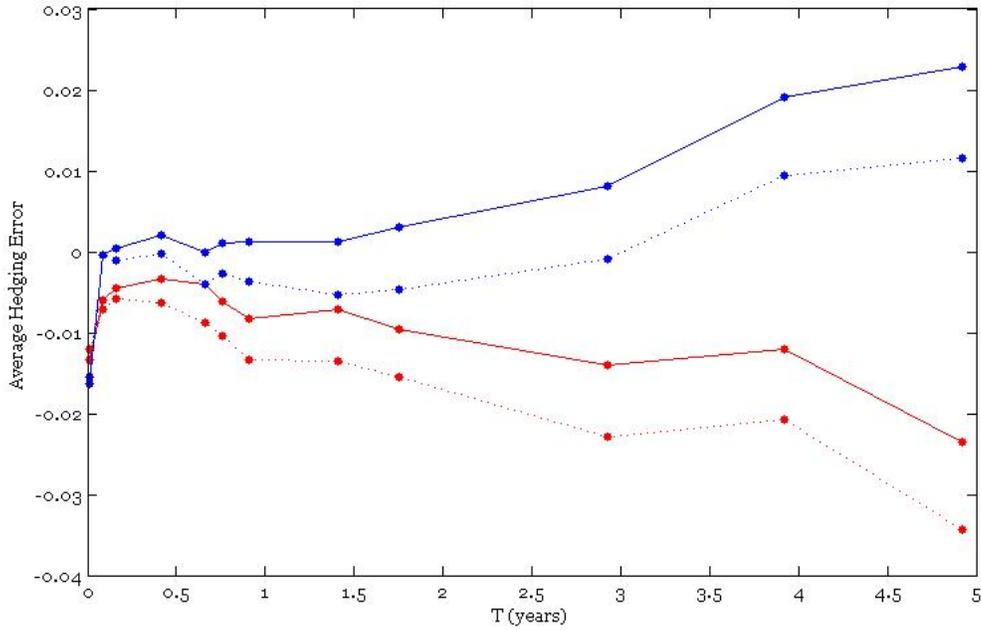
$$\epsilon = \delta V - \Delta \delta S \quad (8.11)$$

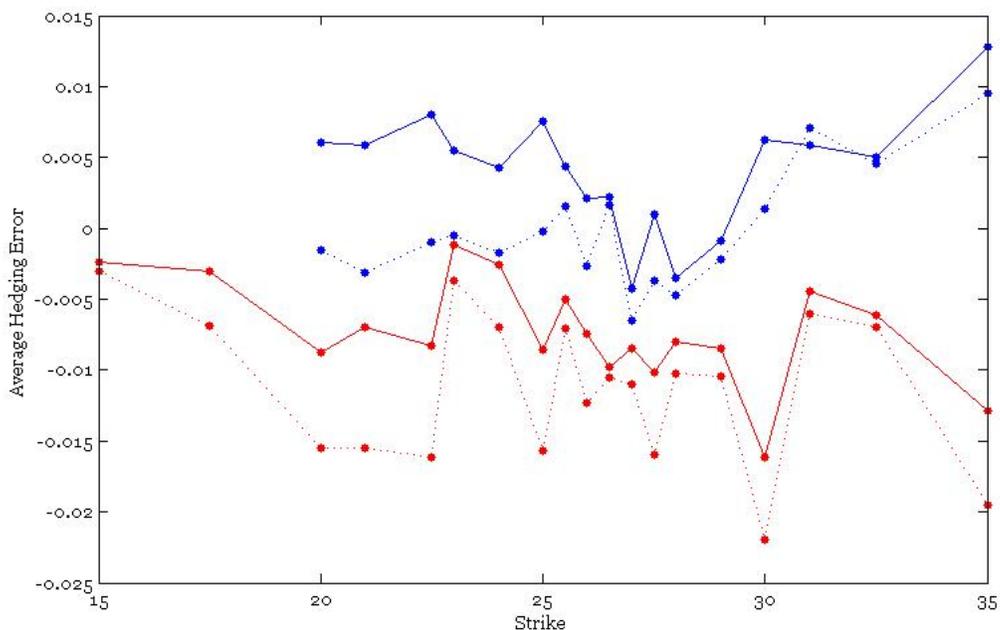
This is the error that would occur if the option was delta hedged with this particular delta. A lower absolute value for ϵ indicates a better hedge.

The performance of the local delta is observed over a period of seven days: January 2, 3, 4, 5, 9, 10 and 18. Only options for which market data is available on all these days are considered. Each option is considered to be delta hedged at every day. For each option, the average of the error, $\bar{\epsilon}$, and the standard deviation, σ_{ϵ} , over these days indicate how well the hedging scheme works. The performance of the local delta is compared to the Black-Scholes delta. Since the options on RDSA are American and there are discrete dividends, no explicit Black-Scholes delta formula exists. But it can be calculated by using a flat local volatility surface, with values equal to the implied volatility

of the option, and then calculating the delta in the same way the local delta was calculated. Since this is not the classical Black-Scholes delta, a more accurate name would be implied delta.

The following graphs show how big the hedging errors are with respect to maturity and strike. The plotted values are the average values of $\bar{\epsilon}$. In the first graph the values at a particular maturity are averaged over all strikes and the resulting values are plotted against time to maturity. In the second graph they are averaged over different maturities and plotted against the strike price.





In these graphs the normal lines denote the local delta hedge, the dashed lines denote the implied delta hedges. Red lines represent call options, blue ones denote put options. It should be noted that during the time period over which the data was observed the stock price did not vary much, it stayed between 26 and 27 euros. This means the same strikes represented about the same moneyness over the different days. The complete set of data can be found in the Appendix.

From these graphs the following observations can be made about the hedging errors:

- The hedging errors for call options are usually negative, for put options they are usually positive.
- The largest errors occur for the options with the largest maturities.
- For short maturities the local delta and implied delta give rise to approximately the same errors.
- For larger maturities the local delta gives smaller errors than the implied delta for call options, but larger errors for put options.
- A general statement on the error as a function of the strike, or moneyness, cannot be made. The results shown do not show a clear dependency on the strike price.

Given these observations a clear cut general statement on the performance of the local delta over the implied delta cannot be made.

9 Conclusion

The aim of this report was to investigate the local volatility model. It was found in section 2 that although the Dupire formula gives a nice theoretical method to extract the local volatility surface from option prices, in practice it is better to extract the surface from implied volatilities. To achieve this, the Dupire formula was transformed into a formula that relates local volatility to implied volatility.

To price options valuation trees were used. By adjusting the traditional binomial tree to accommodate for local volatility, it was found in section 3 that the binomial tree was unstable. Since this is unrealistic and gives unstable results, it was deemed insufficiently appropriate. The trinomial tree model provides more flexibility (section 4). Some adjustments had to be made to avoid negative transition probabilities which occurred because more flexibility was needed in some places in the tree. By transforming the tree into a quintinomial tree in those places, the negative probabilities were resolved. Discrete dividends were incorporated into the tree using the Vellekoop-Nieuwenhuis method (section 5).

The numerical results presented in section 8 show that the model works well for both European and American options. For European options the results were compared to Black-Scholes prices, for American options they were compared to Monte Carlo simulated prices (section 7). By increasing the number of steps in the tree, the difference between these values could be made less than one cent.

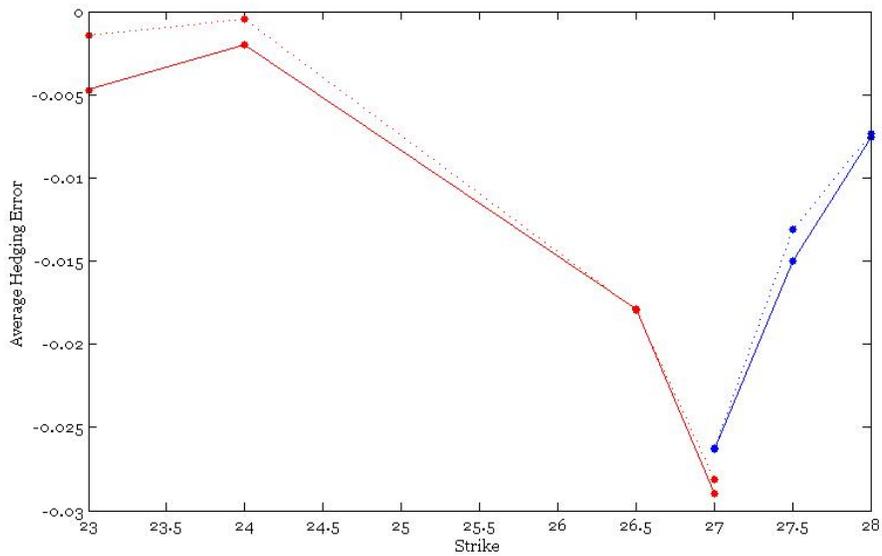
When the model was used on market data for options on Royal Dutch/Shell Class A shares (section 8.4), it was found that the prices fell within the bid ask spread most of the time, but not always. This means that there exists a significant difference between the local volatility model and the model to price options currently used by AllOptions, which was inverted to generate implied volatility data. Whether one or the other is better cannot be stated, since the latter model was not investigated in this report. When the local volatility surface, generated at a given point in time, was used to price options at later times it was found that the differences between the models was fairly constant. Perhaps more significantly it was found that recalibration of the local volatility surface at regular intervals during the day did not have a significant impact, indicating that recalibrating during the day is not needed.

The literature suggests the delta's generated by the local volatility model, local delta's, perform as better delta hedges than Black-Scholes or implied delta's (section 6). For the data observed this seems true for call options with large maturities, but the reverse seems to be the case for put options with large maturities. For short maturities the delta's give approximately the same result. As such no conclusive statement can be made about the performance of the local delta. The data used was limited (it included seven trading days) and a more extensive investigation of the local delta is needed to get a more thorough understanding of its performance.

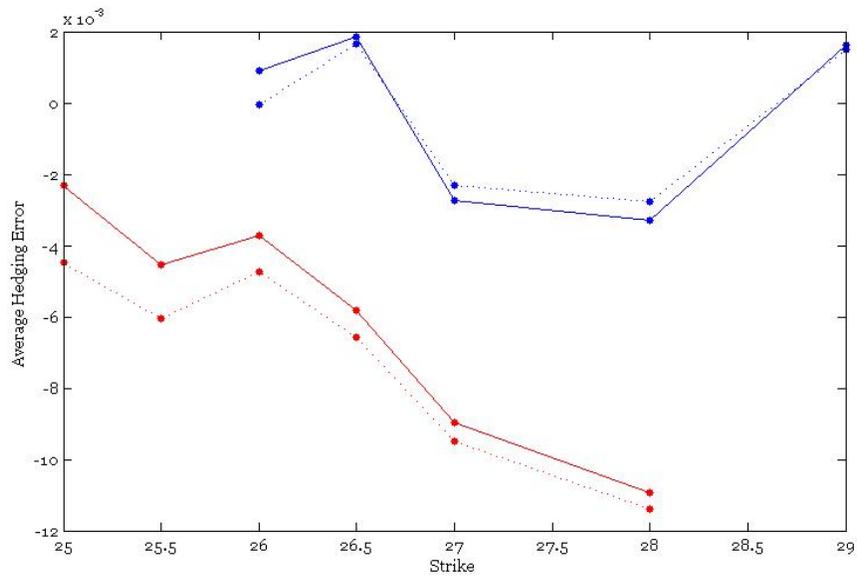
A Appendix: Delta Hedging Results

Below are the graphs of the average delta hedging error, $\bar{\epsilon}$ as a function of the different strikes, for a given maturity. The normal lines denote data for the local delta hedge, the dashed lines denote data for the implied delta hedges. Red lines represent call options, blue ones denote put options. The maturities mentioned below are the times to maturity, in years, as seen from the first day market data was observed, January 2 2006.

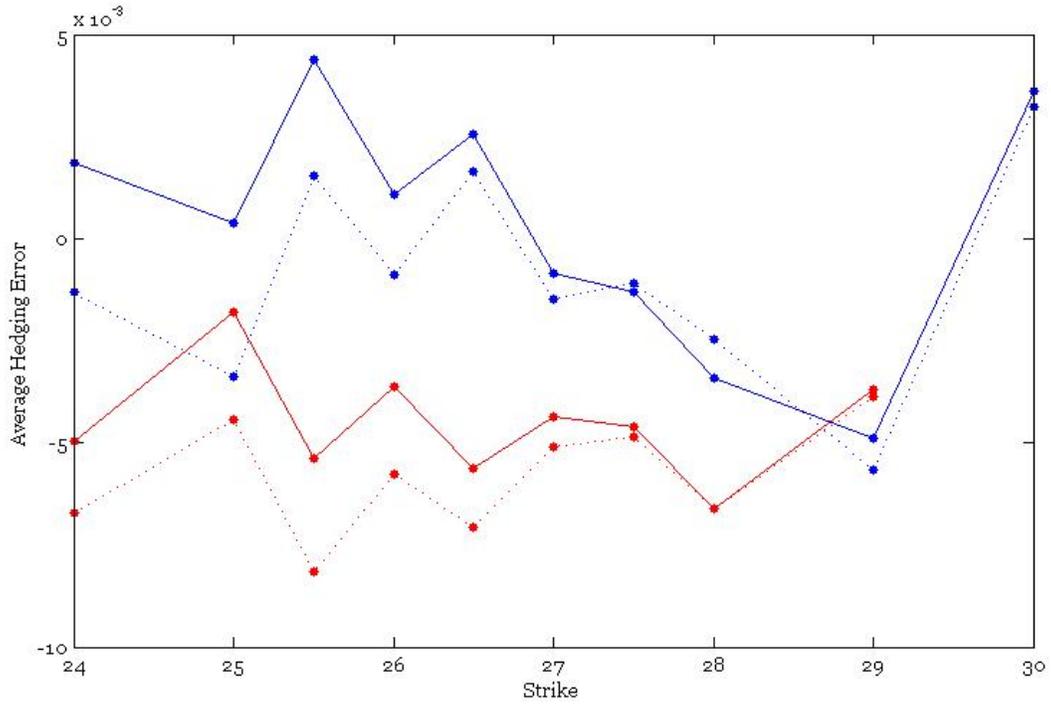
$$T = 0.0537$$



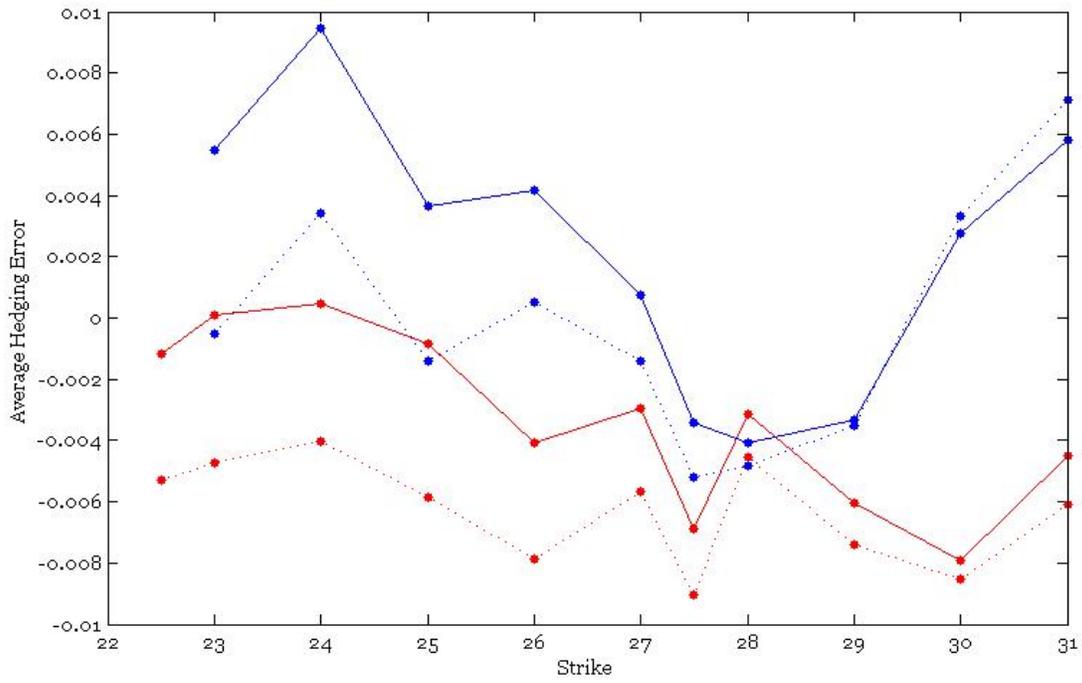
$$T = 0.1304$$



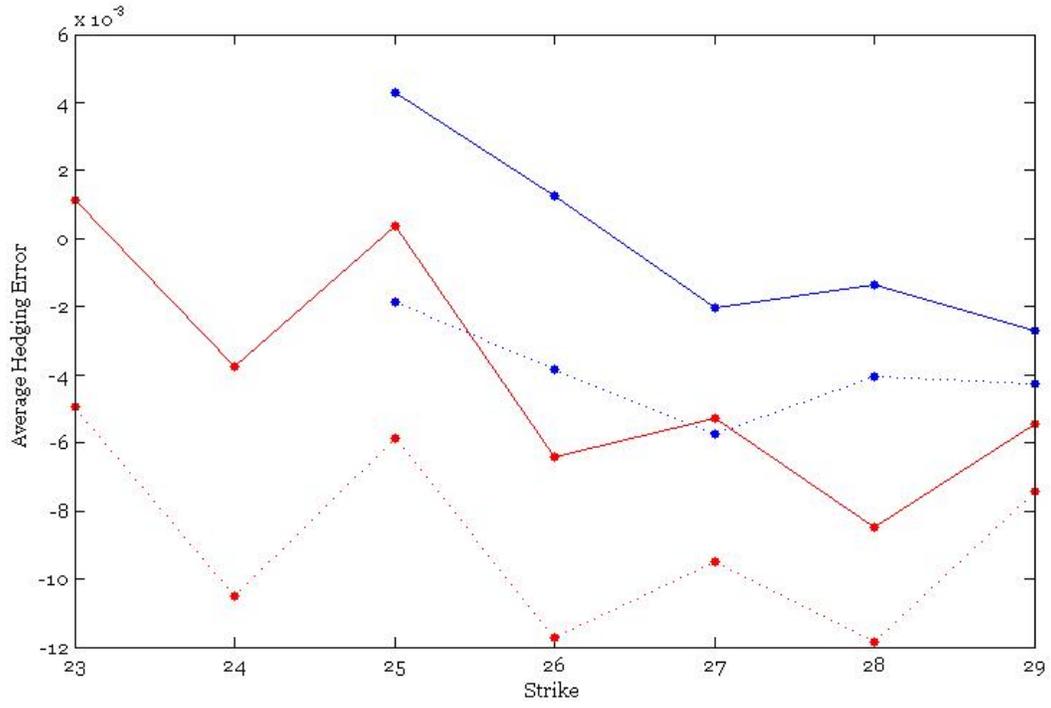
$T = 0.2071$



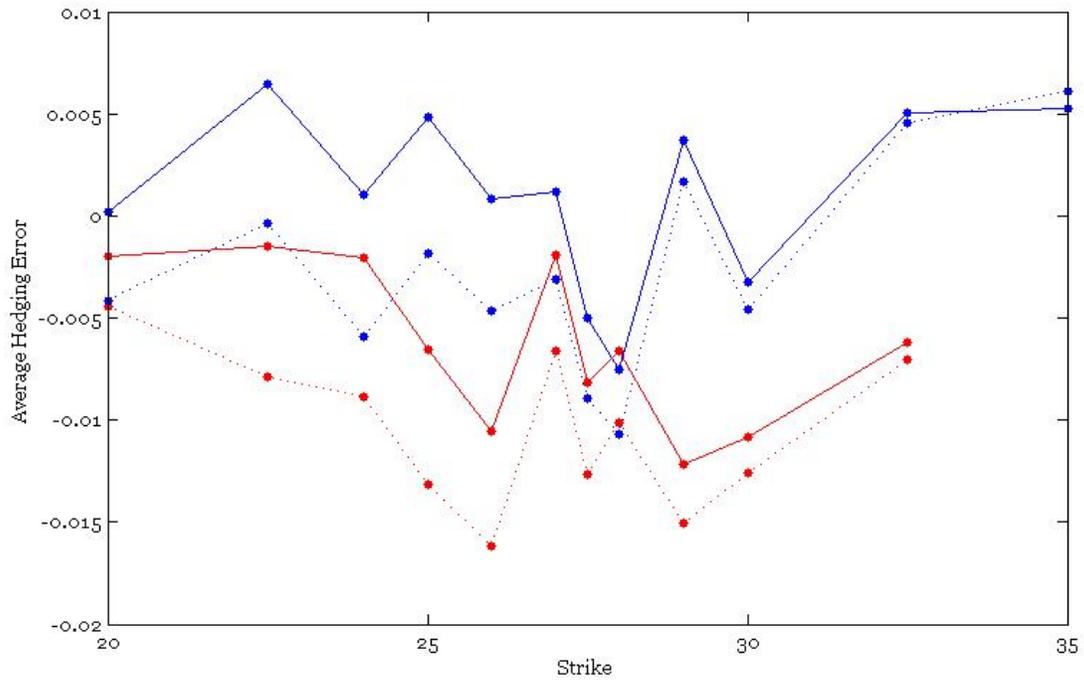
$T = 0.4564$



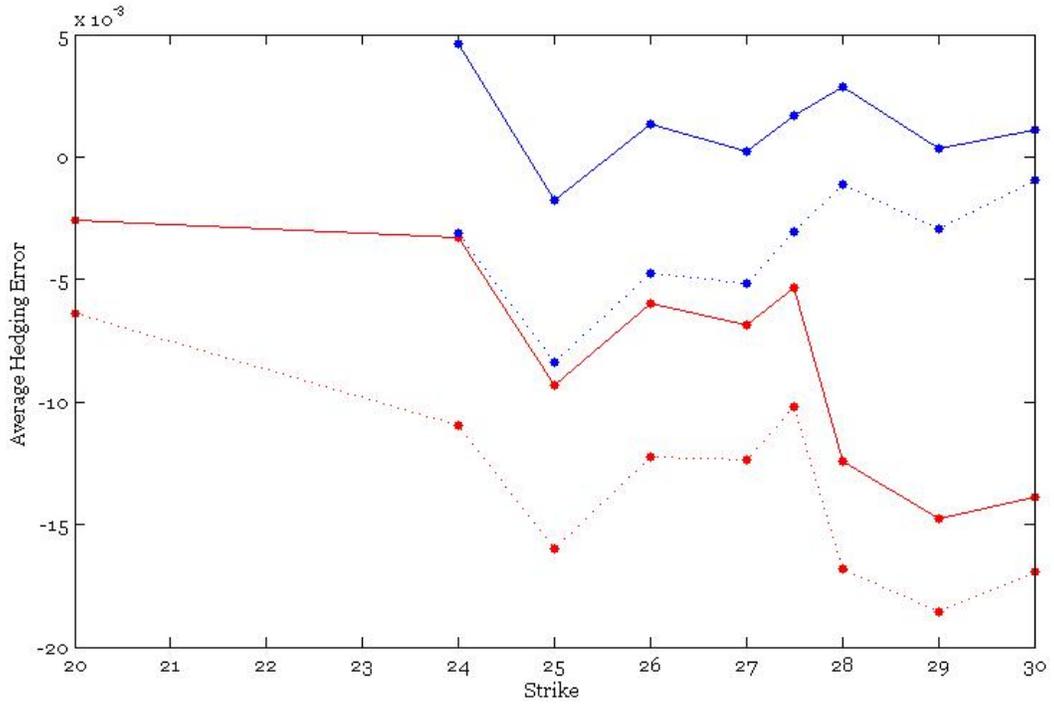
$T = 0.7058$



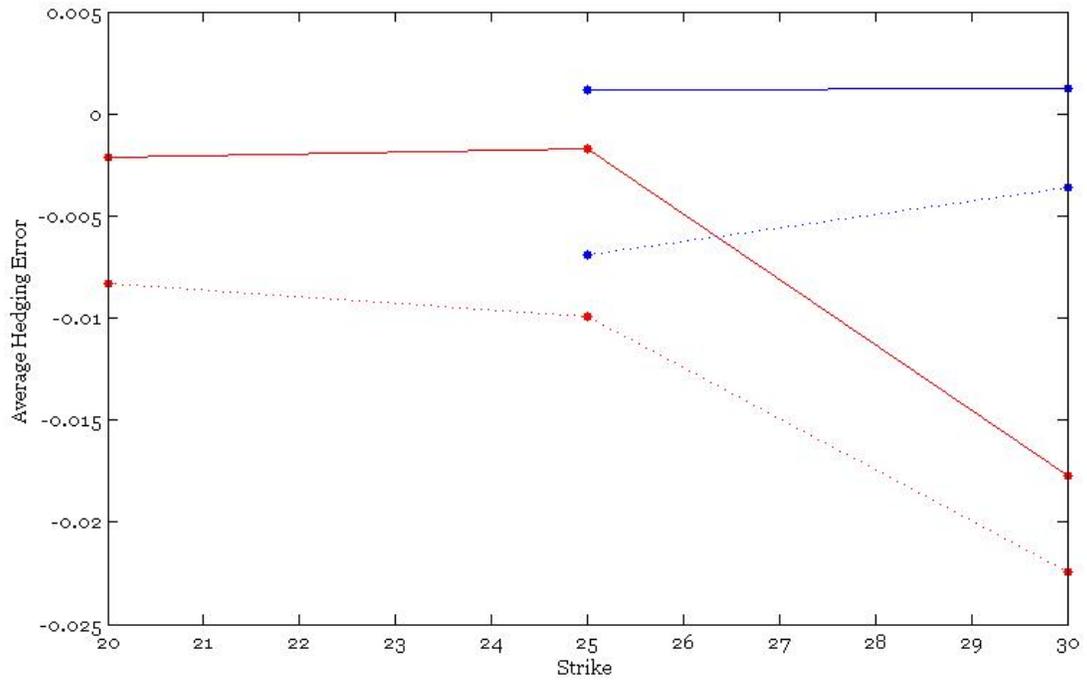
$T = 0.8016$



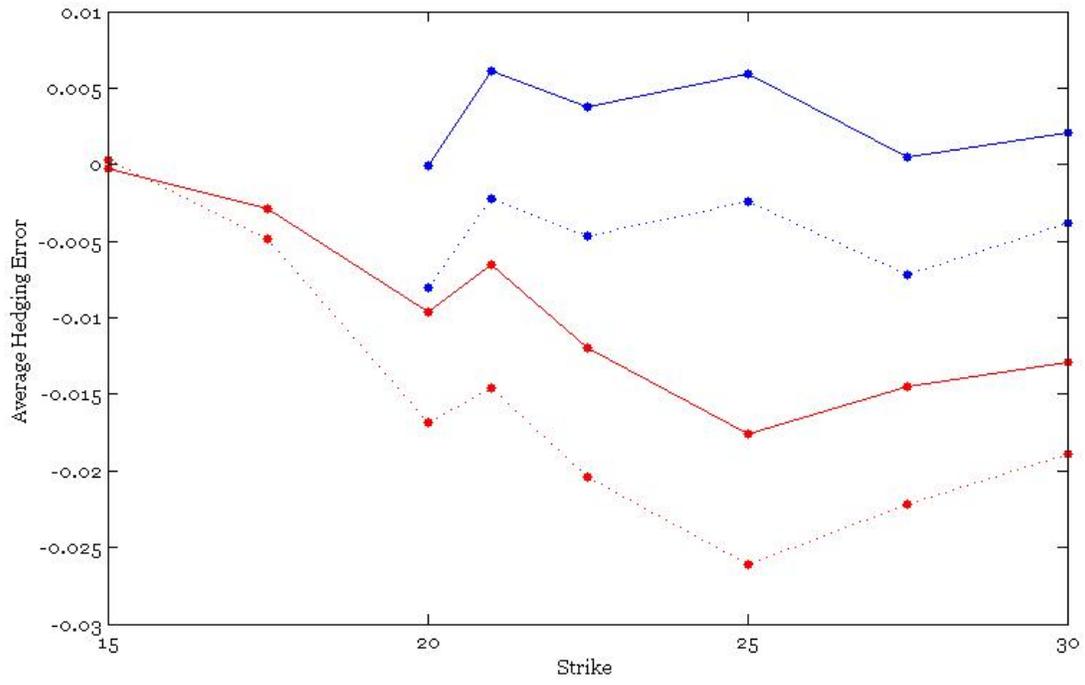
$T = 0.9551$



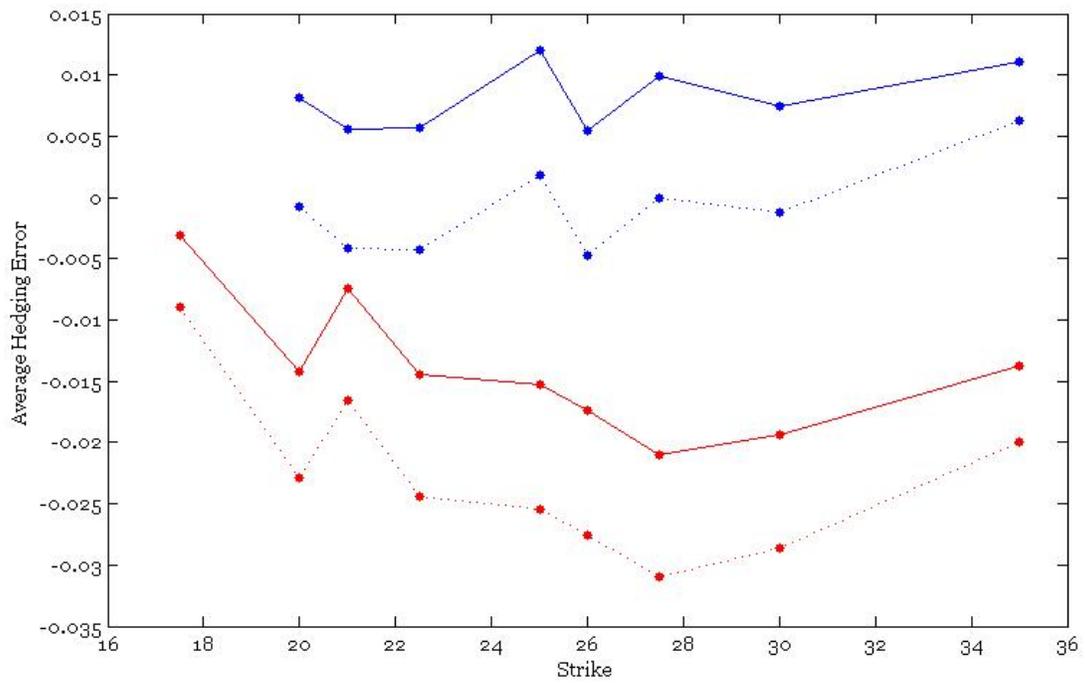
$T = 1.4537$



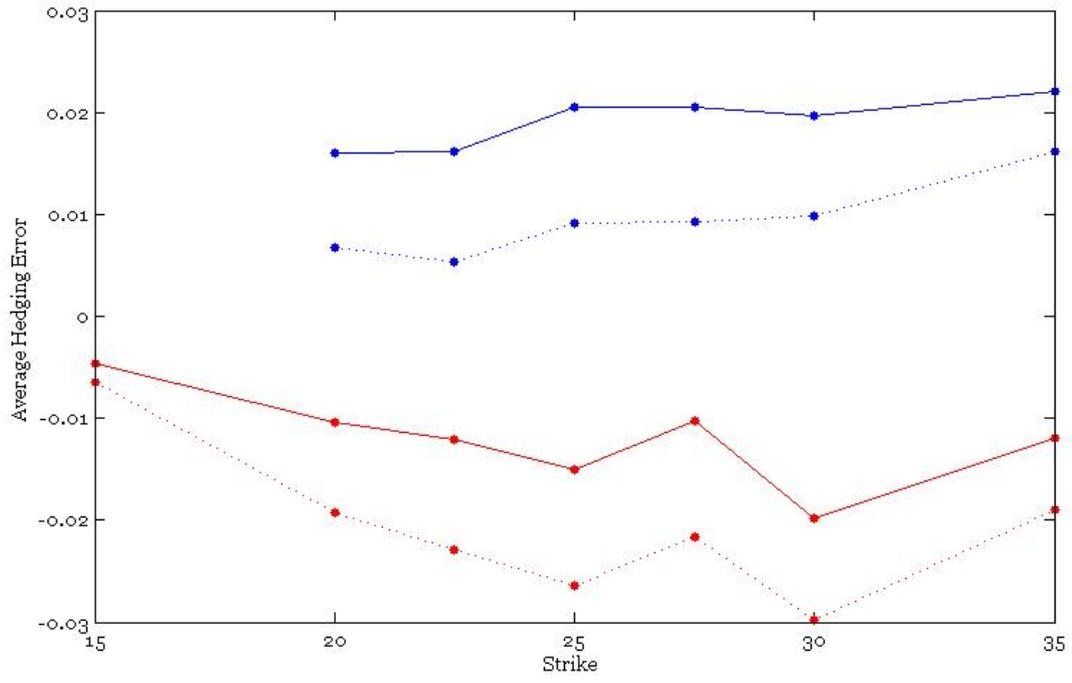
$T = 1.7898$



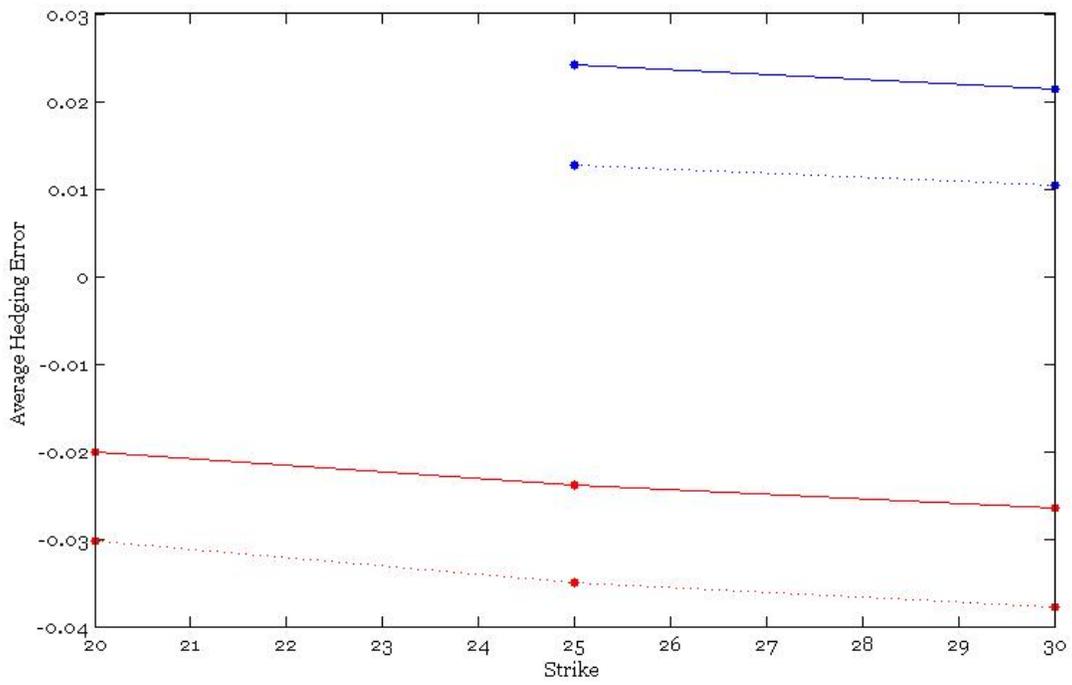
$T = 2.9688$



$T = 3.9660$



$T = 4.9633$



References

- [1] Alexander, C. and Nogueira, L., ‘Optimal Hedging and Scale Invariance: A Taxonomy of Option Pricing Models’, *ICMA Centre Discussion Papers in Finance*, 2005
- [2] Andersen, L. and Andreasen, J., ‘Jump-Diffusion Processes: Volatility Smile Fitting and Numerical Methods for Option Pricing’, *Review of Derivatives Research*, no. 4, pp. 231-261, 2000
- [3] Avellaneda, M., Friedman, C., Holmes, R. and Samperi, D., ‘Calibrating Volatility Surfaces via Relative-Entropy Minimization’, *Applied Mathematical Finance*, vol. 4, no. 1, pp. 37-64, 1997
- [4] Ayache, E., Henrotte, P., Nassar, S. and Wang, X., ‘Can Anyone Solve the Smile Problem?’, *Wilmott Magazine*, Jan., pp. 78-96, 2004
- [5] Bakshi, G. and Kapadia, N., ‘Delta Hedged Gains and the Negative Volatility Risk Premium’, *Review of Financial Studies*, no. 16, pp. 527-566
- [6] Berestycki, H., Busca, J. and Florent, I., ‘Computing the Implied Volatility in Stochastic Volatility Models’, *Communications on Pure Applied Mathematics*, vol 57, no. 10, pp 1352-1373, 2004
- [7] Björk, T., ‘Arbitrage Theory in Continuous Time’, 2nd Edition, Oxford University Press, 2004
- [8] Black, F. and Scholes, M., ‘The Pricing of Options and Corporate Liabilities’, *Journal of Political Economy*, vol 81, no. 3, 1973
- [9] Bos, M. and Vandermark, S., ‘Finessing Fixed Dividends’, *Risk Magazine*, Sept., pp. 157-158, 2002
- [10] Brecher, D., ‘Pushing the Limits of Local Volatility in Option Pricing’, *Wilmott Magazine*, Sept., pp. 6-15, 2006
- [11] Clément, E., Lamberton, D. and Protter, P., ‘An Analysis of a Least Squares Regression Method for American Option Pricing’, *Finance and Stochastics*, vol. 6, no. 4, pp. 449-471, 2002
- [12] Coleman, T., Kim, Y., Li, Y. and Verma, A., ‘Dynamic Hedging With a Deterministic Local Volatility Function Model’, *Journal of Risk*, no. 4, pp.63-89, 2000
- [13] Coleman, T., Li, Y. and Verma, A., ‘Reconstructing the Unknown Volatility Function’, *Journal of Computational Finance*, vol. 2, no. 3, pp. 77-102, 1999
- [14] Cox, J., Ross, S. and Rubinstein, M., ‘Option Pricing: A Simplified Approach’, *Journal of Financial Economics*, no.7, pp. 229-263, 1979
- [15] Crépey, S., ‘Calibration of the Local Volatility in a Trinomial Tree Using Tikhonov Regularization’, *SIAM Journal on Mathematical Analysis*, vol. 34, no. 5, pp. 1183-1206, 2003
- [16] Crépey, S., ‘Delta Hedging Vega Risk?’, *Quantitative Finance*, vol. 4, no. 5, pp. 559-597, 2004
- [17] Derman, E., ‘Regimes of Volatility’, *RISK*, vol. 12, no. 4, pp. 55-59, 1999

- [18] Derman, E. and Kani, I., 'Riding on a Smile', *RISK*, vol. 7, no. 2, pp. 32-39, 1994
- [19] Derman, E. and Kani, I., 'Stochastic Implied Trees: Arbitrage Pricing with Stochastic Term and Strike Structure of Volatility', *International Journal of Theoretical and Applied Finance*, no.1, pp. 61-110, 1998
- [20] Derman, E., Kani, I. and Chriss, N., 'Implied Trinomial Trees of the Volatility Smile', *Journal of Derivatives*, vol. 3, no.4, pp.7-22, 1996
- [21] Derman, E., Kani, I., and Kamal, M., 'Trading and Hedging Local Volatility', *Quantitative Strategies Research Notes, Goldman Sachs & Co*, 1996
- [22] Doran, S., 'The Influence of Tracking Error on Volatility Risk Premium Estimation', Working Paper, 2006
- [23] Duchon, J., 'Interpolation des Fonctions de Deux Variables Suivant le Principe de la Flexion des Planques Minces', *RAIRO Analyse Numérique*, vol. 10, pp. 5-12, 1976
- [24] Dumas, B., Fleming, J. and Whaley, R., 'Implied Volatility Functions: Empirical Tests', *Journal of Finance*, vol. 53, no. 6, pp. 2059-2106, 1998
- [25] Dupire, B., 'Pricing and Hedging with Smiles', *Paribas Capital Markets Swaps and Options Research Team*, 1993
- [26] Dupire, B., 'Pricing with a Smile', *RISK Magazine*, no. 7, pp. 18-20, 1994
- [27] Engelmann, B. and Fongler, M., 'Better than its Reputation: An Empirical Hedging Analysis of the Local Volatility Model for Barrier Options', Working Paper, 2006
- [28] Fongler, M., 'Semiparametric Modeling of Implied Volatility', Springer-Verlag, 2005
- [29] Fongler, M., 'Arbitrage-Free Smoothing on the Implied Volatility Surface', Working Paper, 2005
- [30] Fries, C., 'The Foresight Bias in Monte-Carlo Pricing of Options with Early Exercise: Classification, Calculation & Removal', Working Paper, 2005
- [31] Frishling, V., 'A Discrete Question', *Risk Magazine*, Jan., pp. 115-116, 2002
- [32] Gatheral, J., 'The Volatility Surface: A Practitioner's Guide', John Wiley & Sons, 2006
- [33] Glasserman, P., 'Monte Carlo Methods in Financial Engineering', Springer-Verlag, 2004
- [34] Hagan, P., Kumar, D., Lesniewski, A. and Woodward, D., 'Managing Smile Risk', *Wilmott Magazine*, pp. 84-108, 2002
- [35] Hanke, M. and Rosler, E., 'Computation of Local Volatilities from Regularized Dupire Equations', *International Journal of Theoretical and Applied Finance*, vol. 8, no. 2, pp. 207-222, 2005
- [36] Haug, E., Haug, J. and Lewis, A., 'Back to Basics: a New Approach to the Discrete Dividend Problem', *Wilmott Magazine*, Sept., pp. 37-47, 2003

- [37] He, H., ‘Convergence from Discrete- to Continuous-time Contingent Claims Prices’, *Review of Financial Studies*, vol. 3, pp. 523-546, 1990
- [38] Heston, S., ‘A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options’, *Review of Financial Studies*, vol. 6, no. 2, pp. 327-343, 1993
- [39] Hull, J., ‘Options, Futures and Other Derivatives’, 6th Edition, Pearson Prentice Hall, 2006
- [40] Hull, J. and White, A., ‘The Pricing of Options on Assets with Stochastic Volatilities’, *Journal of Finance*, vol. 42, no. 2, pp. 281-300, 1987
- [41] Hull, J. and White, A., ‘Efficient Procedures for Valuing European and American Path-Dependent Options’, *Journal of Derivatives*, vol. 1, pp. 21-31, 1993
- [42] Labordère, P.H., ‘A General Asymptotic Implied Volatility for Stochastic Volatility Models’, Working Paper, 2005
- [43] Longstaff, F. and Schwartz, E., ‘Valuing American Options by Simulation: A Simple Least Squares Approach’, *The Review of Financial Studies*, vol. 14, no. 1, pp. 113-147, 2001
- [44] Meinguet, J., ‘Multivariate Interpolation at Arbitrary Points Made Simple’, *Zeitschrift für Angewandte Mathematik und Physik*, vol. 30, pp. 292-304, 1979
- [45] Merton, R., ‘Option Pricing when Underlying Stock Returns are Discontinuous’, *Journal of Financial Economics*, vol. 3, pp. 125-144, 1976
- [46] Merton, R., ‘Theory of Rational Option Pricing’, *Bell Journal of Economics and Management Science*, vol. 4, no. 1, pp. 141-183
- [47] Musiela, M. and Rutkowski, M., ‘Martingale Methods in Financial Modelling’, Springer, 1997
- [48] Obłój, J., ‘Fine-Tune Your Smile: Correction to Hagan et al.’, Working Paper, 2008
- [49] Rebonato, R., ‘Volatility and Correlation: the Perfect Hedger and the Fox’, 2nd Edition, John Wiley & Sons, 2004
- [50] Rubinstein, M., ‘Implied Binomial Trees’, *Journal of Finance*, vol. 69, no. 3, pp. 771-818, 1994
- [51] Samperi, D., ‘Calibrating a Diffusion Pricing Model with Uncertain Volatility: Regularization and Stability’, *Mathematical Finance*, vol. 12, no. 1, pp. 71-87, 2002
- [52] Stentoft, L., ‘Convergence of the Least Squares Monte-Carlo Approach to American Option Valuation’, *Management Science*, vol. 50, no. 9, pp. 1193-1203, 2004
- [53] Tsitsiklis, J. and Van Roy, B., ‘Optimal Stopping of Markov Processes: Hilbert Space Theory, Approximation Algorithms, and an Application to Pricing High-Dimensional Financial Derivatives’, *IEEE Transactions on Automatic Control*, no. 44, pp. 1840-1851
- [54] Tsitsiklis, J. and Van Roy, B., ‘Regression Methods for Pricing Complex American-Style Options’, *IEEE Transactions on Neural Networks*, vol. 12, no. 4, pp. 694-703, 2001

- [55] Vähämaa, S., ‘Delta Hedging with the Smile’, *Financial Markets and Portfolio Management*, vol. 18, no. 3, pp. 241-255, 2004
- [56] Vellekoop, M. and Nieuwenhuis, J., ‘Efficient Pricing of Derivatives on Assets with Discrete Dividends’, *Applied Mathematical Finance*, vol. 13, no. 3, pp. 265-284, 2006
- [57] Wahba, G., ‘Spline Models for Observational Data’, SIAM, 1990
- [58] Wahba, G. and Wendelberger, J., ‘Some new Mathematical Methods for Variational Objective Analysis Using Splines and Cross-Validation’, *Monthly Weather Review*, vol. 108, pp. 1122-1143, 1980
- [59] Wilmott, P., ‘on Quantitative Finance’, 2nd Edition, John Wiley & Sons, 2006
- [60] Zanger, D., ‘Convergence of a Least-Squares Monte Carlo Algorithm for Bounded Approximating Sets’, *Applied Mathematics Finance*, vol. 16, no. 2, 2009 (preprint)