

Implied Volatility Surface Construction

Erik Magnusson

erikpmagnusson@hotmail.se

A thesis presented for the degree
Master of Science and Engineering
in Engineering Physics



UMEÅ UNIVERSITET

Supervisor

Oskar Janson

oskar.janson@cinnober.com

Cinnober Financial Technology

Examiner

Markus Ådahl

markus.adahl@umu.se

Umeå University

February 16, 2018

Abstract

Implied volatility surfaces are central tools used for pricing options. This thesis treats the topic of their construction. The main purpose is to uncover the most appropriate methodology for constructing implied volatility surfaces from discrete data and evaluate how well it performs. First some methods and techniques in use for such surface constructing are presented. Then the most attractive approach, chosen to contain 4 interesting models is studied. The models' performances are tested on two price grids from the EURO STOXX 50 and Nikkei 225 indices. The found implied volatility surfaces give good and decent fits to the data, respectively.

The surfaces are evaluated in terms of presence of static arbitrage and are found to have it, although mostly for strike price and time to maturity combinations which are somewhat abnormal and rarely traded. Arbitrage is found to be more prevalent in surfaces when the input data is of lower quality.

The volatility surfaces' shapes and absolute values are compared in between models but also within models for some of the best-fit producing parameter sets. The surfaces are found to differ in both cases for some strike and maturity combinations - sometimes with relative differences of more than 10%. This suggests that surfaces with good fits to the input data still can produce distinctly differing prices for some options.

Calibrating the models with the chosen approach involves calculations with complex numbers in ways which potentially introduce problematic discontinuities due to branch crossings. This is investigated numerically as well as theoretically for the 4 models and found to be a significant problem in one of them. The three other models are found to avoid these problems under all valid parameter sets.

Sammanfattning

Implicita volatilitetsytor är centrala verktyg vid prissättning av optioner. Detta examensarbete utforskar skapandet av sådana ytor med målen att hitta det mest lämpliga tillvägagångssättet samt att utreda hur väl olika modeller presterar. Till att börja med presenteras några metoder som kan användas för att skapa volatilitetsytor. Sedan väljs det mest lämpliga tillvägagångssättet ut och implementeras med 4 olika modeller för den underliggande tillgången. Modellernas prestanda testas på två dataset med optionsspriser från indexen EURO STOXX 50 samt Nikkei 225. De erhållna volatilitetsytorna ger bra respektive tillräckliga överensstämmanden gentemot dessa dataset för alla 4 modeller.

Ytorna utvärderas genom att testas för statiskt arbitrage i närliggande punkter i ytorna och arbitrage hittas ofta, men mestadels för lösenpris- och löptidskombinationer som sällan handlas. Arbitrage hittas i större utsträckning i ytor skapade från data av relativt låg kvalitet.

De konstruerade ytornas form och absoluta belopp jämförs mellan olika modeller samt även inom enstaka modeller för parametersetet som genererar de bästa matchningarna mellan modellerna och det givna datasetet. Ytorna erhålls skilja sig i båda fallen för vissa lösenpris- och löptidskombinationer - ibland med relativt fel på över 10%. Detta tyder på att ytor som alla stämmer väl överens med de data de kalibrerats mot ändå kan ge tydligt skilda priser för vissa optioner.

Kalibrering av de 4 modellerna genom det valda tillvägagångssättet innebär utförande av beräkningar med komplexa tal på sätt som potentiellt kan leda till att principalgrenen korsas vilket i sin tur leder till felaktiga prissättningar. Detta utforskas teoretiskt och numeriskt i modellerna och konstateras vara ett tydligt problem i en av dem. De tre andra modellerna konstateras inte ge upphov till sådana problem för några tillåtna parameterset.

Acknowledgements

I want to send gratitude to Cinnober for letting me do my thesis at their Umeå office. I had a great time there - learning a lot and spending time with brilliant people. I direct a big "thank you" to my examiner Markus who presented thoughtful input to the work. Finally, I would like to send some special thanks to my supervisor Oskar who poured out valuable ideas whenever I wanted to discuss something.

Contents

1	Background	1
2	Introduction	2
2.1	Thesis structure and goals	2
2.2	What are implied volatility surfaces?	2
3	Empirically based properties of vol surfaces	4
3.1	Findings from the literature	4
3.2	Remarks	5
4	Theoretical conditions related to vol surfaces	6
4.1	Rules of thumb	6
4.1.1	Sticky strike	6
4.1.2	Sticky delta	6
4.1.3	Sticky square-root of time	7
4.1.4	Remarks	7
4.2	Bounds	7
4.2.1	Vanilla prices	7
4.2.2	Implied volatilities	8
4.2.3	Remarks	8
4.3	No-arbitrage conditions	9
4.3.1	Vanilla prices	9
4.3.2	Implied volatilities	10
4.3.3	Remarks	11
4.4	Discussion	11
5	Vol surface modelling approaches	13
5.1	The pricing integral	13
5.2	Monte Carlo simulations	13
5.3	Fourier transform representations	13
5.4	PDE pricing	14
5.5	Tree pricing	15
5.6	Directly modeling surfaces	15
6	On data input and models in general	16
6.1	Data quality	16
6.2	Time-dependent parameters	16
6.3	Calibration to exotics and smile dynamics	16
7	Models for the underlying asset	18
7.1	Local volatility	18
7.1.1	Examples	19
7.2	Stochastic volatility models	19
7.2.1	Examples	19
7.3	Discontinuous jump models	20
7.3.1	Examples	21
7.4	Combined models	22
7.4.1	Examples	22
7.5	Models with stochastic time	23
7.5.1	Examples	23
7.6	Models with historical path-dependence	23
7.7	Fitting higher order moments	24

8 Direct models for vol surfaces	25
8.1 Examples	25
8.1.1 SVI	25
8.1.2 RFSV	25
8.1.3 Historical vol surface decomposition	26
8.1.4 A combined vol surface and S_t model	26
9 My implementation	27
9.1 Model selection	27
9.2 The FFT approach	28
9.3 Chosen models	29
9.3.1 Heston and Heston jump	29
9.3.2 Barndorff-Nielsen-Shephard	30
9.3.3 CGMYSA	31
9.4 Arbitrage	32
9.4.1 Choosing arbitrage check	32
9.4.2 My implemented arbitrage check	32
10 Calibration procedure	34
10.1 The Nelder-Mead algorithm	34
10.2 The BOBYQA algorithm	35
10.3 The CMA-ES algorithm	35
10.4 Objective functions	36
10.5 Calibration validation and pitfalls	37
11 Results	40
11.1 EURO STOXX 50 data	40
11.2 Nikkei 225 data	52
11.3 On discontinuities from complex branch cut crossings	61
11.3.1 Heston and Heston Jump	61
11.3.2 BNS	61
11.3.3 CGMYSA	63
12 Concluding discussion	64
A EURO STOXX 50 call option data	A-1
B Nikkei 225 call option data	B-1

1 Background

Advancements in technology are driving a transformation in finance. Trading is being done in an ever-increasing diverse set of assets and with financial contracts on just about anything whose value is changing in time. One of the main reasons for the existence of such colorful high-pace markets is the subject of risk management, also called hedging, in which market participants aim to achieve their economical goals while being exposed to minimal risk. Hedging is mainly done by decreasing exposure to price movements in certain assets such as stocks, debt, commodities or derivatives by taking positions in derivatives, often the opposite of what one already possesses. An increasingly popular type of derivative is options which are used in hedging due to their suitability to hedge against adverse price movements, for example in equity and commodities.

More complex portfolios have triggered innovations for new derivatives aimed at further reducing risks. Market participants must decide upon what they think such instruments are worth - to quote bid and/or ask values. A problem is that it can be difficult to decide what a reasonable price for an intricate contract is. You can look at what others think an instrument is worth, look at its price history if such exists or have a look at its fundamentals. Methods like these involve some level of subjective thinking which in itself introduce riskiness. Scientists have been trying to eliminate this by attempting to price derivatives theoretically. By making assumptions on the dynamics of the markets and the actions of its participants, theoretical models have been developed where the most well-known and widely used is the Black-Scholes-Merton framework in which the price processes must satisfy the Black-Scholes differential equation. Analytical price expressions for normal (vanilla) options can be found which solve this equation. These are termed the Black-Scholes formulas and give theoretical prices for vanilla put and call options given information available in the option contract and the market. The only required input not directly observed in the market is the volatility of the underlying. To use the theoretical framework one needs to find an appropriate value for the volatility.

The Black-Scholes formulas were happily used for years after their proposal as a guiding tool for pricing vanillas, and maybe because of just that fact, they agreed well with market prices. After the crash in October 1987, however, something happened to the applicability of the model. The market had become more afraid of unusual events and generally put a higher price on riskier derivatives than before. This led to a magnification of phenomena called volatility smiles and skewes, which is a measure of the failure of market prices for vanillas to exactly agree on the assumptions in the theoretical Black-Scholes framework.

Using the Black-Scholes formula for vanillas in reverse one can find a volatility value corresponding to each option. Such volatilities are called implied volatilities and when plotting them against the strike prices of the corresponding options, one can often see that the plot gives a convex set of data points - resembling a smile or a skew in shape, where the Black-Scholes formula predicts a constant. Something similar is seen when checking the implied volatilities' dependence on the option's time to maturity. These effects are arising from the fact that the log-normal distribution of the price of the underlying used in the Black-Scholes formula fails to capture the real probability distribution. Thus smiles and skewes capture the gap between the most commonly applied pricing theory and real markets.

Knowledge of the structure and dynamics of implied volatility is valuable for pricing new options. To price new options, often the discrete implied volatility data from the market is used to construct a similar continuous implied volatility surface - with a value for the implied volatility at every strike and maturity. Such a surface can then be used to price options of any strike and maturity. Accomplishing construction of a surface requires there to be guiding principles for the interpolation and extrapolation of implied volatilities - the topic of this thesis.

2 Introduction

2.1 Thesis structure and goals

In this thesis, guidelines and models for constructing an implied volatility surface (henceforth vol surface) out of a discrete set of option prices are examined. The methods apply to all options, but there is a slight focus on options on equity and indices. Cinnober's standpoint is that they want to have better knowledge of implied volatility surfaces in general and how to create them since that might be something they would like to do on their own in the future.

The pre-study showed that a smattering of models for vol surface construction have been proposed and evaluated. In addition to this, theoretical explorations and empirical investigations on vol surfaces of differing underlyings have exposed interesting results which the models to some extent try to capture. I could however not immediately find any model which seemed superior to the rest and frankly some of the one's used in practice had serious flaws. Since I started this thesis having no knowledge of vol surfaces, I did not want to naively choose some models to investigate which I feared an experienced practitioner would quickly discard. Hence I decided to put some of the efforts in cruising through the literature, reading about proposed models and hopefully learn to discern what makes a vol surface model good or insufficient. Therefore, the first part of this thesis is basically a survey summary of things I have picked up from reading the vol surface literature, divided into sections of empirics, theoretical conditions and models for surface construction. After reading I evaluated some of the models using criteria I set up based on what I had learnt and picked a favorite approach to implement. I implemented this approach - actually involving 4 models - in Java and tested them on two datasets; call options on Nikkei 225 and EURO STOXX 50.

I think this thesis is better read by plowing through the chapters in order. If you are only interested in my implementation and my results, you can jump to section 9.

The remainder of section 2 introduces the subject in further detail. Section 3 presents some empirical results on vol surfaces from the literature. Section 4 presents theoretical conditions imposed on options which can be used when creating vol surfaces. Section 5 introduces the main approaches used for surface construction. Section 6 highlights some things to think about regarding the input data and models. Sections 7 and 8 give examples from two subcategories of surface construction approaches - namely whether to model the underlying asset or the surface directly. In section 9 I present my implementation with the approach and models I have chosen. In section 10 I present the calibration procedure in my implementation. Section 11 shows the results from the calibrations to the two datasets, with some comments included. Section 12 concludes and summarizes.

2.2 What are implied volatility surfaces?

For a call option with price C , its implied volatility can be found from the Black-Scholes formula below. This is done by inputting the price C , strike price K , time to maturity T , current price of the underlying asset S_0 , interest rate r and dividend yield q in,

$$\begin{aligned} C &= \Phi(d_1)S_0 - \Phi(d_2)Ke^{-rT}, \\ d_1 &= \frac{1}{\sigma\sqrt{T}} \left[\ln\left(\frac{S_0}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)T \right], \\ d_2 &= d_1 - \sigma\sqrt{T}, \end{aligned} \tag{1}$$

and then solving for the volatility σ . Here $\Phi(\cdot)$ is the cumulative normal distribution. Solving for σ cannot be done analytically although quickly by numerical methods. If the assumptions in the Black-Scholes model would apply, the implied volatility would be the same for options on the same underlying but different strike prices or times to maturity, but this is not observed in practice. Market participants have realized that the log-normal distribution cannot capture the probability of extreme events such as upswings, or especially crashes occurring while simultaneously capturing normal day to day changes in the underlying asset. In other words, the tails of the real probability distribution of the underlying (S_t) are relatively fat and might have skewness in contrast to the probability distribution implied by Black-Scholes - if this is calibrated to be similar to the real

distribution for normal days, see figure 1.

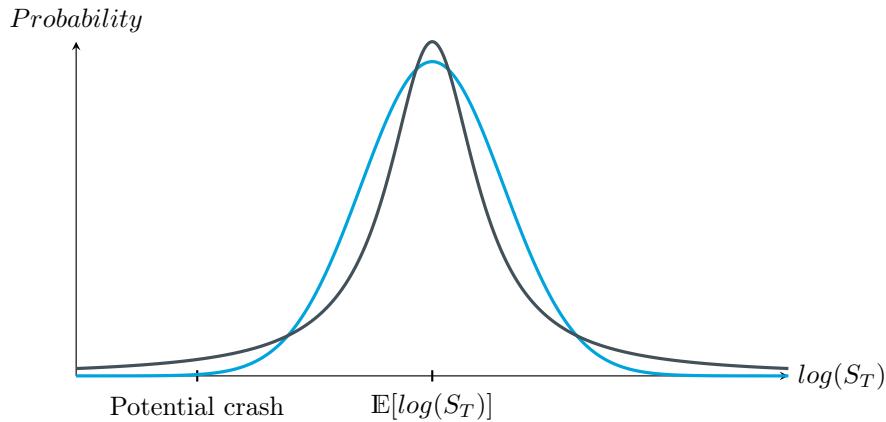


Figure 1: Illustration of a fat-tailed distribution represented by a Cauchy distribution (in black) compared to a normal distribution (in cyan). The probability of a crash is lower for the normal distribution.

The dissimilarity of the real distribution for $\log(S_T)$ and the normal distribution produces implied volatilities that change with strike and maturity. The goal of vol surface construction is to create a well-behaved vol surface defined for all positive K and T of interest - finding $\sigma(K, T)$, see figure 2. To do this one can use historical option prices such as a price grid for recently traded options (for example as in Appendix B) combined with other sources of information such as yield curves and dividend payouts.

A vol surface can be used for pricing vanilla and exotic options, which constitutes an important area in hedging and risk management. One can also extend the problem into finding a vol surface's time dependence. Having a surface which evolves appropriately in time is important for pricing some exotic options such as Napoleons and cliquets.

As will be seen in the following sections, a lot of research has been directed towards creating good vol surfaces, but how important is it to do it well? One indication of the importance can be seen by noticing the size of the gross market values¹ and notional amounts outstanding in the world's OTC markets for equity-related options; 347 and 6'761 billion USD respectively in the end of june in 2016, [1]. This can be compared to the global GDP in 2016 of 75'642 billion USD [2]. Given that vol surfaces are the main tools used to price such options, the importance of doing it right is of immense financial interest. If you find a way to create vol surfaces which are superior to others', you can make a lot of money at a low risk.

¹A sum of absolute values of gains and losses for all active contracts

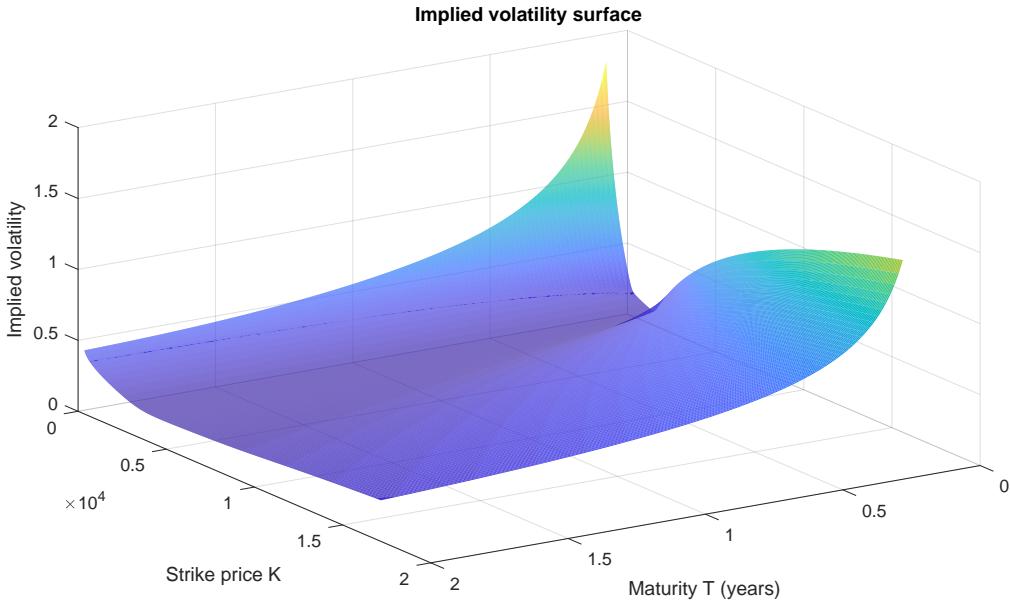


Figure 2: An example of an implied volatility surface. The underlying asset has a current price (S_0) of 2461.44.

3 Empirically based properties of vol surfaces

3.1 Findings from the literature

To get a feel for implied volatilities, we start off by looking at some results from empirical studies on them with the goal of finding out what properties they possess. The studies mentioned here are by no means meant to represent an exhaustive account.

In [3], they study time series of implied volatilities from vanilla prices on the S&P500 and FTSE indices. The following is my shortened version of their summary of the main properties they observed:

1. The vol surface is non-flat and has both term structure (T-dependence) and smile (K-dependence).
2. The vol surface display high autocorrelation and mean reversion.
3. Variations in vol surfaces can be explained by a small number principal components including: The overall shift, the slope and the curvature of the vol surface.
5. Changes in the vol surface cannot be exclusively described through changes in underlying.
6. Global changes in the vol surface are negatively correlated with changes in the underlying.
7. Relative movements within the vol surface have little correlation with the underlying.

In [4] they investigate the implied volatilities from S&P500 from March 2, 2000 to February 2, 2001. They model the them as state variables and the constructed vol surface was seen as a random surface and decomposed using Karhunen-Loëve decomposition. They find that only a few orthogonal components are needed to describe the surface. The 3 main factors could explain 90%, 5% and 1.3% of the data, respectively. They also found indications of the leverage effect - an overall negative correlation between implied volatility and the underlying.

Friction in the markets have been discussed as being part of the explanation for the shapes of vol surfaces. Transaction costs and illiquidity make sure that the notion of arbitrage is complicated and implies several feasible prices. [5] proposes an "unrestricted Black-Scholes model" and in the study of the S&P100 the author concludes "that transaction costs and liquidity effects play a major role in the valuation of index options" (p. 1093).

There should not be any noticeable difference between the implied volatilities of calls and puts

if put-call parity applies. In [6] they observe gaps in the implied volatility scatter plots for DAX put and call options from 1995 to 1999. This the authors state "...can be traced back to a biased index level caused by dividend payments" (p. 9). When the price of the index is increased accordingly, the implied volatilities agree.

In [6] they also analyze option volatilities on DAX. They use data from 1995 to 1999 and conclude that the correlation between at the money-volatility and the volatility index VDAX is almost 1 (0.997). Hence the IVS has a deterministic "point" in time. They use a spline regression model which explains about 95% of the variation in the IVS. They do an extensive study, both in modeling and in thought but conclude "All in all, though the smile pattern can be estimated with great precision, our attempt to uncover the economic variables underlying the dynamics of the smile was not fully successful.", (p. 32). They also mention that the changes in implied volatilities are supposed to vary with the market participants perception of the crash risk, which they point out is naturally difficult to model and predict.

In [6], they find that differences between the implied volatilities for degrees of moneyness of 0.95 and 1.0, as well as 1.0 and 1.05, can be modelled accurately with an AR(1)-process with an auto-correlation coefficient exceeding 0.98 throughout the sample period. This means that the shocks in the smile die out slowly. They find that proxies for volatility of volatility and percentage trading volume of out of the money puts seem to significantly positively affect the degree of skewness in implied volatilities.

In [7], many interesting points on arbitrage in option data are brought up. Computing mid points out of bid-ask prices and collecting data over an entire day can introduce arbitrage even though there were actually none. They mention that the data used for vol surface construction more often contain arbitrage than not.

In [8], they point out that modelling a vol surface is mainly difficult due to the sparse option data. When plotted, the data often looks like strings of pearls of data with strikes for maturities of say 1,2,3,6,9,12 and 24 months. No matter what method you choose, there is going to be gaps where the accuracy of any model can be questioned.

In [9], they analyze the Spanish index IBEX-35 from 1994 to 1996 and present several findings. They find indications for that the smile is statistically different on mondays compared to other week days and that option bid-ask spreads are higher on mondays. They find indications for that the curvature of the smiles are lower in the end on the week than in the beginning. They investigate the effect of several economic variables on smiles and find that the curvature is positively and significantly correlated to bid-ask spreads and negatively and significantly correlated to time to maturity and historical volatility in the underlying asset. Also they find indications for that when the stock market level is better than in the past, on average the curvature of the smile increases. They find indications for trading activity being positively correlated to at-the-money volatility and negatively correlated with the slope of the smile. They do Granger tests and find that the curvature is Granger-caused by bid-ask spreads and vice versa. There is also unidirectional causality between the relative market momentum and the at-the-money implied volatility.

3.2 Remarks

Vol surfaces have many observed properties and as many of these as possible should be included in the models for accuracy. It seems that modeling of vol surfaces is manageable but bringing to light the driving factors of the surface is challenging. It seems to be beneficial to include bid-ask spreads and traded volume in vol surface construction since they have been found to significantly affect the surfaces. Taking into account historical vol surfaces in the modeling of the current surface seems like a good idea given its autocorrelation and mean-reversion. The input data needs to be of good quality, with option prices and price of the underlying gathered at nearby times.

4 Theoretical conditions related to vol surfaces

In this section, theoretical results from the literature on implied volatilities and vanilla prices are presented. I have included constraints and conditions on vanilla prices because of the available map between prices and implied volatilities. The goal of this section is to find conditions for checking whether a surface, once created through fitting to market data, is internally consistent and does not allow unrealistic trading.

A common approach for theoretical investigations in trading of assets is to check if riskless profits can be made through trading. This is called arbitrage and will be of interest throughout this thesis. There are a few types of arbitrage:

Arbitrage: The existence of arbitrage is equivalent with the existence of a portfolio of no initial cost which is guaranteed to not make a future loss and has a non-zero probability of making a future profit. Absence of **static arbitrage** means there is no way to set up a portfolio today in the existing price grid which might generate a profit, without any risk of a loss. If no riskless profit can be made even when the portfolio composition is allowed to change with time, there is absence of **dynamic arbitrage**.

A few "rules of thumb" have been used in practice for describing the evolution of vol surfaces. Such rules could be of good use when one cannot calibrate a new surface to very recent data but has to extrapolate from a previous surface. These rules are presented next, including whether they can be consistent with absence of arbitrage.

4.1 Rules of thumb

The rules of thumb are attempts for describing short-term evolution of vol surfaces in simple ways. The results in this section are based on the work in [10]. In it, the authors assume that the underlying follows a geometric Brownian motion with deterministic dividends and interest rates and that the volatility follows a diffusion process. They further assume that the process of the square of the implied volatility is a linear combination of Wiener processes and a drift term.

4.1.1 Sticky strike

In the sticky strike rule, the implied volatility (σ) does not change with the underlying. In one version of it, σ varies only with K and T . In another version, it is independent of S_t , but possibly dependent on other stochastic variables.

In [10], they theoretically showed that both of these versions of the sticky strike rule is incompatible with any type of smile or skew. They also tested the sticky strike on over-the counter S&P500 options using a model with 6 parameters resulting in an unimpressive coefficient of determination; $R^2 = 27\%$.

4.1.2 Sticky delta

In the most basic form of the sticky delta rule, σ is assumed to be a deterministic function of K/S_t (moneyness) and $T-t$, where t is the current time. The generalized sticky delta rule is where σ depends on K , S_t , T , and t only through K/S_t and $T-t$. Another version, sometimes used by traders, is the relative sticky delta rule which specifies that $\sigma_{K,T}(S_t, t) - \sigma_{F_t,T}(S_t, t)$ is only a function of K/F_t and $T-t$, where F_t is the value of a future or forward contract on S_t , with delivery at T .

In [10], they theoretically showed that the basic form of the sticky delta rule is not internally consistent if there are smiles or skews. The generalized sticky delta rule can be consistent with absence of arbitrage. The relative sticky delta rule can be approximately consistent with no-arbitrage. They tested the relative sticky delta using a model with 6 parameters on the S&P500 options and found a good fit; $R^2 = 94.93\%$.

4.1.3 Sticky square-root of time

The square-root of time rule has sometimes been used by traders. One version of it says that $\frac{\sigma_{K,T}(S,t)}{\sigma_{F,T}(S,t)}$ should only be a function of $\frac{\ln(K/F)}{\sqrt{T-t}}$. Another version is similar and says that $\sigma_{K,T}(S,t) - \sigma_{F,T}(S,t)$ should only be a function of $\frac{\ln(K/F)}{\sqrt{T-t}}$.

The square-root of time rules can be approximately consistent with no-arbitrage according to [10]. They tested the square-root of time rule using a model with 4 parameters on the S&P500 options and found a good fit; $R^2 = 97\%$.

4.1.4 Remarks

The relative sticky delta and the sticky square-root of time rules can be used for describing the dynamics of vol surfaces in between vol surface constructions. When doing so, of course the exact function of the rules must be specified.

The sticky square-root of time is quite similar to the recent results displayed in [11], that σ at-the-money decreases with T as $\frac{1}{T^\alpha}$ for $\alpha \in (0.3, 0.5)$, α usually being close to 0.4. That further justifies the use of this rule.

4.2 Bounds

Now let's look at some bounds for option prices and implied volatilities which can be found useful when evaluating the constructed surfaces. The bounds are mainly based on no-arbitrage arguments. For their motivations and derivations, I refer to the references. The conditions presented for calls can be transformed into conditions for puts and vice versa using put-call parity (equation 6). The mentioned conditions by no means constitute an exhaustive account. They serve to show examples of what can be said to hold for prices and implied volatilities derived from theoretical reasoning.

4.2.1 Vanilla prices

- In [12], equations 2-5 are derived assuming,
 1. There are no transaction costs.
 2. All trading profits are subject to the same tax rates.
 3. Borrowing and lending are possible at the risk-free rate.
 4. There are no arbitrage opportunities.

$$\max(S_0 - Ke^{-rT}, 0) \leq C(K, T) \leq S_0, \quad (2)$$

$$\max(Ke^{-rT} - S_0, 0) \leq P(K, T) \leq Ke^{-rT}. \quad (3)$$

Here C and P are the prices of vanilla calls and puts respectively. S_0 is the current price of the underlying asset and r the risk-free rate. If the underlying is paying dividends, the lower bounds change, becoming

$$\max(S_0 - D - Ke^{-rT}, 0) \leq C(K, T), \quad (4)$$

$$\max(D + Ke^{-rT} - S_0, 0) \leq P(K, T), \quad (5)$$

where D is the present value all future dividends.

- If there is no arbitrage, the following equation (called put-call parity) holds:

$$C(K, T) + D + Ke^{-rT} = P(K, T) + S_0, \quad (6)$$

- From [13]:

$$\max(0, Se^{-r_f T} - Ke^{-r_d T}) \leq C(t, T, K) \leq S_t e^{-r_f T} \quad (7)$$

where the domestic rate r_d and foreign rate r_f can be set equal for non-FX options.

- Using the expression of a call price in terms of its probability distribution and the fact that the payoff is always lower for a higher strike price, [14] motivate,

$$-e^{-r_d T} \leq C'(K) \leq 0, \quad (8)$$

$$C''(K) \geq 0. \quad (9)$$

Furthermore, in [15], it is shown that for $C(K)$,

$$C(0) = Se^{-r_f T}, \quad (10)$$

$$C'(0) = -e^{-r_d T}, \quad (11)$$

$$C''(0) = 0. \quad (12)$$

- Motivated by its portfolio's payoff, in [14] it is shown that,

$$\frac{\partial(\frac{P(K)}{K})}{\partial K} \geq 0 \quad (13)$$

which stated in terms of calls, using put-call parity, is

$$\frac{\partial C(K)}{\partial K} \geq \frac{C(K)}{K} - \frac{Se^{r_f T}}{K}. \quad (14)$$

This is a stronger lower bound than in equation 8.

4.2.2 Implied volatilities

- In [15] and [16] the authors obtain bounds for $\sigma(K, T)$. Assuming $\frac{\partial C}{\partial K} \leq 0$,

$$-\frac{N(-d_2)}{KN'(d_2)\sqrt{T}} \leq \frac{\partial \sigma(K, T)}{\partial K} \leq \frac{N(d_1)}{KN'(d_1)\sqrt{T}} \quad (15)$$

where d_1 and d_2 are the same as in the Black-Scholes formula (equation 1).

- From [14], setting $\frac{\partial^2 C}{\partial K^2} \geq 0$ is equivalent to,

$$\frac{\partial^2 \sigma(K)}{\partial K^2} \geq \frac{1}{K^2 \sigma(K) T} - \frac{2d_1}{\sigma(K) K \sqrt{T}} \frac{\partial \sigma(K)}{\partial K} - \frac{d_1 d_2}{\sigma(K)} \left(\frac{\partial \sigma(K)}{\partial K} \right)^2 \quad (16)$$

- In [17], upper bounds for σ at large and small strikes are found. The large strike rule is: There exists $x_0 > 0$ such that for all $x > x_0$,

$$\sigma(x, T) < \sqrt{2|x|T} \quad (17)$$

where $x = \log(K/F_0)$ and F_0 is the discounted expected value of S_T . The small strike rule is: For any $\beta > 2$, there exists a x_0 , such that for all $x < x_0$,

$$\sigma(x, T) < \sqrt{\beta|x|T}. \quad (18)$$

4.2.3 Remarks

The price and vol bounds introduce quick checks for the soundness of a vol surface. They are hardly difficult to implement, but the question is whether there are more far-reaching checks. Let's see.

4.3 No-arbitrage conditions

Closely related to the bounds in section 4.2 are conditions for absence of arbitrage. These could be used for checking for presence of arbitrage in surfaces. One or two (depending on how you count) of them will be implemented in this thesis.

4.3.1 Vanilla prices

4.3.1.1 Positivity of spread prices

It was shown in [18], building on works by [19] and others, that vanilla prices are free of static arbitrage if butterfly spreads, calendar spreads and vertical spreads all have non-negative prices. This is a powerful result, but the proof they give is based on some unrealistic assumptions. These assumptions are:

- The option prices are on a rectangular grid with no missing data, such that there is a price $C(K_i, T_j)$ for all indices i and j on the grid. The grid has a finite number of times to maturity and countably infinitely many strikes.
- Continuous trading is allowed.
- There are no dividends and no interest rates.
- Call prices tend to 0 as strikes go towards ∞ .

The three spreads are described next.

A vertical spread involves two options of the same maturity but different strikes. Let the strikes be denoted K_i such that they are indexed by i , where $i = 1, 2, 3, \dots$ and the maturities T_j , indexed by $j = 1, 2, 3, \dots, n$. Let $K_{i+1} > K_i$ and $T_{j+1} > T_j$, then the price of a vertical spread is

$$VS = v \times [C(K_i, T_j) - C(K_{i+1}, T_j)], \quad (19)$$

where $v \in \mathbb{R}^+$.

A calendar spread involves two options of the same strike but differing maturities. Its price is:

$$CS = c \times [C(K_i, T_{j+1}) - C(K_i, T_j)], \quad (20)$$

where $c \in \mathbb{R}^+$.

A butterfly spread involves three options with the same maturity but different strikes. Its price is:

$$BS = b \times \left[C(K_i, T_j) - \frac{K_{i+2} - K_i}{K_{i+1} - K_i} C(K_{i+1}, T_j) + C(K_{i+2}, T_j) \right] \quad (21)$$

where $b \in \mathbb{R}^+$.

4.3.1.2 Continuous checks

In [20], necessary conditions for absence of static arbitrage are derived based on assumptions which are similar to those in 4.3.1.1:

- Perfectly liquid vanillas with non-negative prices existing for all real $K > 0$ and $T > 0$.
- There are no interest rates and no dividend yields.
- There are no transaction costs.

The author show that if and only if the following 5 conditions hold:

- (A1) Convexity in K : $C(K)$ is convex for all T .
- (A2) Monotonicity in T : $C(T)$ is non-decreasing for all K .
- (A3) Large strike limit: $C(K, T) \rightarrow 0$ when $K \rightarrow 0$.
- (A4) Bounds: $(S_0 - K)^+ \leq C(K, T) \leq S_0$
- (A5) Expiry value: $C(K, 0) = (S_0 - K)^+$,

the price surface $C(K, T)$ is free of static arbitrage:

4.3.1.3 Vector-matrix multiplication

In [21], a set of arbitrage conditions is based on the assumptions:

- The market is frictionless.
- There is a finite number of European call options in the market.
- Trading can be done with the underlying asset itself and its futures.

It is expressed that a set of call prices, all with the same maturity, with prices $C(K_i)$ is free of arbitrage if and only if

$$\mathbf{v}\mathbf{A}^{-1} > 0 \quad (22)$$

where

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 1 & K_1 & 0 & 0 & 0 & \cdots & 0 \\ 1 & K_2 & K_2 - K_1 & 0 & 0 & \cdots & 0 \\ 1 & K_3 & K_3 - K_1 & K_3 - K_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

and

$$\mathbf{v} = (e^{-rT}, S_0 e^{-qT}, C(K_1), C(K_2), \dots, C(K_N))$$

where q is the dividend yield and r the interest rate. This is equivalent to requiring that all combinations of portfolios which can exist in the set of options together with the underlying and its future cannot produce an outcome with a non-zero probability of a gain and a negative price without a non-zero probability of a loss.

4.3.1.4 Checking conditions on support functions

In [22] a condition for absence of arbitrage is given. It allows for finite amounts of strikes and maturities. The assumptions are:

- The market is frictionless: Rates for borrowing and lending are the same and trades can be made in arbitrary amounts at no costs.
- Dividend yields are deterministic and interest rates have no volatility.

The theorem the authors present is more liberal in its assumptions but probably more difficult to implement compared to the one's in the previous paragraphs. It requires checking conditions on what are called support functions for transformed prices as a function of K/F . The condition is able to distinguish model-independent arbitrage from weak arbitrage opportunities. I stop here and refer to the arbitrage condition in its entirety in [22] for a better explanation.

4.3.2 Implied volatilities

There are not many arbitrage checks known directly applicable to vol surfaces. I will present one of them next.

4.3.2.1 Necessities for absence of static arbitrage

In [20], sufficient and almost necessary conditions for a vol surface to be free of static arbitrage is presented. It is based on the following assumptions:

- The market for call options is perfectly liquid with non-negative prices existing for all real $K > 0$ and $T > 0$.
- Interest rates and dividend yields are zero.
- There are no transaction costs.

Consider the function $\Xi(x, T)$, which is connected to the implied volatility $\sigma(K, T)$ through $\Xi(\log(K/S), T) = \sqrt{T}\sigma(K, T)$. The (six) conditions are written in terms of this transformed version of implied volatility (Ξ):

- (IV1) Smoothness: For every $T > 0$, $\Xi(\cdot, T)$ is twice differentiable.
- (IV2) Positivity: For every $x \in \mathbb{R}$ and $T > 0$, $\Xi(x, T) > 0$.
- (IV3) Durrleman's Condition: For every $T > 0$ and $x \in \mathbb{R}$,

$$0 \leq \left(1 - \frac{x}{\Xi} \frac{\partial \Xi}{\partial x}\right)^2 - \frac{1}{4} \Xi^2 \left(\frac{\partial \Xi}{\partial x}\right)^2 + \Xi \left(\frac{\partial^2 \Xi}{\partial x^2}\right)^2$$

where $\Xi(x, T)$ is written Ξ .

- (IV4) Monotonicity in T : For every $x \in \mathbb{R}$, $\Xi(x, \cdot)$ is non-decreasing.
- (IV5) Large moneyness behaviour: For every $T > 0$,

$$\lim_{x \rightarrow \infty} \sup \left(\frac{\Xi(x, T)}{\sqrt{2x}} \right) \in [0, 1)$$

- (IV6) Value at maturity: For every $x \in \mathbb{R}$, $\Xi(x, 0) = 0$.

- If all conditions (IV1)-(IV6) are satisfied by $\Xi(\log(K/S), T)$, then the call prices found from the implied volatilities $\sigma(K, T)$ constitute a price surface which is free of static arbitrage.
- If $\Xi(\log(K/S), T)$ satisfy (IV1) and (IV2) but violates any of the remaining conditions (IV3)-(IV6), the call prices do not constitute a surface which is free from arbitrage.

In conclusion, if the assumptions apply and a vol surface passes the tests (IV1)-(IV6), one can be certain that it does not contain any static arbitrage opportunities. If (IV1) and (IV2) are satisfied, the rest of the conditions are necessary for there to be no arbitrage.

4.3.3 Remarks

There are promising checks out there which can be used in testing for static arbitrage in a vol surface, as the previous paragraphs display. If these are incorporated in the construction, one can automatically get self-consistent surfaces. I could not find any conditions for avoiding other types of arbitrage - such as dynamic arbitrage.

4.4 Discussion

Many of the bounds and conditions in this section depend on assumptions which are clearly not satisfied in practice. Therefore, I'm pondering the question: Can they be used anyway if with caution? Finance is based on people's often irrational decisions and is hence not a very exact science. There is something to what Paul Wilmott wrote, "every financial axiom... ever seen is demonstrably wrong. The real question is how wrong...", [23]. One can only debunk conditions such as the one's presented, never regard them as true. On the other hand, some assumptions might only be introduced in the efforts to produce a rigorous mathematical proof of a proposed condition - they might actually have no effect on the validity of the condition. In this light, I don't simply discard these nice results but handle them with thoughtfulness.

The assumptions are not simply valid or not but are better visualized as on a spectrum from being generally regarded as fulfilled to being far from fulfilled. On the far end of the former side I would say observables such as deterministic dividends or the ability to trade in future contracts are located. Then comes assumptions such as non-existence of arbitrage and existence of a continuous set of market options. Further to the latter side are more restrictive assumptions including deterministic interest rates and even no interest rates. It is hard to say how applicable conditions with such restrictive assumptions are without furthering the theoretical results.

The conditions presented suggest that there are ways to tell if a set of vanilla prices - or for that matter a constructed complete vol surface is free of static arbitrage. This is good news indeed. Some of the conditions, though, in particular ones involving derivatives are suited for continuous vol surfaces and not directly applicable to market data.

The conditions are not tight enough to be able to completely determine how the interpolation, extrapolation and transformation of the market prices into a vol surface should be performed. To construct a surface, specifically designed methods for surface construction are needed. This is the subject of the next section.

5 Vol surface modelling approaches

In this section, common approaches to vol surface construction found in the literature are presented. It is a non-exhaustive account, but, includes many of the approaches used in practice today.

To start from the beginning, the future payoff of a European vanilla call should determine its price - which determines its implied volatility. The payoff of a vanilla call option is simply,

$$\text{Payoff(European call)} = \max(S_T - K, 0) \quad (23)$$

Knowledge of the future price of the underlying, S_t , is then what stands in our way from being able to find the payoff at any K and T . If we can determine the distribution of the underlying at all times, we can find the payoffs of all vanilla calls on the underlying. Then these payoffs can be converted to prices and then inverted to a continuous vol surface.

This motivates the common approach to constructing vol surfaces: A model for the underlying is created which determines S_t up to a number of model parameters. Next, the model's parameters are calibrated to make the model's prices match the market's option prices, either theoretically or using simulations. The calibrated model parameters now give the complete vol surface.

Next, a few different ways of calculating prices of vanilla option prices based on equation 23 from a model for S_t are presented. Each approach applies better to certain types of S_t -models. Only one of them will be implemented in this thesis.

5.1 The pricing integral

The price of a call option can be stated as an integral over the future price of the underlying:

$$C(K, T) = e^{-rT} \int_K^{\infty} (S_T - K) p(S_T) dS_T, \quad (24)$$

where $p(S_T)$ is the probability density function of S_t at time T . When one can find explicit expressions for S_t , and $p(S_T)$ for a model of choice, equation 24 can be used to find all call prices. Finding the expressions and solving the integral can be difficult for complex models of S_t , but when it can be done accurately, this approach to pricing is efficient.

5.2 Monte Carlo simulations

When S_t exhibit stochastic behaviour, vanilla prices can be found by running Monte Carlo simulations on S_t . A pricing equation often used is,

$$C(K, T) = e^{-rT} \mathbb{E}^Q [\max(S_T - K, 0)]. \quad (25)$$

where the expectation value is taken with respect to a risk neutral measure. This method works regardless of the complexity of the model but it can be computationally demanding and time consuming. Another drawback is that the prices arising from simulations are not exact but can vary between simulations. Such variations are likely to lead to arbitrage-opportunities in the vol surface.

Monte Carlo methods work well for pricing path-dependent exotic options relative to other existing approaches. If prices of exotics are to be included in the vol surface construction, one might want to consider using Monte Carlo simulations.

5.3 Fourier transform representations

When the probability density function of S_t in a model is too complicated to be found or allow for evaluating equation 24 satisfactory, it is often simpler to use pricing based on what are called characteristic functions, defined as follows:

When the probability distribution p of a stochastic process $\log(S_t)$, exists, the process' characteristic equation $\phi_t(u)$ is the Fourier transform of its probability distribution,

$$p(\log(S_t)) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iux} \phi_t(u) du \quad (26)$$

Another way to define the characteristic function of a process $\log(S_t)$ is simply as an expected value,

$$\phi_t(u) = \mathbb{E}[e^{iu \log(S_t)}] \quad (27)$$

Hence, for every probability density function there is a one-to-one correspondence to its characteristic equation, defined by the two equations above.

There are two main methods of vanilla option pricing using characteristic equations and Fourier transforms. The first one, introduced in an option pricing context by Heston in 1993, is described in [24]. Unfortunately, singularities arise in the method's integrals which means the Fourier transform cannot be applied directly.

The authors of [25], introduce two versions of another method in which prices are found from Fourier transforms. The one involving exponentials is briefly described in this section. First, the authors multiply the call price $C(k, T) = C_T(k)$ with an exponential factor:

$$c_T(k) = e^{\alpha k} C_T(k), \quad (28)$$

where $k = \ln(K)$, such that $\psi_T(v)$ is the Fourier transform of the modified price, $c_T(k)$,

$$\psi_T(v) = \int_{-\infty}^{\infty} e^{ivk} c_T(k) dk. \quad (29)$$

This is done to get a square-integrable function in the integral for the call price for large enough α ,

$$C_T(k) = \frac{\exp(-\alpha k)}{\pi} \int_0^{\infty} e^{-ivk} \psi_T(v) dv, \quad (30)$$

where $\psi_T(k)$ is shown to be

$$\psi_T(k) = \frac{e^{-rT} \phi_T(v - (\alpha + 1)i)}{\alpha^2 + \alpha - v^2 + i(2\alpha + 1)v}. \quad (31)$$

When the characteristic equation of S_t is known, prices for all K and T can be found by taking the Fourier transform of equation 30 together with 31.

5.4 PDE pricing

Expressions can be found where the vanilla prices are stated as solutions to partial differential equations. As an example, using the Black-Scholes dynamics,

$$dS_t = \mu S_t dt + \sigma S_t dW$$

and creating a risk-less portfolio, [12] arrives at the Black-Scholes-Merton equation,

$$\frac{\partial C}{\partial t} + rS_t \frac{\partial C}{\partial S_t} + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 C}{\partial S_t^2} = rC. \quad (32)$$

By solving this with the derivative-specific boundary conditions for a vanilla option, the Black-Scholes formula emerges. For other dynamics of S_t , similar PDE's can be found and solved numerically or analytically - giving prices for all K and T .

5.5 Tree pricing

The evolution of the price of an underlying can be described using paths in a tree where the price starts at the root and can transition into a few new values by following the branches in the tree. The case of lowest branching dimension is a binomial tree where from one point in time to the next, the underlying can move up or down, see figure 3. For a time step like this, a risk-free portfolio of the underlying and a vanilla option can be set up. The relative sizes of the current price and the two potential future prices determine the amount of vanillas there should be in a portfolio for each underlying. Such a risk-free portfolio must earn the risk-free rate and from this the option price can be determined. In this way options can be priced for different strikes and maturities.

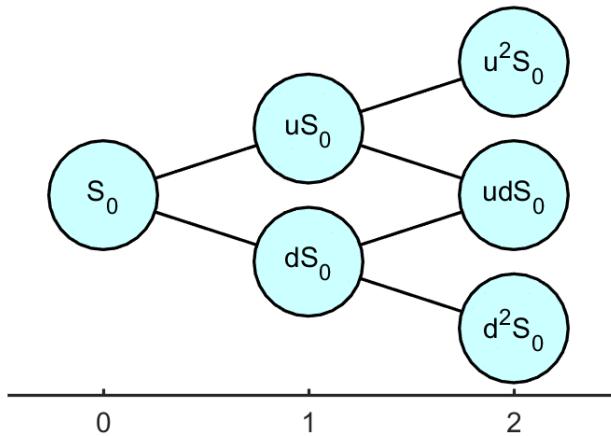


Figure 3: Example of a tree with a branching of 2 displayed for 2 time steps. Multiplication with the factors u and d represent up and down movements in S , respectively.

5.6 Directly modeling surfaces

In section 3.1 as well as in [26], it was pointed out that option prices can be affected by factors beyond the underlying. In addition, the information contained in option prices is not enough to infer the dynamics of the underlying. This is a fundamental reason for why the problem of calibrating models for S_t to option prices in general is ill-posed.

This problem can be avoided if the vol surface is modeled directly without paying attention to the underlying. The approach is often carried out by setting some functional form for the surface and optimizing the parameters to fit the model to the market prices in the most appropriate way. One way is to set the surface to be a sum of orthogonal surfaces. Another is to use parametric functions of either strike or maturity and use some simpler interpolation strategy in the other dimension.

6 On data input and models in general

Before going into models for vol surface construction, I want to bring forth some points regarding data input and vol surface modelling in general.

6.1 Data quality

Deciding upon what data should be used as input in vol surface construction is not that straightforward. Here are some points to take into consideration.

- If prices for an entire day are saved to later be used in vol surface construction together with one value representing the price of the underlying for the day, arbitrage might arise which was not there when the trades took place.
- To get prices where there are none, an idea could be to calculate mid prices from bid and ask spreads. This calculation might lead to data of insufficient quality.
- There is an inherent weighting problem within the option price grid. Some options are traded very frequently and these data are probably more reliable compared to infrequently traded options. Perhaps weighting with respect to traded volume or number of trades in a given time period leads to more realistic vol surfaces.
- For options with low prices, the tick size might be a significant fraction of the price and thus produce large uncertainties in prices.
- Option data points with high bid-ask spreads are likely less reliable since the market is uncertain of their prices.

6.2 Time-dependent parameters

When creating a model for some phenomenon, sometimes time-dependence are introduced in model parameters. This is an indication that the underlying process is not accurately captured by the model and that the modeller is trying to save the model by incorporating inhomogeneity. Take for example the modelling of the orbits of the planets around the sun. The orbit of the earth around the sun could be modelled by almost any function, for example, in any reference system, $\mathbf{r} = a(t)x^2\hat{\mathbf{x}} + b(t)(\cos(y) + 2)\hat{\mathbf{y}} + c(t)e^{3z+2x}\hat{\mathbf{z}}$ could make a perfect fit to the past path if the parameters a, b and c are calibrated well, but that does not make the model good because it would be difficult to use it to produce accurate future predictions. Another model for the problem could be set up solely based on Newton's law of gravity which has no time-inhomogeneity. Newton's law will be simpler and captures the underlying dynamics of the problem without any parameters to be fitted, except if you will, the masses of the surrounding planets and the sun.

Moreover, getting an accurate fit to the past and present with a homogeneous model is a strong indication for that the structure of the underlying problem has been captured.

6.3 Calibration to exotics and smile dynamics

A vol surface is uniquely determined if there is reliable information to calculate all future prices of the underlying asset. Unfortunately there is not. Instead, to get the best possible vol surface, one could argue that as much information about the future of the underlying should be included in the construction process.

Exotic options which depend on the future path of the underlying contain information about the underlying asset's expected price process which vanillas are missing. The price of a European vanilla option is only determined by the probability distribution of the underlying at maturity. The price of an exotic option on the other hand, for instance a barrier option, depends on the transitional probabilities between all possible price paths of the underlying until maturity. This is a significantly larger chunk of information and is a reason for why it is much more difficult to price exotics. It is also a reason for the inclusion of reliable exotic prices in model calibration for the underlying asset. To do it, a map between the vol surface or the underlying asset and the exotic

option of interest is needed. Such a mapping can in some cases be found semi-theoretically, but in many cases, one has to resort to numerical methods, for instance Monte Carlo simulations.

When calibrating to exotics, similar measures of fit as for vanilla market prices can be used. It may be suitable to put some weight on exotic option data to either make these points more or less influential than vanilla option points, depending on the relative quality of the data.

In [27] together with [28], the authors bring up the relevant problem that it is one thing to calibrate a model of a vol surface to vanilla prices and something completely different to get the dynamics of a vol surface right and price path-dependent exotic options accurately. They also point out the huge parametrization step done when choosing a model, because then all the complexities in the market is reduced to a few model parameters which are to produce fits for all prices. The authors mention that the difficulties with pricing exotic options does not lie in that models differ slightly in the pricing of vanillas and that these differences blow up when pricing exotics. It is rather that different models infer different translational probabilities for future prices. In other words, the models have differing smile dynamics which does not show up when pricing vanillas.

Another point worth mentioning is that there are options (including barriers, Napoleons and cliques) whose prices depend on the evolution of vol surfaces, so if one's vol surface does not evolve in a realistic way, one might be arbitAGED.

7 Models for the underlying asset

In this section I present a few commonly used models for the underlying asset (denoted S_t) found in the literature followed by examples of their implementations and performances in practice. The models' main purposes here are to be used within the pricing approaches of the previous section to construct vol surfaces. Some of the models will be implemented in this thesis.

The majority of the studies I encountered were on foreign-exchange rates, equities and equity indices. For these underlyings, any models presented in this section can be used for better or for worse. When the underlying asset is more intricate, one just have to make sure that the model could capture the properties of the evolution of the asset price somehow. For instance, for an interest rate, it might be appropriate to not have a drift term, but only mean-reversion.

I have included many examples in this section which I read about in the literature study. If you get sick of reading them, I suggest you to just gather the main ideas behind the different models.

7.1 Local volatility

In local volatility models, the volatility of the underlying at time t is set to only depend on S_t and the time t . If the procedure of using local volatility for getting implied volatilities can be followed, the result is prices for each K and T with a perfect fit to the market prices. The procedure is such that the local volatility surface is computed from option prices and from this, model prices for options are found. Many of the models, simply extend the Black-Scholes by replacing the constant volatility by local volatility;

$$\frac{dS}{S} = (r - q)dt + \sigma(S_t, t)dW_t, \quad (33)$$

where W_t is a Wiener process. Using this model for the underlying, the local volatilities can be found for all call options $C(K, T)$ using Dupire's equation;

$$\sigma(S, T) = 2 \frac{\frac{\partial C(K, T)}{\partial T} + \mu K \frac{\partial C(K, T)}{\partial K} + r_f C(K, T)}{K^2 \frac{\partial^2 C(K, T)}{\partial K^2}} \quad (34)$$

This is the local volatility surface. Once this surface is found, model prices can be found from the Black-Scholes equation. Note that 34 requires knowledge of a complete set of options. Hence some interpolations have to be performed in order to find option prices.

Local volatility, at first sight might seem a bit restrictive and arbitrary but there is a connection to reality: It can be shown that the local variance $\sigma^2(S, t)$ is the expected value of the future real instantaneous variance $\sigma^2(t)$ at time t , conditional on the asset having value S at time t ;

$$\sigma^2(K, T) = \mathbb{E}[\sigma^2(T)|S(T) = K] \quad (35)$$

This is a constraint which must be satisfied for all models of S_t and local volatility models satisfies it in the simplest way; without any stochastic distribution around the expected value.

When fitting local volatility models, the common approach is to use a cross-section of option prices at the same maturity and create an implied tree. When constructing the trees, there are as many degrees of freedom in specifying the tree as there are option prices. This leads to a perfect fit to the market data but also a considerable risk of overfitting the data.

Local volatility models have been popular in vol surface construction since their introduction, even though local volatility was not intended for this in the first place. An advantage of using local volatility is the perfect fit, but there are serious caveats. The local volatility models require that the input market data is arbitrage-free. If it is not, negative translational probabilities can occur and mispricings follow [7]. It is well-known that local volatility models yields bad future predictions.

7.1.1 Examples

In [29], the author proposes a 3-step method in which he begins by smoothing the vanilla prices to obtain a price surface which is definitely free of arbitrage. Then he inverts the price surface using the Black-Scholes to get implied volatilities and then makes a few final adjustments on the vol surface to make sure it is globally arbitrage-free. The procedure does require that the vanilla prices are arbitrage-free to begin with. [7] gives a simple algorithm based on Kahalé's work of local volatilities which does not require the initial price data to be arbitrage-free.

In [30] an algorithm based on differential evolution and a local-diffusion model is used to fit vol surfaces. The method assures no arbitrage in the time or strike dimensions. The authors test their method on prices for put and call options on the Nikkei 225 index and the match to the data is very good.

In [31], the authors use a model based on local volatility calculated by using implied trees. They assume the existence of a complete, spanning set of European call option prices. One has to use interpolation and extrapolation to get such data and this is what the authors do. They assume geometric Brownian motion of the underlying, deterministic dividends and interest rates. They use a finite-difference scheme with a fast reliable calibration to vanilla prices. Then they use their method to price knock-out calls using the local volatilities calculated from the vanillas without checking the accuracy relative to market prices.

7.2 Stochastic volatility models

In stochastic volatility models, the volatility of the underlying is a stochastic variable with its own process, specified with a separate equation. These models introduce a new dynamic for the price of the underlying in contrast to local volatility models. They are popular in practice and in general they are good at fitting option prices with long maturities, but can yield worse fits for short maturities.

7.2.1 Examples

7.2.1.1 Heston

The Heston model is a stochastic volatility model which is of common use. It is in its basic form driven by the equations,

$$\begin{aligned} dS_t &= \mu S_t dt + \sqrt{\sigma_t^2} S_t dW_t \\ d\sigma_t^2 &= \kappa(\theta - \sigma_t^2)dt + \eta\sqrt{\sigma_t^2} dZ_t \\ dW_t dZ_t &= \rho dt. \end{aligned} \tag{36}$$

The first line describes the dynamics of the underlying. The stochastic behavior of the volatility σ_t is specified in the second equation. This process for the stochastic volatility is called a Cox-Ingersoll-Ross (CIR) process. W_t and Z_t are Wiener processes for the underlying and the stochastic volatility, having a correlation ρ . κ , θ and η are the other model parameters together with the initial volatility σ_0 . The Heston model models volatility as a mean-reverting process and introduces continuous shocks correlated between the underlying and volatility.

The Heston model can be, but is not necessarily arbitrage-free. It follows the sticky-strike rule and the smiles/skews decrease with increasing T as 1/T, [32].

The Heston model was compared to arguably more complex models involving stochastic time and jump processes in [33] on pricing of vanillas and exotics. It agreed well on the prices on vanillas, and all the models were differing on the exotics.

7.2.1.2 SABR

The SABR is a stochastic volatility model for a forward contract. It models the change in the forward price as the volatility multiplied by a power of the forward price weighted by increments

in a Wiener process. It is defined by the relatively simple equations,

$$\begin{aligned} dF_t &= \sigma_t F_t^\beta dW_t, \\ d\sigma_t &= \alpha \sigma_t dZ_t, \end{aligned} \tag{37}$$

where W_t and Z_t are correlated through $\rho \in (-1, 1)$ and $\beta \in [0, 1]$ and $\alpha \geq 0$ are the two other model parameters. SABR has been widely used in practice due to the existence of analytical approximate solutions to F_t , [34]. No exact solutions for the model's implied volatilities have been found so one has to resort to approximations. When these break down, which happens far in the money, serious mispricings and arbitrage arise from concave option prices and negative probability distributions, [35]. It is commonly known that SABR is not always arbitrage-free.

The SABR model can be extended by letting the parameters ρ , α and β be time-dependent. This complicates the analytical procedure and calibration, and introduces time-inhomogeneity but leads to better accuracy.

7.3 Discontinuous jump models

Brownian motions based on Wiener processes are involved in all models we have introduced so far. These have jumps which are infinitesimal increments that occur continuously, see figure 4. When zooming in on any short time span in such a process, the likelihood of a large jump becomes very small. Introducing finite jumps at discrete times in the price processes of financial assets such as from the process in figure 5 gives a finite probability of very large jumps in any given time span. Instead of continuous noise ultimately producing observable change, there are here periods of no change combined with single events of observable change. This is attractive, considering that the events in the markets happen irregularly and sometimes drastically and that trading is done discretely. A common distribution used for including discontinuous jumps is the Poisson distribution.

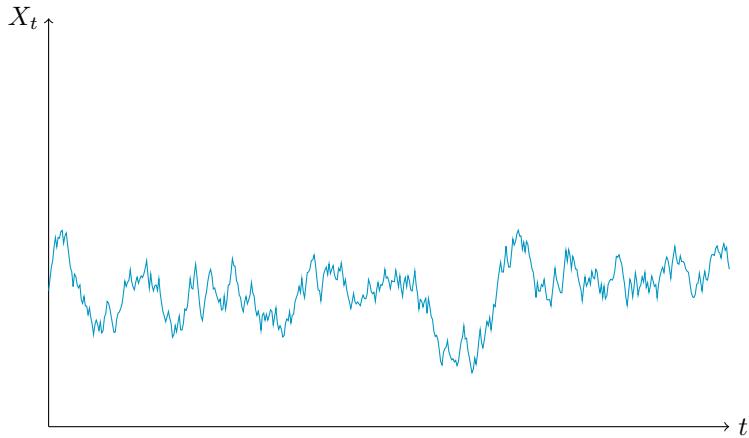


Figure 4: An example of a Brownian motion, X_t .

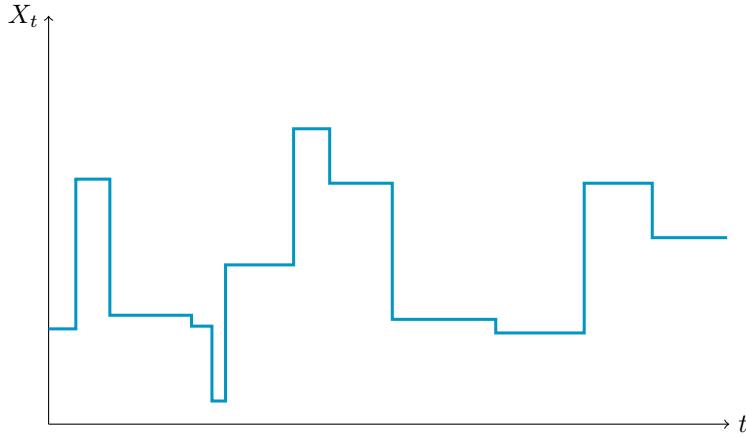


Figure 5: An example of a discontinuous jump process, X_t .

Stochastic volatility models often have trouble matching prices for out-of-the money options with short maturities, [36]. Discontinuous jump models are good at this their addition is needed for good accuracy, although pure discontinuous jump models can have difficulties with matching long maturities, see for example: [37], [38], [39], [40], [41], [42], [43], [44] and [45].

Discontinuous jump models with jump-sizes independent of S follow the sticky-delta rule and their skew decays as $1/T$ with increasing T in first order, [32].

7.3.1 Examples

7.3.1.1 Merton

The Merton jump-diffusion model involves a drift term, a diffusion term and a discontinuous jump term. One form of the process is given by the equation

$$\frac{dS_t}{S_t} = (r - \lambda m)dt + \sigma dW_t + (e^j - 1)dN_t, \quad (38)$$

where N is a Poisson process with some intensity λ , W is a Wiener process independent of N , j is a random logarithmic jump size, m is the expected value of $(e^j - 1)$, σ is a constant and r is the interest rate, [27].

7.3.1.2 CGMY

The CGMY model models the price of the underlying by

$$S_t = S_0 \exp [(\mu + \omega)t + X_{CGMY}(t, C, G, M, Y)] \quad (39)$$

where $X_{CGMY}(t, C, G, M, Y)$ is the CGMY process with parameters C, G, M and Y , μ is the mean rate of return and ω a is term obtained from the characteristic equation of the process, [46]. The CGMY-process is specified by its Lévy density:

$$k_{CGMY}(x) = \begin{cases} C|x|^{-(1+Y)} e^{-G|x|}, & \text{for } x < 0 \\ C|x|^{-(1+Y)} e^{-M|x|}, & \text{for } x > 0 \end{cases} \quad (40)$$

The process can capture both finite and infinite activity depending on the sign of Y . The parameters of the process further allow the jump component to have either finite or infinite variation. In [46], the model is extended by adding a diffusion component. The authors state that they find that this diffusion component is missing for market indices.

7.4 Combined models

The three types; local volatility, stochastic volatility and discontinuous jump processes can be combined to more complex and far-reaching models. One motivation for such combining is that stochastic volatility models tend to underestimate smile convexity at short maturities and jump-diffusion models as well as local volatility models tend to underestimate smile convexity at long maturities. When combined, they help with each others shortcomings. In, [44] the authors claim that stochastic volatility - jump diffusion models matches the market better than the use one of the three types alone.

A universal model is a term of a model where local volatility, stochastic volatility and discontinuous jumps are all combined. The large scope of these models increases the chances of a good fit but also of overfitting, which is when a model is too dependent on the training data to give accurate predictions outside this data. A model gains nothing by being overly complex and its inherent degrees of freedom are preferably not similar to the number of market data points to avoid this. In [47], the author notes that it is unfortunately very difficult to find explicitly solvable universal models. Combining jump diffusions and stochastic volatilities is straightforward but mixing stochastic and local volatilities requires a considerable effort.

7.4.1 Examples

7.4.1.1 The Barndorff-Nielsen-Shephard model

In [48] a model is described where the stochastic volatility is modeled with positive discontinuous jumps and exponential decay. The underlying is correlated with the jumps and otherwise follows a Brownian motion.

7.4.1.2 A model of Ayache, Henrotte, Nassar and Wang

The model named "body" presented in [27] involves 3 regimes of volatility with their own parameter values and 6 discontinuous jump processes between these, each with its own parameter for intensity and size. When only calibrating to the vanillas with 2 different parameter guesses, they obtain differences in barrier option prices but agreement on the vanillas.

7.4.1.3 A model of Gallucio and Cam

In [44], a model with discontinuous jumps and diffusion for the underlying and a correlated diffusion process and mean-reversion for the stochastic volatility is studied. They first calibrate the discontinuous jump part for short maturities. Then a regime-switch is introduced and the stochastic volatility-part is calibrated to long maturities. Time-inhomogeneity is introduced in the parameters. They indicate that their model gives an accurate and meaningfully fitted vol surface and that the algorithmic complexity is such that generalizing the approach to more complex models might be difficult to achieve.

7.4.1.4 The Heston model with discontinuous jumps

This model extends the regular Heston by simply adding a term with a Poisson process in the first row of equation 36. It was shown in [33] to give a good fit for one data set of vanillas.

7.4.1.5 JP Morgan

In [49], the authors aim to accurately fit their model to exotic options. They use a model with one stochastic volatility part which is the same as in the Heston model and one part of local volatility. It is matched to the vanilla prices and then used to calculate prices of different exotics. When the stochastic part matches the market data, the deterministic local part becomes redundant, otherwise, the local part makes the final tweak to match the market prices and indeed the model produces prices which are very close to the market prices for JPY-USD American binaries. They use a two-dimensional trinomial tree with one dimension being the underlying and the other the stochastic volatility part. The authors, who at that time were at JP Morgan, state that at the time of publication, their method was used by the JP Morgan foreign exchange options trading desk.

7.4.1.6 Blacher's model

Blacher's model, according to [27], is driven by the dynamics

$$\begin{aligned}\frac{dS_t}{S_t} &= rdt + \sigma_t [1 + \alpha(S_t - S_0) + \beta(S_t - S_0)^2] dW_t \\ d\sigma_t &= \kappa(\theta - \sigma_t)dt + \eta\sigma_t dZ_t\end{aligned}\tag{41}$$

where W_t and Z_t are Wiener processes correlated through ρ and $\alpha, \beta, \eta, \kappa$ and θ are the remaining model parameters. The model is somewhat local, in that the volatility explicitly depends on the size of S_t .

7.4.1.7 Lipton's model

According to [47], Lipton's model has the dynamics

$$\begin{aligned}\frac{dS_t}{S_t} &= (r^{01} - \lambda\nu)dt + \sqrt{v_t} \sigma_L(t, S_t) dW_t + (e^j - 1)dN_t, \\ dv_t &= \kappa(\theta - v_t)dt + \eta\sqrt{v_t} dZ_t\end{aligned}\tag{42}$$

Where N_t is a Poisson process with frequency λ , $j > 0$ is a random logarithmic jump size, r^{01} is the difference between interest rates in the accounting and underlying currencies and κ, θ, η are other model parameters. There is a local component in the volatility σ_L , stochastic volatility component in the $v(t)$ and a discontinuous jump process in N_t . The Lipton model is difficult to solve analytically, but it can be done.

7.5 Models with stochastic time

One way to introduce stochastic behaviour is to set up a stochastic process for the speed of passing of time. Then faster passing of time increases the volatility. The following two examples are models that are explained and evaluated in [33]. They are found to fit the vanillas well and do not vary extremely much on the tested exotics.

7.5.1 Examples

7.5.1.1 The NIG-OUT model

In this model, the underlying follows a normal inverse Gaussian process which is a pure jump process with 3 parameters. Time is set to follow a discontinuous jump process with exponential decay having 4 parameters.

7.5.1.2 The VG-CIR model

In this model, the underlying is modeled to follow a variance gamma process which also is a pure jump process with 3 parameters. Time is set to follow a CIR process - as the variance does in the Heston model, see equation 36.

7.6 Models with historical path-dependence

In [50], the author investigates a general path-dependent volatility model of the form

$$\frac{dS_t}{S_t} = \sigma(t, S_t, X_t) l(t, S_t) dW_t\tag{43}$$

where l is an arbitrary function of t and S_t . X_t is a finite set of path-dependent variables; for instance running averages or historical maxima and minima of the underlying. To use the model, one has to choose the particular path-dependent variables; to specify X_t . This type of model could have a lot of input data and it could be beneficial to include some weighting of the data's specific significance for vol surface construction at some stage. In my opinion, inclusion of historical data is a good addition since the future is often similar to the past.

7.7 Fitting higher order moments

In the Black-Scholes model, $\log(S_t)$ is assumed to be normally distributed and its first and second moments are extracted and used to specify the process of S_t . This is not good enough (hence this thesis has the given topic). The distributions implied by the markets have fatter tails and skews. A better fit to the market implied $\log(S_t)$ can be found by explicitly fitting higher order non-normal moments (such as skew and kurtosis) to market prices.

In [51], the authors suggest an extended version of the Black-Scholes model to account for biases induced by non-normal skewness and kurtosis in stock return distributions. Their pricing formula is given by the sum of the Black-Scholes option price plus adjustment terms for skewness and kurtosis. In particular, the probability density function of the underlying, which in the Black-Scholes model is only the normal distribution, $n(z)$, is in the model in [51] given by,

$$g(z) = z \left(1 + \frac{u_3}{3!}(z^3 - 3z) + \frac{u_4 - 3}{4!}(z^4 - 6z^2 + 3) \right), \quad (44)$$

where

$$z = \frac{\ln(S_t/S_0) - (r - \sigma^2/2)t}{\sigma\sqrt{t}}.$$

and u_i is moment i. A natural potential improvement of this model would be to incorporate higher order moments (μ_5, μ_6, \dots), but it is not quite that straight-forward. This is because problems of collinearity arise among additional pricing terms of odd moments and among even moments. Additionally, large data sets are needed to accurately estimate higher-order moments.

8 Direct models for vol surfaces

Instead of modelling the underlying and then investigating what vol surface this implies one can model the surface directly. This especially important given that option prices have been shown to not be completely determined by the price of the underlying asset. The methodology could be appropriate if one wants fast solutions, which fit the market data accurately enough, and better control of the overall shape of the surface. Drawbacks include that many models are not very flexible in their shapes, such that a model might not work for all underlyings. On top of this, creating a new model requires solid knowledge of the form of the vol surface for the underlyings of interest.

8.1 Examples

8.1.1 SVI

The stochastic volatility inspired (SVI) parameterization of the implied volatility smile was originally devised in 1999 and has been popular ever since. It is a practitioner-developed model where each smile/skew is fitted separately. In other words, it is calibrated separately for each T. The "raw" parametrization of the SVI, written in terms of $x = \ln(K/F)$ is defined as

$$\sigma^2(x) = a + b \left[\rho(x - m) + \sqrt{(x - m)^2 + s^2} \right], \quad (45)$$

where a, b, ρ , m and s are model parameters to be calibrated to market prices. The parameters each have straightforward effects on the smile. a affects the vertical alignment and m the horizontal alignment by setting the mid point of the smile. b affects the angle between the tails, ρ rotates the smile and s affects the curvature around m. There are other versions and extensions of the parametrization which are not presented here.

The original SVI model is known to not always be arbitrage-free, but additional conditions can be imposed on the parameters to exclude calendar and strike arbitrage, [52]. Doing this shrinks the set possible parameter values but not so much that the model have trouble fitting differing data sets.

A good property of the SVI model is that it automatically satisfies Lee's Moment formulas (equations 17 and 18) which set necessary bounds on the vol surface to be arbitrage-free in the high and low moneyness limits. Another being that it is quite easy to fit get good fits while avoiding calendar spread arbitrage, [53]. One downside is the amount of parameters needed - about 5 for each time to maturities. Another downside is that it is difficult to include data points with only a few strikes per maturity.

In [53], the authors introduce time-dependent parameters and present conditions for their SVI model to be free from static arbitrage. In [54] these results are extended further. In conclusion, the SVI model, especially versions of it which avoids all kinds of static arbitrage is a promising candidate for modeling vol surfaces directly.

8.1.2 RFSV

In [55], a model is proposed which is based on empirical findings on the distribution of the differences $\log(\sigma_{t+\Delta t}) - \log(\sigma_t)$. The model is built on what the authors call rough volatility - where no discontinuity is involved, but the process still has a weaker form of continuity than Brownian motion. This is achieved by including fractional Brownian motions as sources of stochasticity in the distribution for the differences mentioned above. The model for the volatility is simply stated as

$$\begin{aligned} \sigma_t &= \exp \{X_t\}, \\ dX_t &= \nu dW_t^H - \alpha(X_t - m) dt \end{aligned} \quad (46)$$

where W_t^H is a fractional Brownian motion with Hurst parameter H . H has empirically been found to be quite stable and around 0.13 for most indices and controls the decay in the ATM skew for very short maturities. In, [11] it is claimed that they cannot yet do a fast calibration of the model. This is not a disaster though, since the few parameters can be guessed to obtain good fits. By simply guessing the 3 parameters in the "sub-model rBergomi", they get a vol surface which is almost inseparable from the discrete SPX option data from August 14, 2013 which they base their calculations on, [56].

8.1.3 Historical vol surface decomposition

In [26], the authors introduce a model where the vol surface is represented by orthogonal basis functions, which are Ornstein-Uhlenbeck processes with noise terms made of Wiener or jump processes. In the model, the principal directions of deformation of historical forms of a vol surface can be found by principal component analysis and then be used as a basis for predicting the future shape of such a vol surface. The basis functions for the principal directions of deformation has to be specified by the modeller beforehand.

In order to decompose historical vol surfaces, such have to be found or constructed beforehand. This makes the method less of a vol surface construction model and more of a complementary model for predicting the future.

8.1.4 A combined vol surface and S_t model

In [57] an approach is investigated where the vol surface and the underlying are modelled in correlation;

$$\begin{aligned} dS_t &= S_t \sqrt{v_t} dW_t \\ d\sigma_t(K, T) &= \mu_t dt + \omega_t dZ_t \end{aligned} \tag{47}$$

where W_t and Z_t are Wiener processes with correlation ρ . v_t sets the variance of S_t and $\sigma_t(K, T)$ is the vol surface at time t .

The model has a framework which integrates how investors manage their option positions. They find that given their assumptions, the entire shape of the vol surface can be obtained by solving a simple quadratic equation, thus dramatically decreasing the computational time relative to many other models. The model allows the surface to be based on many state variables without having any model parameters. This allows for a continuous recalibration of the model as the time series of the state variables evolve, [57]. The authors Carr and Wu analyze the S&P500 index and its options within the model and calculate realized and expected vols. The model of the implied vols explain on average 98% of the variation in the data.

In [58], a theoretical investigation is made to check whether two specific versions of the model are consistent with no-arbitrage. The author assumes an interest rate of zero and finds that the two versions are not consistent with absence of arbitrage. He finds that negative vertical spreads can arise for strikes not very far from the money in these cases. Further, for a quite general case of the model where,

$$d\sigma_t(K, T) = \mu_t \sigma_t^p(K, T) dt + \omega_t \sigma_t^q(K, T) dW_t, \tag{48}$$

where $p \in [0, 1]$, $q \in [1/2, 1]$ and μ_t and ω_t do not depend on K , T or $\sigma(K, T)$, if a risk-neutral condition is fulfilled, this model cannot fulfill a necessary condition for absence of static arbitrage.

9 My implementation

9.1 Model selection

Now a smattering of models have been introduced briefly. Let's have a break from introducing new things and discuss what we have found. In this master's thesis, apart from getting an overview of the problem, I also want to implement and test models. It is then time to decide upon what model/models I want to investigate. The premise is that I want to find a way to create a vol surface with appropriate properties from option prices. Which properties? Well, I set up that I ideally wanted a model to fulfill the following criteria:

- Ability to fit market vanillas well.
- Ability to incorporate and accurately price exotics.
- Absence of arbitrage.
- Ease of implementation.

The last point is included based on the time restrictions and the need for at least one correctly implemented model. This is of course highly dependent on the complexity of implementation. Some models, I found, would be good but too time-consuming to study in further detail and to implement, relative to other choices. Let's now evaluate the methods according to my standing point and the criteria.

I was instantly attracted to modeling surfaces directly. This choice is motivated by the results in section 3.1 indicating that modelling the underlying alone is not accurate enough for constructing vol surfaces. The new RFSV model is promising but I don't like that it cannot yet be calibrated quickly such that instead, the parameters have to be guessed. Another choice is the SVI. This is a popular model and could be a good choice because of its simplicity. There might be arbitrage though, unless time-dependent parameters are introduced.

The Monte Carlo approach is very flexible. Here I could set up my own model with any properties that I want; including bid-ask spread data, traded volumes, trading costs and prices of exotics for better accuracy. I think these properties could be incorporated in the other approaches too, as added terms in the pricing equations, for example on the right-hand-side of equation 30. The fact that such inclusions can be made in other models as well combined with that Monte Carlo methods are highly based on statistical methods lead me to not implement this approach. The randomness of Monte Carlo likely introduces more uncertainty than in other methods, adding another level on top of the numerical approximations made in all models. This likely introduces minor arbitrage opportunities and a rugged surface.

I was also attracted to the Fourier transform approach. In it, once the framework is set up, many different S_t -models can be evaluated simply by interchanging the characteristic equations. The calculations in this method could be done quickly and with good accuracy while the surfaces produced should have neat shapes. This combined with the expected speed and ease of implementation made me go with this choice of approach. Direct surface modeling has to wait until next time. Next, I wanted to decide upon which S_t -models to use.

I wanted to have models without local volatility and inhomogeneity in time. Stochastic volatility and jumps have been used quite successfully and seem to be good choices. I decided to use both the Heston model and the Heston model with a discontinuous jump term added to the evolution of S_t - a model I call Heston Jump. In this way, I can investigate the claimed improvements of introducing discontinuous jumps on top of stochastic volatility. I found the Bernhard-Nielsen-Shephard model attractive based on that it only has 5 parameters but still captures stochastic volatility, discontinuous jumps and a drift for S_t . I chose to include this one too. Finally, I wanted to have a model with a larger scope, if possible. In [59], 25 models are studied. The best-performing is called CGMYSA and models S_t as a CGMY-process with time varying stochastically. This model can capture finite and infinite activity as well as finite and infinite variation (not simultaneously though) in S_t using 9 parameters. The model is hence quite general and somewhat more mathematically complex.

I could not get my hands on any reliable exotic option data. Hence I had to exclude calibration to and pricing of them in my implementation. I aimed to model the surface of Nikkei 225 and found daily closing prices and traded volumes for each option during a day. I decided to calibrate the models to Nikkei option data from one (random) day. I would have wanted to include a linear term in bid-ask spread in percentage of option price in equation 30, but could not gather bid-ask spread data. The quality of the data is not very good. Nevertheless, I wanted to test the models on some low-quality data and see whether they can perform adequately. Though, as a start, I decided to check the models on the nice looking grid of implied vols found in [33] and presented in Appendix A, with the hope to get the models working and a happy spirit before tackling the Nikkei data.

On top of this I decided to implement a check for arbitrage; both for checking the quality of the input data and for sanity testing of constructed surfaces. This check can be found in section 9.4.

Now let's have a closer look at the (fast) Fourier transform approach I have implemented.

9.2 The FFT approach

The idea is to use one of the two methods introduced in [25] to calculate option prices. The first method was briefly described in section 5.3. It states that European call prices can be found through,

$$C_T(k) = \frac{\exp(-\alpha k)}{\pi} \int_0^\infty e^{-ivk} \psi_T(v) dv \quad (49)$$

where the model specific characteristic equation $\phi_T(u)$ of $\log(S_t)$ is incorporated through,

$$\psi_T(v) = \frac{e^{-rT} \phi_T(v - (\alpha + 1)i)}{\alpha^2 + \alpha - v^2 + i(2\alpha + 1)v} \quad (50)$$

Here k is the log-Strike; $k = \log(K)$. Equation 49 can be solved for all strikes simultaneously using the fast Fourier transform (FFT) algorithm, but to do so, it needs to be discretized and put on the Fourier transform form

$$X_t = \sum_{n=0}^{N-1} x_n e^{-i2\pi kn/N}$$

I choose to discretize k and v as

$$\begin{aligned} k_u &= -b + \lambda(u - 1), \\ v_j &= \eta(j - 1), \\ \eta\lambda &= \frac{2\pi}{N}, \end{aligned} \quad (51)$$

where $u = 1, 2, 3, \dots, N$ and $j = 1, 2, 3, \dots, N$. The constant b can be tweaked to focus on a desirable range of strikes. λ and η sets the spacing in $\log(K)$ - and integration variable-space, respectively. The last line in equation 51 is needed to get the right relationship between the "signal variable" v and the "frequency variable" k in the FFT. Using the trapezoid rule, equation 49 can now be written in a discretized form,

$$C_T(k_u) = \frac{\exp(-\alpha k_u)}{\pi} \sum_{j=1}^N e^{-i\frac{2\pi}{N}(u-1)(j-1)} e^{ibv_j} \psi_T(v_j) [1 - \delta]\eta \quad (52)$$

where $\delta = 1/2$ for $j = 1$ and $j = N$, and otherwise 0. This is a direct application of the FFT algorithm for finding prices at many strikes for a chosen maturity. When the $\phi_T(\cdot)$ for a model is known, equation 52, together with equation 50 can be used to calculate vanilla call prices in a discrete but arbitrarily dense grid (K,T). The output from the FFT is equidistant prices for $\log(K)$ - not for K. To find prices for a specific (K,T) using the FFT then, some interpolation is usually needed between nearby prices.

The integral in equation 49 can also be solved directly - without using an FFT. In this way one avoids the step of interpolating prices when calculating the price for a chosen (K,T). Which method is faster depends on the size of the data and the interpolation method used after the FFT. If N is the number of points needed to evaluate the integral accurately, the FFT has a complexity of $N \log_2(N)$. Direct integration has a complexity of ND , where D is the number of market data options. For good accuracy of the FFT and the subsequent interpolation step, at least 1000 points are needed for evaluating the integral for most (K,T) in the market data. $\log_2(1000) \approx 10$, so the FFT might be faster for data sets with (significantly) more prices per maturity than 10. But, interpolation has to be done after the FFT algorithm has been run, which takes some time. Since the speeds of computation are comparable but evaluating the integral directly leads to more accurate model price estimates, I choose to do the calibration using the direct integral evaluation.

When plotting a surface from model parameters, prices are needed for a dense (K,T) space. For this purpose the FFT is faster than direct integration so I use this method for that purpose. In the step of creating a surface from model parameters, I observed that it was difficult to calculate accurate prices for the combination of high strikes and low maturities. Sometimes the prices would even turn out to be negative - which is not unimaginable given that the true prices for such (K,T) could be around (say) 10^{-5} (Yen) which are to be pin-pointed by an oscillating integral running towards infinity. The second method in [25] is better suited for pricing out of the money strikes so I used this instead when creating a surface from parameters. In this method, instead of multiplying with an exponential, a hyperbolic function is used. It is defined by the similar-looking pricing equations

$$C_T(k) = \frac{1}{\pi \sinh(\alpha k)} \int_0^\infty e^{-ivk} \gamma_T(v) dv, \quad (53)$$

where

$$\begin{aligned} \gamma_T(v) &= \frac{\zeta_T(v - i\alpha) - \zeta_T(v + i\alpha)}{2} \\ \zeta_T(w) &= e^{-rT} \left[\frac{1}{1 + iw} - \frac{e^{rT}}{iw} - \frac{\phi_T(w - i)}{w^2 - iw} \right]. \end{aligned} \quad (54)$$

Equation 53 could then be discretized similarly using the trapezoid rule and equation 51 to yield an equation where the FFT is directly applied:

$$C_T(k_u) = \frac{1}{\pi \sinh(\alpha k)} \sum_{j=1}^N e^{-i\frac{2\pi}{N}(u-1)(j-1)} e^{ibv_j} \gamma_T(v_j) [1 - \delta] \eta \quad (55)$$

That is how I have chosen to go from S_t -models to option prices. Now, let's look at the S_t -models I have implemented in more depth. We start with the historically popular Heston model and its extension.

9.3 Chosen models

9.3.1 Heston and Heston jump

The Heston jump model is an extended version of Heston. It is exactly as Heston apart from an additional term in the process of the underlying which also affects the drift. The governing relations are here

$$\begin{aligned} \frac{dS_t}{S_t} &= (r - q - \lambda \mu_J) dt + \sigma_t dW_t + J_t dN_t \\ d\sigma_t^2 &= \kappa(\theta - \sigma_t^2) dt + \eta \sqrt{\sigma_t^2} dZ_t \\ dW_t dZ_t &= \rho dt. \end{aligned} \quad (56)$$

$N = \{N_t, t \geq 0\}$ is an independent Poisson process with intensity parameter λ . Given that a jump occurs, the jump size in percent is driven by the process J_t :

$$\log(1 + J_t) \sim n \left(\log(1 + \mu_J) - \sigma_J^2 / 2, \sigma_J^2 \right) \quad (57)$$

where n is the normal distribution such that μ_J is the mean jump in percent and σ_J the standard deviation of jumps. The characteristic equation for the Heston jump model then involves the 7 parameters $\{\rho, \theta, \eta, \kappa, \sigma_0, \lambda, \mu_J, \sigma_J\}$, where σ_0 is the current volatility of the underlying. In terms of $\log(S_t)$ it is specified by ([33]),

$$\begin{aligned}\phi_t(u) = & \exp(iu(\log(S_0) + (r - q)t)) \\ & \times \exp(\eta\kappa\theta^{-2}((\kappa - \rho\theta ui - d)t - 2\log((1 - ge^{-dt})/(1 - g)))) \\ & \times \exp(\sigma_0^2\theta^{-2}(\kappa - \rho\theta ui - d)(1 - e^{-dt})/(1 - ge^{-dt})) \\ & \times \exp(-\lambda\mu_J iut + \lambda t((1 + \mu_J)^{iu} \exp(\sigma_J^2(iu - 1)) - 1)),\end{aligned}\quad (58)$$

where

$$\begin{aligned}d &= \sqrt{(\rho\theta ui - \kappa)^2 - \theta^2(-iu - u^2)}, \\ g &= (\kappa - \rho\theta ui - d)/(\kappa - \rho\theta ui + d).\end{aligned}\quad (59)$$

When searching for optimal fits in parameter space it is sometimes needed to define parameter bounds. I set the lower parameter bounds for the Heston Jump model to $\{-1, 0, 0, 0, 0, 0, 0, 0\}$ and the upper bounds to $\{1, 4, 2, 1, 1, 5, 1, 1\}$ with the initial guess $\{-0.9, 0.1, 0.1, 0.1, 0.1, 1, 0.1, 0.1\}$. The upper bounds are chosen somewhat arbitrarily: Hopefully not too restrictive while still not allowing a too large search space. Note that setting a lower bound of $\mu_J = 0$, favors jumps in the positive direction.

The characteristic equation for the pure Heston model is simply given by equation 58 without the fourth exp-term with the jump parameters. The pure Heston model's parameters are $\{\rho, \theta, \eta, \kappa, \sigma_0\}$. I set the lower parameter bounds to $\{-1, 0, 0, 0, 0\}$ and the upper bounds to $\{1, 4, 2, 1, 1\}$ with the initial guess $\{-0.9, 0.1, 0.1, 0.1, 0.1\}$.

9.3.2 Barndorff-Nielsen-Shephard

In the BNS model, volatility is modeled by an Ornstein-Uhlenbeck process driven by a subordinator for which I will choose the Gamma-OU process. In the model, volatility can only jump upwards and then it decays exponentially until the next jump. The underlying asset has a co-movement with the volatility through the parameter ρ . Between jumps, the price of the underlying will diffuse. The equations describing the model are, ([33]):

$$\begin{aligned}d\log(S_t) &= (r - q - \lambda k(-\rho) - \sigma_t^2) dt + \sigma_t dW_t + \rho dz_{\lambda t} \\ d\sigma_t^2 &= -\lambda \sigma_t^2 dt + dz_{\lambda t}\end{aligned}\quad (60)$$

where $\lambda > 0$, $k(u) = \log(\mathbb{E}[\exp(-uz_1)])$ and z_t is a subordinator independent of the Wiener process W_t . The Gamma-Ornstein-Uhlenbeck process is expressed by

$$z_t = \sum_{n=1}^{N_t} x_n \quad (61)$$

where $N = \{N_t, t \geq 0\}$ is a Poisson process. N_t has an intensity parameter a , such that $\mathbb{E}[N_t] = at$. $x = \{x_n, n = 1, 2, 3, \dots\}$ is an independent and identically distributed sequence. All x_n follow an exponential law with mean $1/b$. This leaves σ_t as a stationary process with a marginal law that follows a Gamma distribution with mean a and variance a/b .

The model's implied characteristic equation for $\log(S_t)$ then has 5 parameters; $\{\rho, a, b, \lambda, \sigma_0\}$ and is specified by, [33],

$$\begin{aligned}\phi(u, t) = & \exp(iu(\log(S_0) + (r - q - a\lambda\rho(b - \rho)^{-1})t)) \\ & \times \exp(-\lambda^{-1}(u^2 + iu)(1 - \exp(-\lambda t))\sigma_0^2/2) \\ & \times \exp\left(\frac{a}{b - f_2(u)} \left(b \log\left(\frac{b - f_1(u)}{b - iu\rho}\right) + f_2\lambda t\right)\right)\end{aligned}\quad (62)$$

with,

$$\begin{aligned} f_1(u) &= iu\rho - \lambda^{-1}(u^2 + iu)(1 - \exp(-\lambda t))/2 \\ f_2(u) &= iu\rho - \lambda^{-1}(u^2 + iu)/2 \end{aligned} \quad (63)$$

For the BNS model, I set the lower parameter bounds to $\{-10, 0, 0, 0, 0\}$, the upper bounds to $\{10, 30, 30, 10, 1\}$ and the initial guess to $\{-5, 1, 10, 2, 0.2\}$.

9.3.3 CGMYSA

The CGMYSA model is based on the CGMY process. This process can capture all combinations of finite or infinite variation & activity, thus being more general than the Variance Gamma, Normal Inverse Gaussian or a compound Poisson process, each of these adhering to one of the choices alone, [59]. The process' characteristic function is, [59],

$$\begin{aligned} \phi_{CGMY}(u; C_p, G, M, Y_p, Y_n, \zeta) &= C_p [\Gamma(-Y_p) ((M - iu)^{Y_p} - M^{Y_p}) \\ &\quad + \zeta \Gamma(-Y_n) ((G + iu)^{Y_n} - G^{Y_n})] \end{aligned} \quad (64)$$

In models with stochastic time, often a stochastic clock in the form of an increasing yet stochastic process made from Lévy processes is introduced as a subordinator to another Lévy process which models S_t . The link between the stochastic clock and the subordinated Lévy process can be found through their characteristic equations. Let's denote the stochastic clock's elapsed time by $Y(t)$ and set it on the form

$$Y(t) = \int_0^t y(s)ds \quad (65)$$

where $y(s)$ is a stochastic process. Denoting by $X(t)$ the subordinated process, the resulting S_t -modelling process with the subordinator $Y(t)$ is $Z = X(Y(t))$ with a characteristic function of,

$$\mathbb{E}[\exp(iu(Z(t)))] = \mathbb{E}[\exp(iu(X(Y(t)))] = \mathbb{E}[\exp(Y(t)\phi_X(u))] = \phi_Y(-i\phi_X(u), t, y(0)). \quad (66)$$

Let's say we model S_t as a Lévy process subordinated by a stochastic clock. Then S_t can be chosen to move according to the mean-corrected expression ([33], [59]),

$$S_t = S_0 \frac{\exp((r - q)t)}{\mathbb{E}[\exp(X(Y(t))|y_0)]} \exp(X(Y(t))) \quad (67)$$

Such that the characteristic equation of $\log(S_t)$ is

$$\phi_{\log(S_t)}(u, t) = \exp(iu(\log(S_0) + (r - q)t)) \frac{\phi_Y(-i\phi_X(u), t, y_0)}{[\phi_Y(-i\phi_X(-i), t, y_0)]^{iu}} \quad (68)$$

In CGMYSA, S_t is modeled to follow the Lévy process CGMY subject to a subordinating CIR stochastic clock. You could say that we pick $Y = CIR$ and $X = CGMY$. S_t is chosen to be mean-corrected as in equation 67. The CIR process is the same as the process for the variance in the Heston model (equation 36). CIR's characteristic equation is given by ([33], [59])

$$\begin{aligned} \phi_{CIR}(u, t; \kappa, \eta, \lambda, y_0) &= \frac{\exp(\kappa^2 \eta t / \lambda^2) \exp(2y_0 iu / (\kappa + \gamma \coth(\gamma t/2)))}{(\cosh(\gamma t/2) + \kappa \sinh(\gamma t/2) / \gamma)^{2\kappa\eta/\lambda^2}} \\ \gamma &= \sqrt{\kappa^2 - 2\lambda^2 iu}. \end{aligned} \quad (69)$$

The complete characteristic equation of $\log(S_t)$ in the CGMYSA model is then given by equations 64 and 69 inserted in equation 68 in this way:

$$\phi_{CGMYSA}(u, t) = \exp(iu(\log(S_0)) + (r - q)t) \times \frac{\phi_{CIR}(-i\phi_{CGMY}(u; 1, G, M, Y_p, Y_n, \zeta), t; \kappa, \eta, \lambda, C)}{[\phi_{CIR}(-i\phi_{CGMY}(-i; 1, G, M, Y_p, Y_n, \zeta), t; \kappa, \eta, \lambda, C)]^{iu}} \quad (70)$$

The parameters in the CGMYSA model are $\{C, G, M, Y_p, Y_n, \zeta, \kappa, \eta, \lambda\}$. In [46] the authors give an account of the meaning of the parameters: C is setting the overall activity. G and M control the exponential decay of the right and left sides of the mean of the Lévy density. $G < M$ leads to a heavier left tail than the right tail in the distribution of the CGMY process and vice versa. The

parameters should satisfy $C > 0$, $G \geq 0$, $M \geq 0$ and $Y < 2$.

The gamma function blows up (and down) at $\Gamma(n)$ for $n = 0, -1, -2, \dots$. This means ϕ_{CGMY} blows up (and down) for $Y_n = 0, 1, 2$ and $Y_p = 0, 1, 2$. Hence one must be careful around these regions in the search for optima in parameter space.

Motivated by some numerical results in [59], I tried the following parameter bounds. Lower bounds: $\{0, 0, 0, -2, -2, 0, 0, 0, 0\}$, upper bounds: $\{100, 400, 400, 0.99, 0.99, 2, 30, 40, 60\}$ and the initial guess: $\{10, 20, 40, 0.4, 0.4, 0.6, 2.0, 5.0, 8.0\}$. The upper bounds and guesses here are a bit arbitrary. I did not want to construct a too large search space to get lost there, but of course not too limiting such that good optima would be excluded.

9.4 Arbitrage

9.4.1 Choosing arbitrage check

Absence of arbitrage has been a hot topic in this thesis so far. I wanted to implement a check for testing whether a surface is arbitrage-free and hence self-consistent and realistic. I realized that it would also be of interest to check the input data for arbitrage as well and that this could be done with the same check.

The intention is that the test should be run on the market data before the calibration to see whether there are any bad data points which are better left out. For, instance perhaps a large butterfly spread arbitrage opportunity exists in the data, but by removing one point, the data is free of static arbitrage. After the calibration, the test can be run on the vol surface to see for what (K, T) , the surface fails to perform and get indications for where it is self-consistent.

I aimed to implement simple checks for which the underlying assumptions could at least be partially fulfilled. For market data, finding derivatives in terms of K or T by approximations would not lead to very accurate results, so conditions involving derivatives are difficult to include. In addition, conditions of continuity cannot be checked numerically. Because of this I chose not to implement the otherwise promising check on the vol surface itself in section 4.3.2.1. The check in section 4.3.1.4 might be good but I think any potential excellent results would be outweighed by the amount of work required to get it right, in this case. The matrix check in section 4.3.1.3 requires an input of equidistant strikes which is sometimes not found for the market data. This check could be used for a constructed surface though if some interpolation is performed after the FFT, but I think such a check is similar to what I decided to implement. Instead of using these, I decided to implement the check of whether spreads have positive prices in section 4.3.1.1 and the related check in section 4.3.1.2. This is likely a decent check for presence of *static* arbitrage.

9.4.2 My implemented arbitrage check

I chose to implement the checks of positive prices for calendar, vertical and butterfly spreads and the conditions (A1)-(A5). There is though some overlap between these two. (A2) is taken care of by positivity of calendar spreads. (A3) and (A5) deals with limits of strikes and maturities and are not relevant for market data. (A4) on the other hand sets boundaries for the option prices and is not covered by the spreads. The details of the implemented check are explained in the box below.

Arbitrage check

A grid for the market price data in (K,T) is set up and the strikes are denoted $K_1, K_2, K_3, \dots, K_{kmax}$ and the maturities $T_1, T_2, T_3, \dots, T_{tmax}$, both in ascending order. Next the option prices are looped through using $i = 1, 2, 3, \dots, kmax$ and $j = 1, 2, 3, \dots, tmax$ and the following tests are made at each point:

1. Positivity of the calendar spread: $C(K_i, T_{j+1}) - C(K_i, T_j) > 0$.
2. Positivity of the vertical spread: $C(K_i, T_j) - C(K_{i+1}, T_j) > 0$.
3. Positivity of the butterfly spread: $C(K_i, T_j) - \frac{K_{i+2} - K_i}{K_{i+1} - K_i} C(K_{i+1}, T_j) + C(K_{i+2}, T_j) > 0$
4. Bounds for the price, (A4): $\max(S_0 - K_i, 0) \leq C(K_i, T_j) \leq S_0$

If any of the conditions are not fulfilled for a data point, the lowest price offset in percent could be printed out together with the current (K,T) and a boolean for arbitrage could be returned as false.

10 Calibration procedure

The task in the calibration step is to find the optimal parameters of the models I have chosen for a specific data set by searching the parameter space. It is an optimization problem. The optimal parameters are the ones which produce the best model to market pricing match, measured by some function called the objective function.

As can be seen in equation 49 and its inputs, equation 50 and for instance equations 58 and 59, the optimization problem is highly non-linear and finding analytical gradients of the objective function is likely too big of a task. In this section I want to introduce the basic ideas of three direct black-box optimization algorithms which can perform the calibration step without the need of gradients. I tried them all by using their java implementations, found in a package for multivariable non-linear optimization: org.apache.commons.math3.optim.scalar.noderiv.

Before applying the optimization algorithms, I mapped all the lower and upper parameter bounds to 0 and 10, respectively. This was done to help the algorithms by introducing an equal weighing between the parameters, cancelling out the different lengths of ranges between the lower and upper bounds. If this is not done, it is difficult to calibrate parameters with tight bounds.

To not get stuck with one bad local minimum as a solution, I wrapped each optimization algorithm in a "MultiStartMultivariateOptimizer" which runs the optimization algorithm several times, each with different initial guesses. I chose the second to fifth initial guesses randomly and uniformly in the parameter space.

10.1 The Nelder-Mead algorithm

The Nelder-Mead algorithm is a deterministic algorithm attempting to find a minimum of an objective function. It does this by using a simplex of $n + 1$ dimensions, which can be thought of as a high-dimensional polygon, located in the parameter space. For each iteration, the worst vertex in the simplex is removed and replaced by a better point. In figure 6, a small simplex with 3 points; p_1 , p_2 and p_3 together with possible choices in an iteration step is depicted. In this iteration, the worst point p_1 is reflected in the centroid of the remaining points giving the point p_r . The objective function is evaluated there and if the value is good enough (but not too good) it replaces the worst point and the next iteration starts. If p_r is the current best, an expansion is made to point p_e and the better of the two replaces the worst point. If p_r is the second worst or the worst point, contractions are made to points p_c or p_{cc} , respectively. If these steps have failed to find a better point, the simplex shrinks around the best point, which in this case is set to be p_2 . A more detailed description of the algorithm can be found in [60].

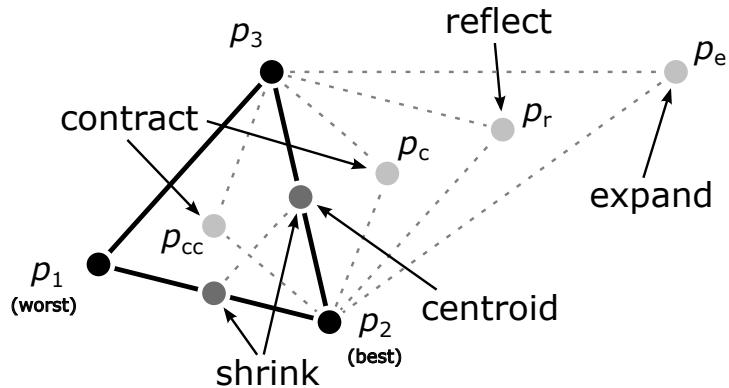


Figure 6: An illustration of one iteration of the Nelder-Mead algorithm where the worst point p_1 is to be replaced by some of the light gray points on the line connecting p_1 and the centroid, or the points p_1 and p_3 will be replaced by the dark gray points.

The Nelder-Mead algorithm does not directly support bounded constraints. This can however be circumvented by wrapping a penalty function around the objective function which penalizes parameter choices outside the bounds. This is how I did it in my calibration. I used the "MultiVariateFunctionPenaltyAdapter" in org.apache.commons.math3.optim.nonlinear.scalar.

The inputs to the algorithm are an initial parameter set guess, a simplex and relative and absolute stopping tolerances. I used $\max(10^{-9}, \sqrt{N} \times 10^{-8})$ where N is the number of model parameters as both the relative and absolute tolerances.

10.2 The BOBYQA algorithm

The name BOBYQA stands for bounded optimization by quadratic approximations. In each iteration a quadratic approximation Q to the objective function is created, which exactly fits the objective function in m points. $m = 2N+1$ is a typical value, where N is the number of parameters, [61]. Q is updated in each iteration by minimizing the Frobenius norm of the change in the second derivative matrix of Q . A number called the trust region radius is calculated in each iteration which indicates where the quadratic approximation can be trusted. Then a step in the parameter space is taken whose magnitude is smaller than the trust region radius.

The algorithm is suitable for optimization problems with intricate search spaces, where derivatives are not available. It can take simple bounded constraints on the parameters x_i of the form

$$a_i \leq x_i \leq b_i, \quad i = 1, 2, 3, \dots, n$$

as inputs. The algorithm takes only two input parameters: the initial and stopping trust region radius. The stopping trust region radius sets the minimum trust region radius allowed. In numerical experiments, the distance from the final parameter set and a local minimum is less than 10 times the stopping trust region radius, unless the radius is so small that this is impossible to achieve, [61].

In my use of the algorithm, I set the initial trust region radius to the square-root of the number of parameters. Remember, all parameters are transformed to within [0,10]. I used $\max(10^{-9}, \text{initial radius} \times 10^{-8})$ as a stopping trust region radius.

10.3 The CMA-ES algorithm

CMA-ES stands for covariance matrix adaptation evolution strategy. It is a randomized direct search method and particularly designed to tackle ill-conditioned-problems with rugged-search landscapes [62]. The algorithm models a species living in parameter space which evolves according to an evolution strategy. See figure 7 for an illustration of a few iterations of the algorithm. In each iteration, a new generation of offsprings (points) is generated from the parental generation (of points), which dies. The offsprings are generated by recombinations of and mutations in the parents. Only points with high fitness (low objective function values) are allowed to produce offsprings. The distribution determining the location of the next generation g is

$$\mathbf{x}_k^{(g+1)} \sim \mathbf{m}^{(g)} + \sigma^{(g)} N(\mathbf{0}, \mathbf{C}^{(g)}), \quad \text{for } k = 1, 2, 3, \dots, \lambda.$$

\mathbf{C} is a covariance matrix which contains information of pairwise dependencies between the model parameters based on an approximation of the objective function. \mathbf{m} is the current mean location of the population, σ is the current standard deviation and λ is the population size. All of these are updated in each iteration after they have been used to create a new population.

The inputs to the algorithm are an initial guess and initial standard deviations for all model parameters. Non-default values for the stopping criteria "stopping fitness" and a population size can be set. Particularly in this algorithm it is of interest to map all parameter bounds into the same interval to not focus on parameters with bigger ranges of allowed values.

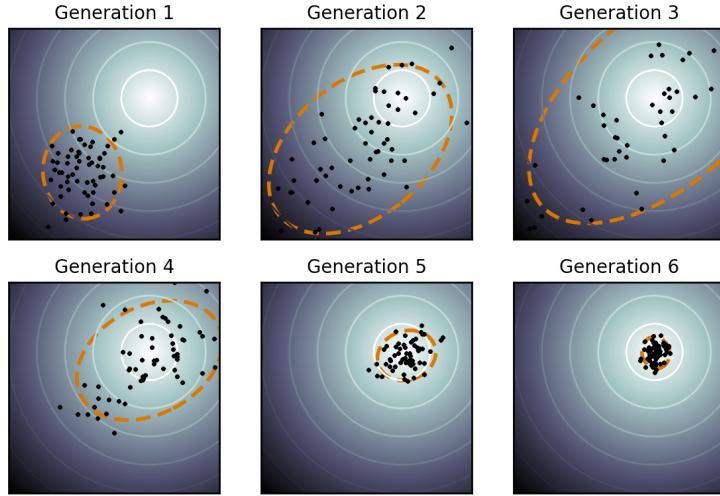


Figure 7: A few iterations of the CMA-ES algorithm in two dimensions using a relatively large population size. The species quickly center around the bright minimum of the objective function.

10.4 Objective functions

Choosing a suitable objective function is important both for creating a not too rugged search space and for being representative of the goodness of fit for the model. I tried the common average relative percentage errors (arpe),

$$\begin{aligned} \text{arpe for options} &= \frac{1}{\text{number of options}} \sum_{\text{Options}} \frac{|\text{market price-model price}|}{\text{market price}} \\ \text{arpe for implied volatilities} &= \frac{1}{\text{number of options}} \sum_{\text{Options}} \frac{|\text{market vol-model vol}|}{\text{market vol}} \end{aligned} \quad (71)$$

Calibration to implied volatilities at first seems like the superior choice since the goal is to construct a good vol surface, but does this prolong the calibration time significantly? The computational complexity of one objective function evaluation in my code, where N is the number of options and v the number of integration points needed in equation 49 is about $3Nv$ for calibrating to option prices and about $3Nv+5N$ for calibrating to implied volatilities. The extra $5N$ comes from the last step; inverting the Black-Scholes formula (equation 1) for implied volatilities for every option. Since $v \gg 5$, there is no significant time loss for calibrating to implied volatilities. However, the search-landscape becomes more rugged and there is a trade-off between this and the supposedly better fit found, if the algorithms succeed. See figure 8 for an example of how a very rugged search-landscape might look. I evaluated both methods and the results can be found in the results section.

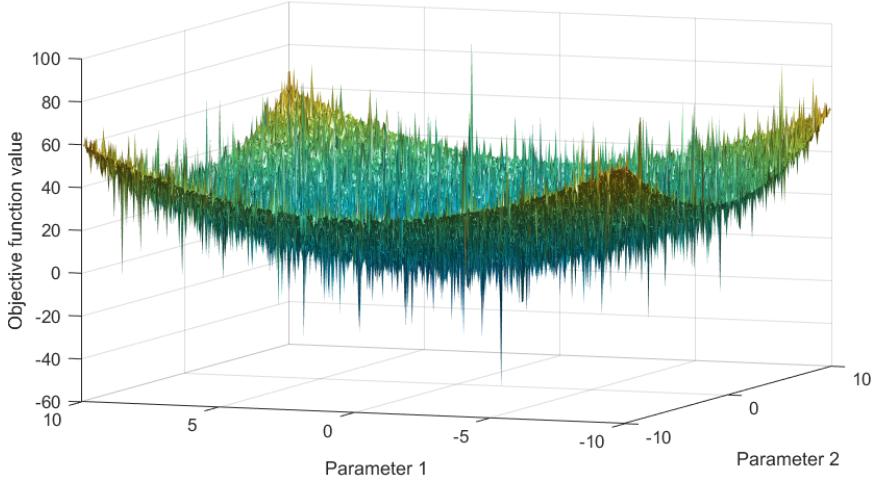


Figure 8: Perhaps calibrating to implied volatilities rugges the search space too much.

10.5 Calibration validation and pitfalls

In this section I want to bring into light some further choices I made in the implementation and some pitfalls one can fall into when implementing the FFT approach.

I found that using rectangular integration in comparison to Simpson's or the trapezoidal rule when doing the direct numerical integration of equation 49 lead to very poor convergence. This could be due to the integrand peaking around 0 and then falling sharply. When using either of the other two integration methods, the step size in the integration variable could be orders of magnitude higher which lead to a significant speed up in the code. This difference was so big that I expected there to be some bug in the code.

When integrating, I used a fixed step size and limits of integration and I suspect that using sophisticated methods which adjusts the integration with respect to the particular (K,T) can lead to even faster convergence. Since I have not been very interested in optimizing for speed in my implementation, I skipped such integration customization.

The choice of the parameter α has to be OK to ensure convergence of the integrand. It has been stated that $\alpha = 0.75$ or similar values can be used effectively, [33]. Checking the dependence of the prices on α should be a good check for determining if the prices have converged such that the integrals have been evaluated properly. Figure 9 is created from pricing with my code and shows that the choice of α is not important within certain bounds, as it should not. The plot in the figure continues to be flat at least until $\alpha = 50$ and indicates that any value above 0.1 and say below 50, can be used with confidence for non-extreme (K,T) . Though for extreme options, such as with a moneyness (K/S_0) of 5 and maturity the next day, it is hard to get convergence, but it is likely not due to α .

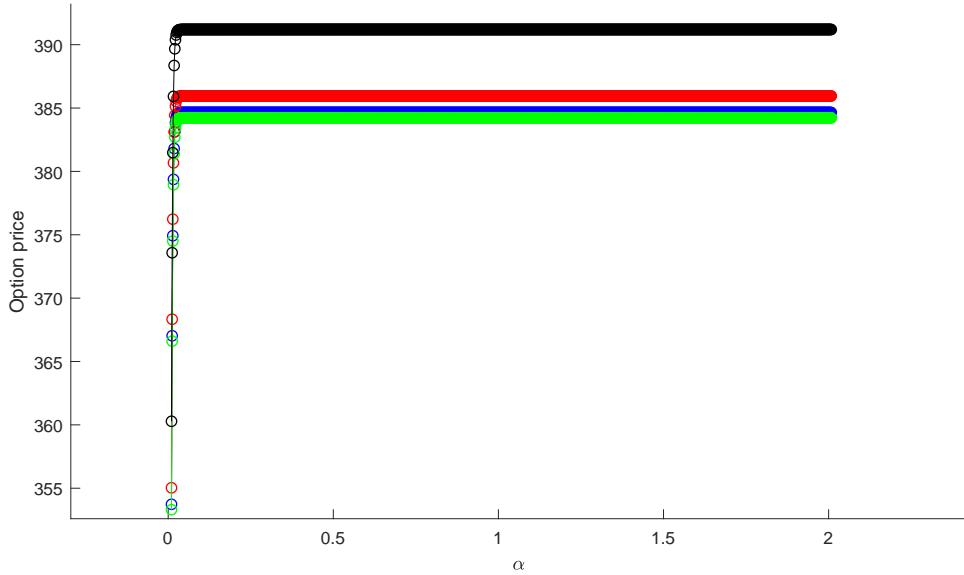


Figure 9: A plot of option price as a function of α for a moneyness of 0.85 ($K = 85\%$ of S_0) and $T = 0.2$. The prices are obtained using direct integration in equation 53. The parameters are obtained from table 1. They are all the best optima found by Nelder-Mead for the 4 models. Heston is in blue, Heston jump in red, CGMYSA in green and BNS in black.

In figure 9, it can be seen that differences in between models are far greater than any differences within a model when changing α . This indicates that the integrands have converged and that the step in the code from $\phi_T(\cdot)$ to option prices works.

In [63] the authors point to the problem of crossing complex branches in calculations with complex numbers in the Heston characteristic equation. There are two main versions of the Heston characteristic equation. The one in this report and one with the signs in front of the d's flipped, [63]. In the other version, when the variable of integration is increased, the branch cuts of the complex logarithm in the Heston characteristic equation can be crossed, leading to numerical instabilities and mispricings. This problem never occurs in the first version - not for any parameter set if one uses the principal values and a branch cut along the negative real axis, [63].

Presence of branch crossings is a potential problem in models other than Heston. Care should be maintained when handling complex numbers, in particular when taking logarithms and exponentiating with exponents less than one, where the specification of the phase angles of complex numbers affect the output. One way to check whether branch crosses can occur in a model is to simply plot $\phi_T(\cdot)$, $\psi_T(\cdot)$ or the integrand in equation 49. Figure 10 shows a plot of $\psi(v)$ in CGMYSA for some values of v which are within normal integration bounds. The model parameters used are the ones taken from the initial parameter set guess, stated in section 9.3.3. Figure 11 shows a plot of the corresponding integrand.

In addition to not crossing complex branch cuts, it is important to make sure that in the implementation, all the functions used which handle complex numbers map to the same argument range. The interval $[-\pi, \pi]$ is the standard and is the default in the math package in Java I used.

One should carefully handle discontinuous functions in the characteristic functions as well. For instance, as previously noted, in the CGMYSA model, there is a gamma function which blows up (and down), being discontinuous for an input of 0 or a negative whole number.

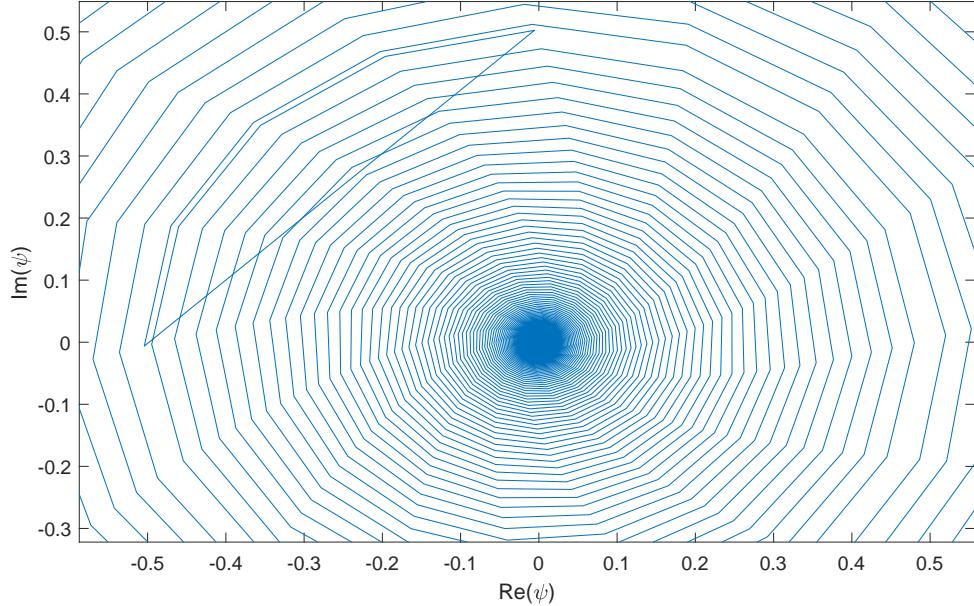


Figure 10: A plot of $\psi(v)$ for CGMYSA using the initial parameter guess $\{10.0, 20.0, 40.0, 0.4, 0.4, 0.6, 2.0, 5.0, 8.0\}$. It shows a discontinuity. It's notable that the very similar optimum $\{10.664, 18.213, 40.081, 0.3949, 0.4273, 0.6008, 1.9910, 5.0510, 7.9878\}$ found by BOBYQA and presented in table 1 does not have such a discontinuity.

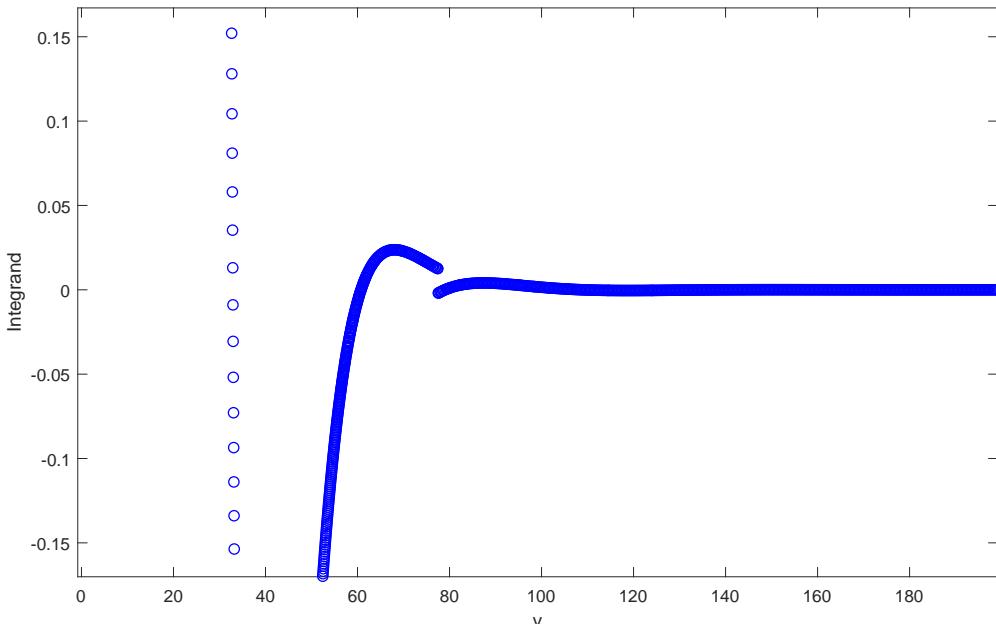


Figure 11: A plot of the integrand in equation 49 corresponding to the $\psi(v)$ in figure 10. It shows a discontinuity near $v = 77$. For $v < 50$, the values are too high/low to fit in the figure.

Now, let's see how the approach and models perform.

11 Results

11.1 EURO STOXX 50 data

I begin to test the models on the relatively well-behaved vanilla call option data on the EURO STOXX 50 on October 7th, 2003, displayed in Appendix A and extracted from [33]. As in [33], I use $S_0 = 2461.44$, $r = 0.03$ and $q = 0$. Plotting the implied volatilities against K for each T , gives figure 12.

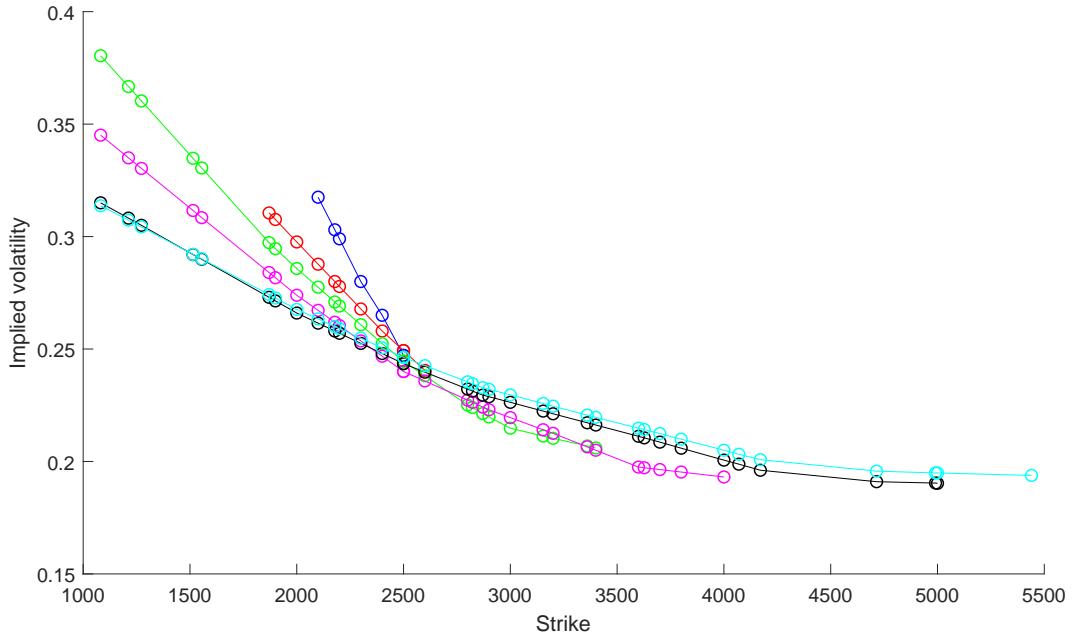


Figure 12: Market implied volatilities for the EURO STOXX 50 data in Appendix A. In ascending T 's (in years), the smiles are in blue = 0.0361, red = 0.2, green = 1.1944, magenta = 2.1916, black = 4.2056 and cyan = 5.1639.

I first tested for arbitrage in the data using the implementation of the arbitrage check in section 9.4.2. There were only 5 very small butterflies where the mid price is off by about 0.01%. These are so small that I decide that it is better to use all data points than removing points introducing such little arbitrage. 4 of those butterflies involve points with very close strikes; 2499.76 and 2500. Getting a perfect relative price for such data is difficult. In addition, whether there is arbitrage or not depends on r and q and these choices are uncertain. The fifth butterfly also involves strikes of close proximity; 4990.91 and 5000.

Then I ran model calibrations to the data, testing the optimization algorithms Nelder-Mead, CMA-ES and BOBYQA. I used arpe of option prices as objective function and ran all the models: Heston, Heston-Jump, BNS and CGMYSA. The algorithms were run 5 times. The first with the initial guesses of section 9.3 and the 4 remaining with randomly chosen initial guesses. I used a limit of 8000 maximum objective function evaluations and algorithm iterations. I did the integral in equation 49 from $v = 10^{-8}$ to $v = 250$, with $\Delta v = 0.02$. α was set to 0.75. I did check that these v -values gives well-converged integrals for the data implied (K,T) . The resulting parameters and fits are captured in table 1. I prioritized accuracy over speed in the calibrations. I have evaluated the integrals in more points than necessary, hence the speed is quite slow. With some speed ups, I think a calibration time of around 2 minutes should be attainable when calibrating one model using one algorithm and perhaps one run if the code is optimized.

Table 1: Results from calibrations to arpe of option prices for the EURO STOXX 50 data. When several optima were found, also the best non-degenerate optimum is displayed (for instance indicated by a 2 in contrast to a 1). Though, sometimes the maximum amount of evaluations were used up before 5 runs could be done. Time is the total time elapsed for all runs. Evaluations is the total number of objective function evaluations used for all runs. Hence time and evaluations are equal for a model-algorithm pair but differing optima. Arpe(vols) are calculated by using a different objective function.

Model - Algorithm: Optimum	Model parameters	arpe (prices)	arpe (vols)	Time (s)	Evaluations
BNS - Nelder-Mead: 1	{-2.6369, 0.5854, 19.381, 1.5168, 0.2264}	0.0126	0.0285	586	1237
BNS - Nelder-Mead: 2	{-2.9957, 0.3180, 27.929, 6.4433, 0.3133}	0.0387	0.0678	586	1237
BNS - CMA-ES: 1	{-1.4000, 0.3815, 9.5272, 1.6866, 0.2377}	0.0099	0.0207	4510	10245
BNS - CMA-ES: 2	{-1.1434, 0.4110, 18.498, 9.9806, 0.3008}	0.0503	0.0799	4510	10245
BNS - BOBYQA: 1	{-4.5425, 0.6245, 18.287, 0.5541, 0.2018}	0.0192	0.0289	341	653
BNS - BOBYQA: 2	{-1.3537, 4.1884, 29.937, 0.1721, 0.1492}	0.1571	0.1473	341	653
Heston - Nelder-Mead: 1	{-0.6690, 0.2294, 0.1577, 0.1065, 0.2451}	0.0112	0.0215	1175	1951
Heston - Nelder-Mead: 2	{-0.6543, 0.4370, 0.9352, 0.0664, 0.2733}	0.0127	0.0135	1175	1951
Heston - CMA-ES: 1	{-0.6576, 0.3075, 0.4523, 0.0756, 0.2541}	0.0076	0.0104	5016	12030
Heston - CMA-ES: 2	{-0.6319, 0.5000, 1.0087, 0.0672, 0.2770}	0.0147	0.0170	5016	12030
Heston - BOBYQA: 1	{-0.6575, 0.4000, 0.8601, 0.0664, 0.2682}	0.0116	0.0122	1673	2707
Heston - BOBYQA: 2	{-0.8945, 0.0099, 0.1877, 0.1137, 0.1023}	0.1938	0.2866	1673	2707
Heston Jump - Nelder-Mead: 1	{-0.7903, 0.2829, 0.0927, 0.1601, 0.2351, 1.0905, 0.0802, 0.0349}	0.0125	0.0229	1375	2783
Heston Jump - Nelder-Mead: 2	{-0.6997, 3.9246, 1.8170, 0.1563, 0.5561, 1.03E-5, 3.3E-7, 0.3063}	0.0973	0.2056	1375	2783
Heston Jump - CMA-ES: 1	{-0.9384, 0.2729, 0.4941, 0.0644, 0.2386, 0.4388, 0.1500, 0.0000}	0.0073	0.0107	5035	40005
Heston Jump - BOBYQA: 1	{-0.8913, 0.2894, 0.2143, 0.0946, 0.2328, 1.0915, 0.0892, 0.0471}	0.0105	0.0187	4710	7557
Heston Jump - BOBYQA: 2	{-0.9035, 3.3516, 0.2561, 0.5905, 0.3388, 1.4806, 0.1156, 0.0037}	0.0474	0.1052	4710	7557
CGMYSA - Nelder-Mead: 1	{12.152, 16.437, 31.109, 0.2306, 0.2810, 0.5587, 2.0602, 5.5482, 7.3410}	0.0120	0.0239	7141	30846
CGMYSA - CMA-ES: 1	{297.46, 121.29, 199.12, 0.3555, -0.0418, 2.000, 30.00, 38.792, 17.227}	0.0788	0.2295	6896	40005
CGMYSA - BOBYQA: 1	{10.664, 18.213, 40.081, 0.3949, 0.4273, 0.6008, 1.9910, 5.0510, 7.9878}	0.0119	0.0254	605	750
CGMYSA - BOBYQA: 2	{68.156, 279.56, 240.23, 0.050, -0.8307, 1.1954, 24.281, 33.917, 51.195}	0.3801	1.7343	605	750

In table 1, it can be seen that fits to the data (arpes) for option prices as well as implied volatilities of 1-2% is attainable for all models. The fits to vols are slightly worse than the fits to prices which is to be expected since the calibration is to prices. All model-algorithm combinations find parameter sets which give good fits to the data, apart from the non-impressive fit found by CGMYSA - CMA-ES. CMA-ES finds the best optimum for the rest of the models, though. It can be seen that the code is quite computationally demanding. One objective function evaluation takes more than one second. The CMA-ES algorithm runs for longer than the other algorithms. This is highly dependent on the stopping criteria.

The optima found give quite different parameter sets despite their similar fitting measures. See for example how the best optima found in Heston Jump from Nelder-Mead and CMA-ES differ, even though their arpes are both low. I investigate the price and vol model-market matches of these two optima in figures 13, 14, 15 and 16.

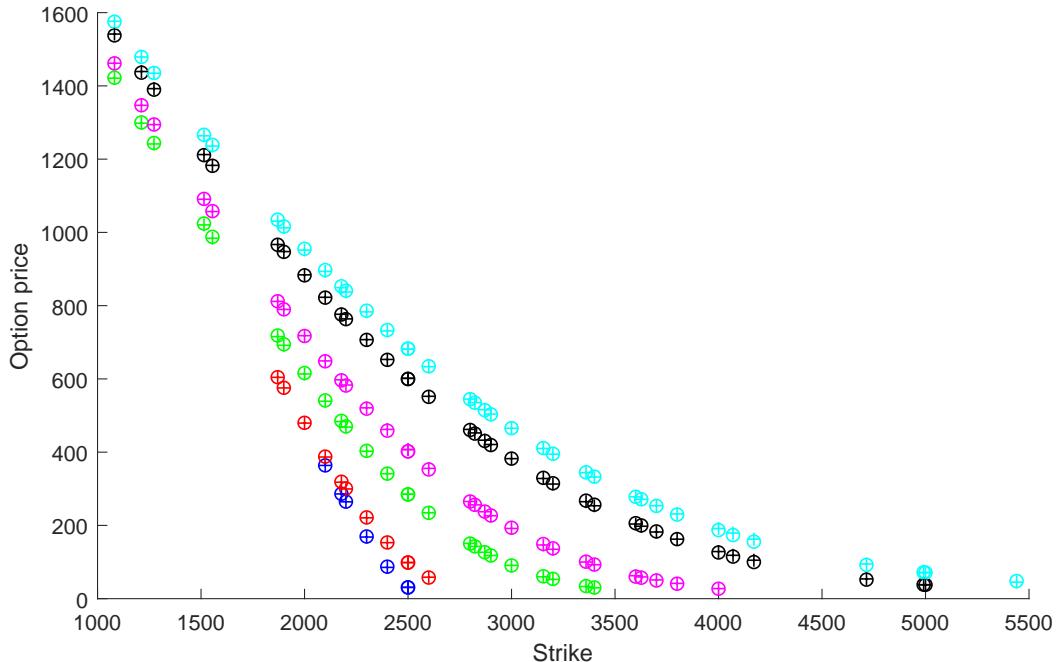


Figure 13: Model - market price match for Heston Jump - CMA-ES: 1. Market prices are represented by circles and the corresponding model prices by plus signs. The same color map as in figure 12 is used.

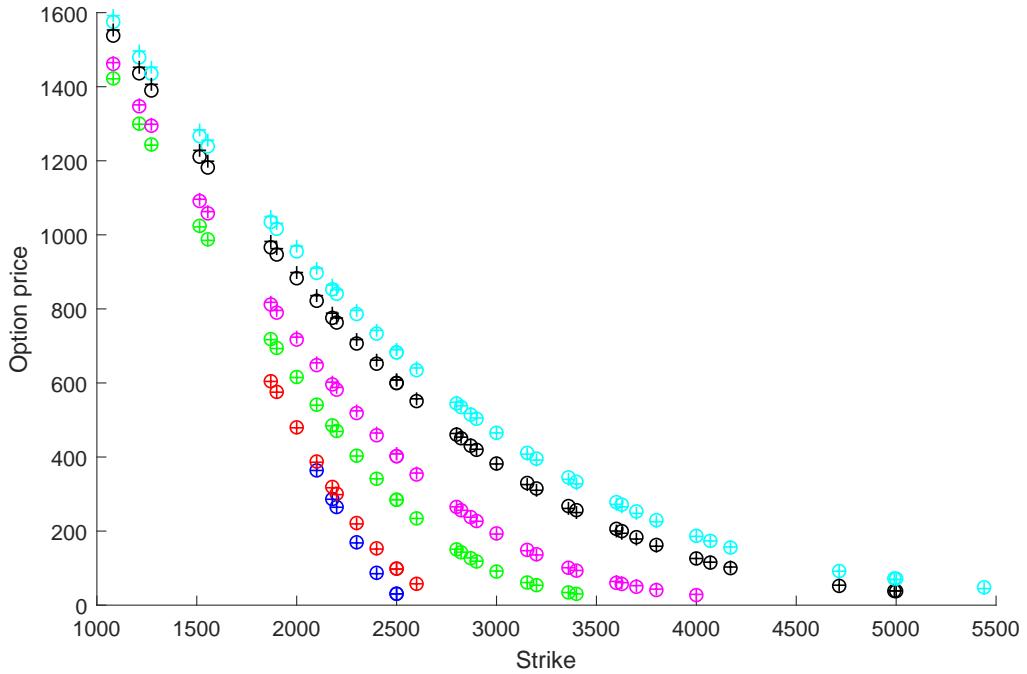


Figure 14: Model - market price match for Heston Jump - Nelder-Mead: 1. Market prices are represented by circles and the corresponding model prices by plus signs.

The higher arpe(prices) in the Nelder-Mead compared to the CMA-ES optimum is visible in the above figures. When plotting the market-model vol matches, it can be seen that there are worse matches for low strikes:

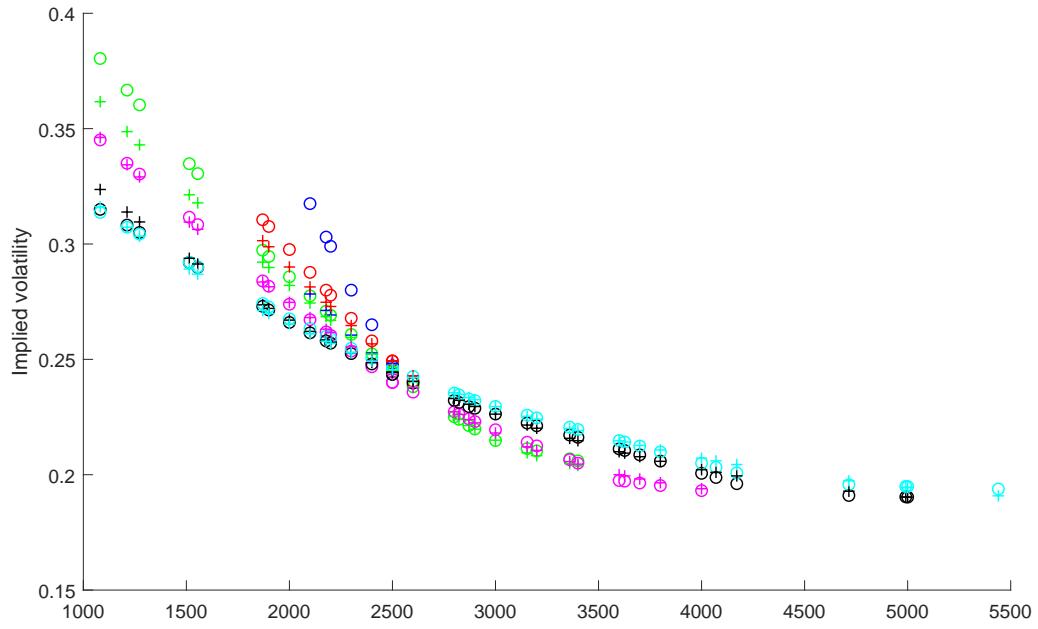


Figure 15: Model - market vol match for Heston Jump - CMA-ES: 1. Market vols are represented by circles and the corresponding model vols by plus signs.

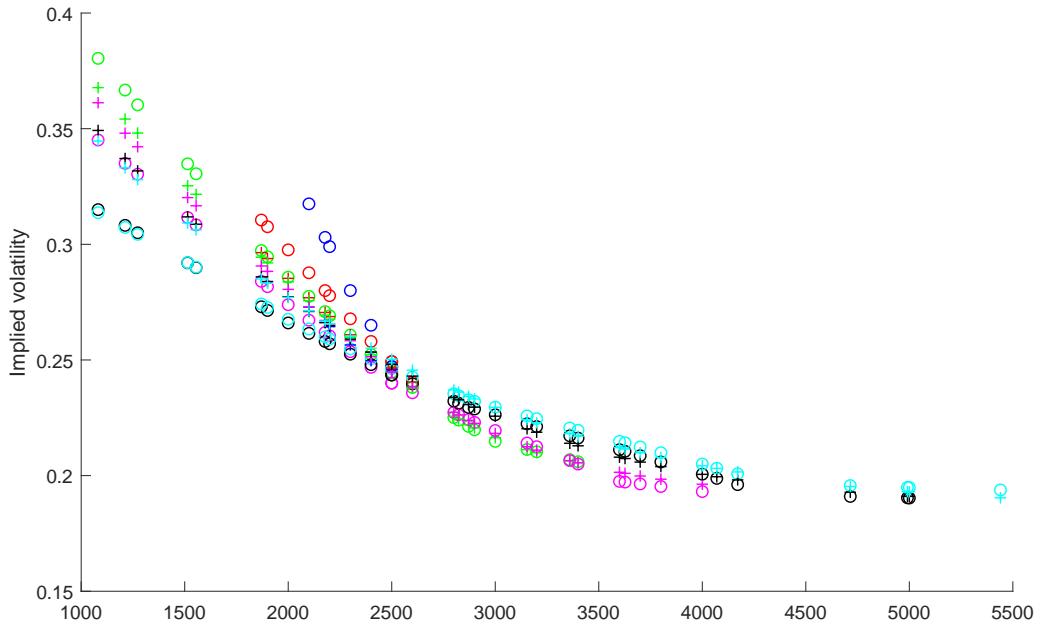


Figure 16: Model - market vol match for Heston Jump - Nelder-Mead: 1. Market vols are represented by circles and the corresponding model vols by plus signs.

Being curious how the corresponding surfaces differ, I plotted the surfaces obtained from the parameters in Heston-Jump - Nelder-Mead: 1 and Heston Jump - CMA-ES: 1 in terms of log(moneyess). These can be seen in figures 17 and 18.

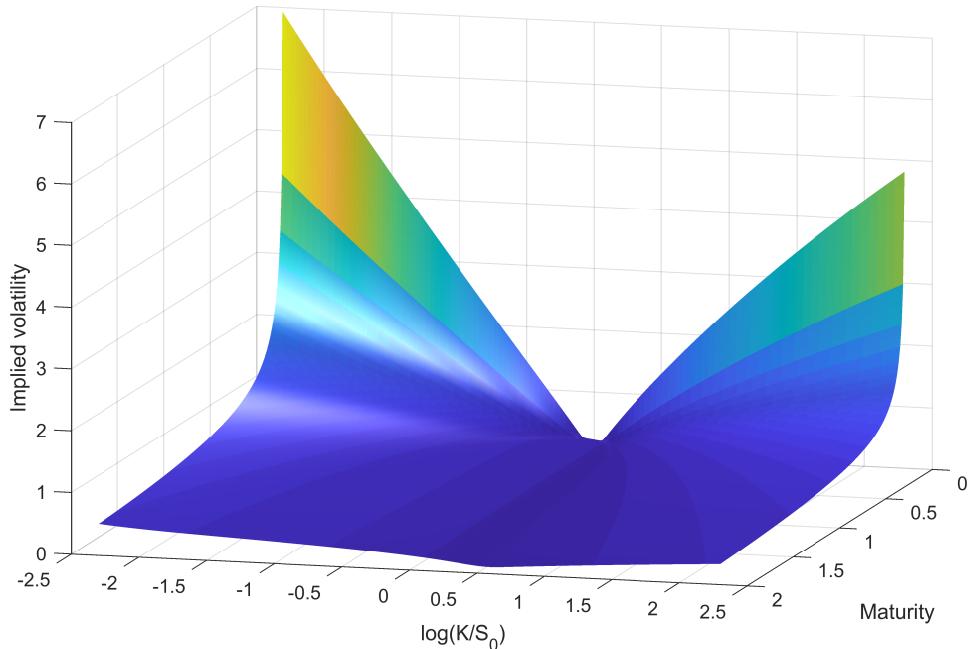


Figure 17: The Heston Jump vol surface obtained from CMA-ES with the parameters $\{-0.9384, 0.2729, 0.4941, 0.0644, 0.2386, 0.4388, 0.1500, 0.0000\}$.

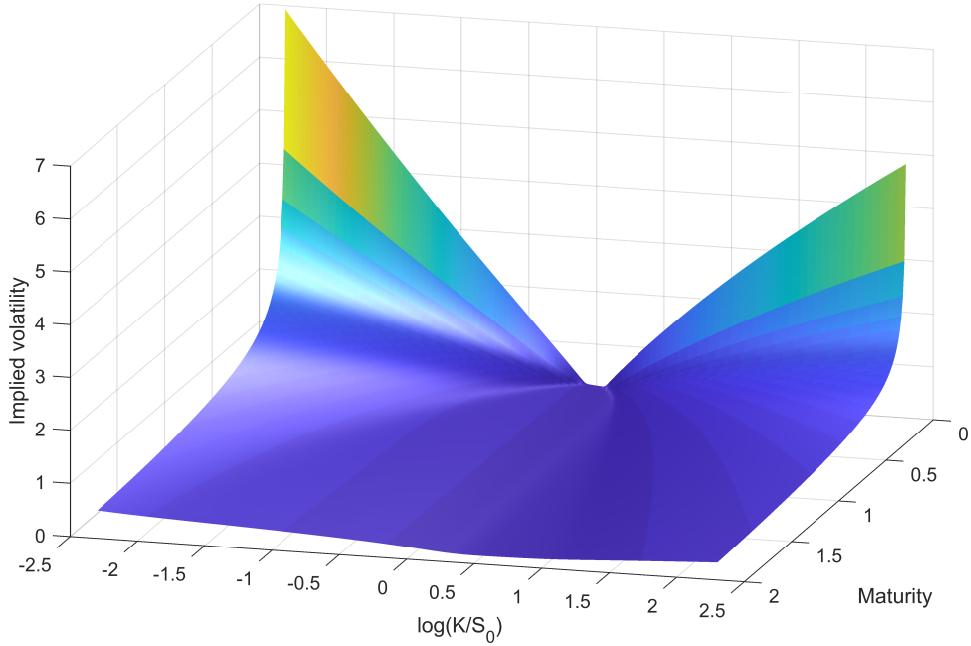


Figure 18: The Heston Jump vol surface obtained from Nelder-Mead with the parameters $\{-0.7903, 0.2829, 0.0927, 0.1601, 0.2351, 1.0905, 0.0802, 0.0349\}$.

The surfaces have a very similar shape but there are bigger differences between them in absolute values than I would have anticipated. Subtracting the surface in figure 18 from figure 17, dividing by the values in figure 18 and multiplying by 100 to get percentage differences produces figure 19.

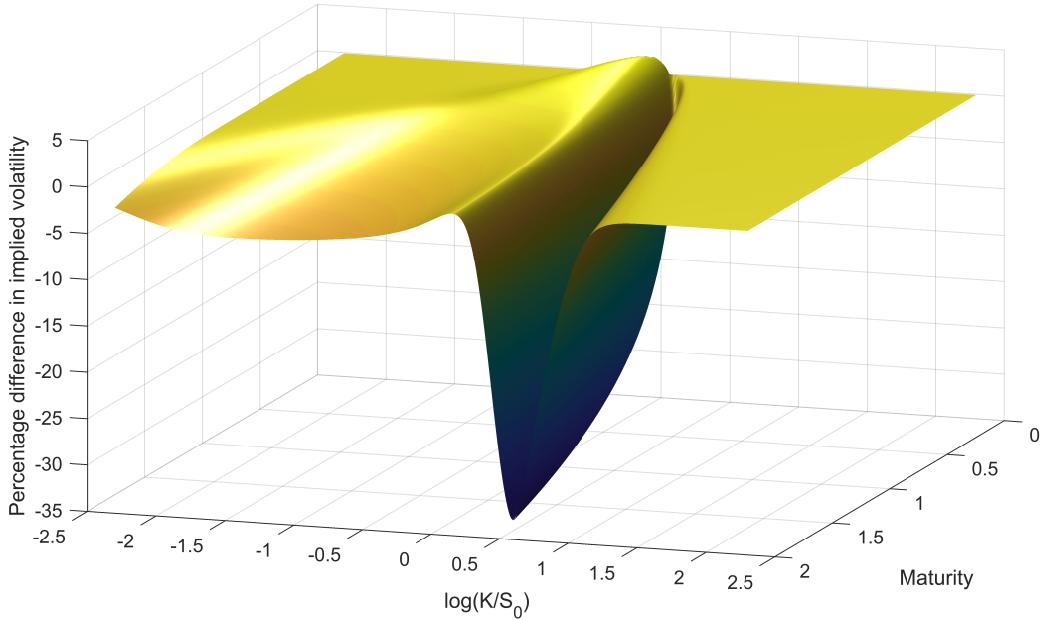


Figure 19: Percentage differences in the two vol surfaces in figures 17 and 18 produced by the Heston jump model.

The differences are drastic for some non-extreme (K, T) . This indicates that there could exist many surfaces within a model which can fit the data well and yet differ at typical points of interest, if one considers a 2% arpe a good fit.

In the cases we have investigated, the match to implied volatilities is worse for small strikes. This could be due to calibration being done to prices and vega varying for different strikes. In Black-Scholes the vega for a European call is

$$vega = \frac{\partial C}{\partial \sigma} = Se^{-qT}n\left(\frac{\ln(S_0/K) + (r - q + \sigma^2/2)T}{\sigma\sqrt{T}}\right)\sqrt{T} \quad (72)$$

where $n(\cdot)$ is the probability density function of the normal distribution. From this equation it can be seen that high or low moneyness leads to small vegas. Plotting vega for the specific EURO STOXX numbers with K varying from 0 to 8000, produces figure 20.

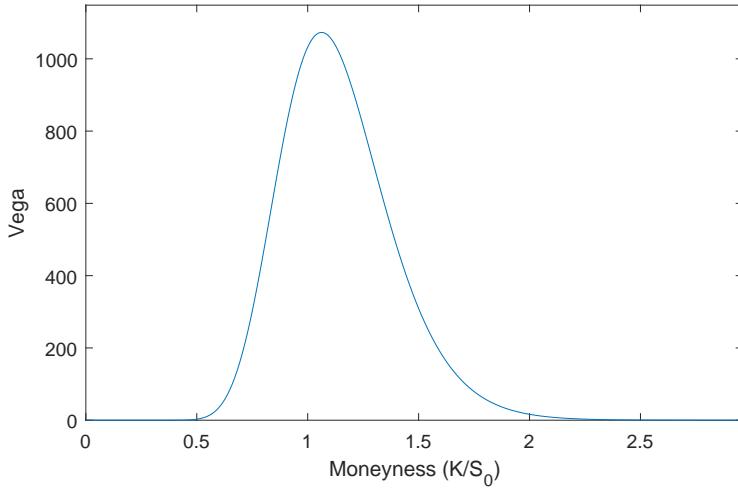


Figure 20: Vega for the EURO STOXX 50 data, using $r = 0.03$, $q = 0.0$, $T = 1.1944$, $S_0 = 2461.3$ and $\sigma = 0.2$.

It can be seen that the vega is highest close to at-the-money and decreases sharply around its peak. One market data point has $K = 1081$ and there, from the figure, $vega = 0.32$ which is about a factor of 3000 smaller than at the peak. A high $\frac{\partial C}{\partial \sigma}$ gives good conversion stability from option prices to implied volatility. That is, a slight mispricing does not produce a large miss when converted to implied volatility. A low vega on the other hand means a small mispricing blows up to bigger misses for implied vols. This suggests that calibration to implied vols leads to a more well distributed fit across all data points. I will test this further down.

Before that, I check the validity of some of the found optima in table 1. I start by naively checking the continuity of the integrands for the randomly chosen $K = 5 \times S_0$ and $T = 2$. I show no plots but simply claim that I could not find any large discontinuities for the best optima found by all model-algorithm combinations up to $v = 300$, except for CGMYSA - CMA-ES: 1. Figure 21 shows the discontinuity in the integral from the parameters in CGMYSA - CMA-ES: 1.

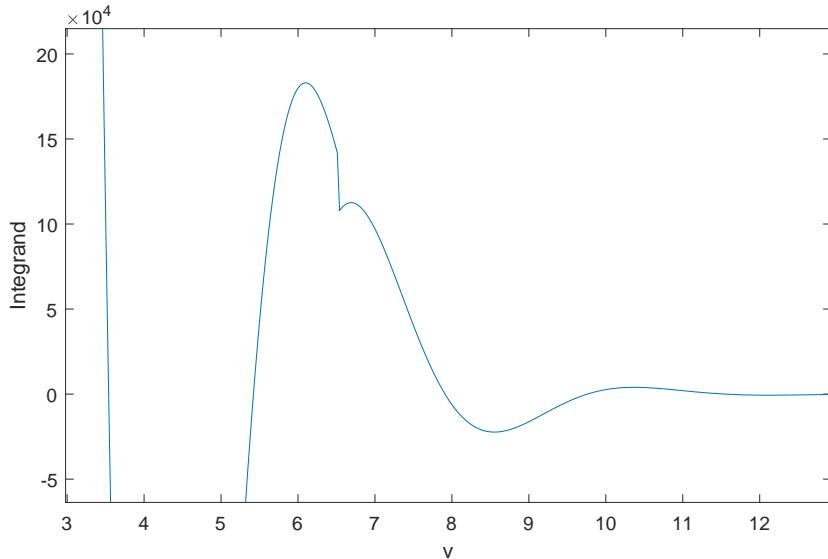


Figure 21: CGMYSA - CMA-ES: 1 having a discontinuous integral for a small v .

The value of the integral is about of the order 10^7 . Hence small discontinutities such as this one does not lead to serious mispricings, although combined with ones for other (K, T) might give rise to arbitrage opportunities.

Interestingly, the speed of convergence of the integrals in the models vary significantly. BNS is converging much faster than Heston and Heston Jump, and CGMYSA converging much slower.

To check for arbitrage in the surfaces, I ran my arbitrage check. Figures 22 and 23 show top-down views of the Heston Jump - CMA-ES: 1 and Heston Jump - Nelder-Mead: 1 surfaces of table 1, with arbitrage in yellow.

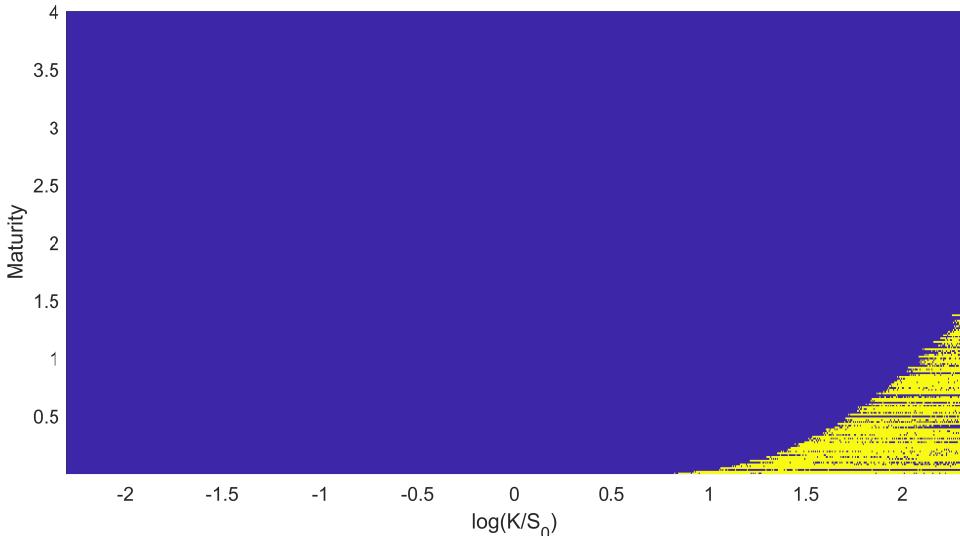


Figure 22: Arbitrage in the vol surface from Heston Jump - CMA-ES: 1, in table 1. The yellow points are points which fail the check and hence have static arbitrage. The indigo points are points which pass the check.

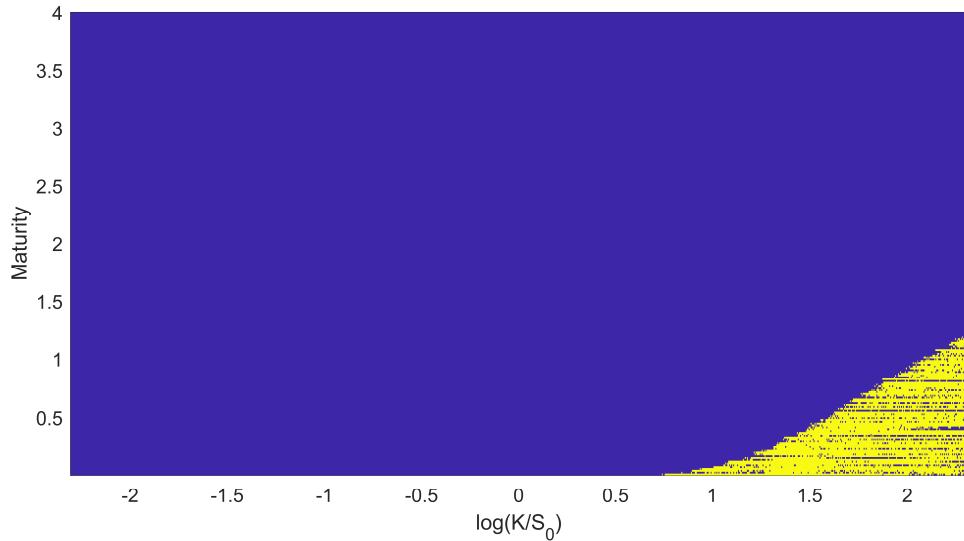


Figure 23: Arbitrage in the vol surface from Heston Jump - Nelder-Mead: 1, in table 1. The yellow points are points which fail the check and hence have static arbitrage. The purple points are points which pass the check.

As can be seen in these figures, for normal (K, T) the surfaces are arbitrage-free. Only for high strikes and short maturities, where prices are hard to pinpoint, there is arbitrage.

I now calibrate to arpe of implied volatilities for the same data set. This should not be much slower since what takes time is calculating the integrals for each option using a thousand points or more, not doing the Black-Scholes inversion using Brent's method, commonly requiring less than 10 iterations for good accuracy. I used the same calibration settings as for calibrating to prices, apart from doubling the maximum number of allowed objective function evaluations and iterations, to see if more time gives better results. The results are shown in table 2 below.

Table 2: Results from calibrations to arpe of implied volatilities for the EURO STOXX 50 data. When several optima were found, also the best non-degenerate optimum is displayed (for instance indicated by a 4 in contrast to a 1). In the Heston model, Nelder-Mead and CMA-ES found the best optima three times. Hence the forth optima are presented. Time is the total time elapsed for all runs. Evaluations is the total number of objective function evaluations used for all runs. Hence time and evaluations are equal for a model-algorithm pair but differing optima. Arpe(prices) are calculated by using a different objective function.

Model - Method: Optimum	Model parameters	arpe (prices)	arpe (vols)	Time (s)	Evaluations
BNS - Nelder Mead: 1	{-3.4592, 0.5997, 18.368, 0.8663, 0.2126}	0.0163	0.0241	1163	2772
BNS - Nelder Mead: 2	{-0.3274, 1.5180, 29.999, 5.2919, 0.2909}	0.1302	0.1002	1163	2772
BNS - CMA-ES: 1	{-1.4309, 0.3822, 9.1560, 1.4383, 0.2377}	0.0114	0.0197	8866	18629
BNS - BOBYQA: 1	{-3.6591, 0.6110, 17.949, 0.7305, 0.2106}	0.0178	0.0244	1533	2807
BNS - BOBYQA: 2	{-5.2615, 8.9769, 19.647, 0.0195, 0.1848}	0.1026	0.0725	1533	2807
Heston - Nelder-Mead: 1	{-0.6529, 0.3451, 0.5645, 0.0724, 0.2592}	0.0082	0.0084	1332	2141
Heston - Nelder-Mead: 4	{-0.4018, 1.5282, 0.2402, 0.3513, 0.2587}	0.1466	0.0800	1332	2141
Heston - CMA-ES: 1	{-0.6529, 0.3448, 0.5638, 0.0725, 0.2592}	0.0082	0.0084	9961	20106
Heston - CMA-ES: 4	{-0.4139, 1.5057, 0.1935, 0.4254, 0.2567}	0.1443	0.0800	9961	20106
Heston - BOBYQA: 1	{-0.6520, 0.3416, 0.5371, 0.0733, 0.2587}	0.0081	0.0085	2847	4440
Heston - BOBYQA: 2	{-0.7292, 0.5574, 1.6497, 0.0675, 0.2630}	0.0328	0.0217	2847	4440
Heston Jump - Nelder-Mead: 1	{-0.8053, 0.3304, 0.5768, 0.0657, 0.2472, 1.0575, 0.0572, 0.0554}	0.0080	0.0085	3654	6780
Heston Jump - Nelder-Mead: 2	{-0.7518, 0.2080, 0.0127, 0.5635, 0.2484, 0.0046, 0.5900, 0.3691}	0.0268	0.0306	3654	6780
Heston Jump - CMA-ES: 1	{-0.7986, 0.3243, 0.4583, 0.0661, 0.2371, 1.8087, 0.0000, 0.0738}	0.0076	0.0069	10111	80005
Heston Jump - BOBYQA: 1	{-0.7878, 0.3224, 0.4172, 0.0678, 0.2371, 1.6593, 0.0000, 0.0768}	0.0073	0.0071	10175	16001
Heston Jump - BOBYQA: 2	{-0.8665, 0.3196, 0.5318, 0.0646, 0.2415, 1.0852, 0.0919, 0.0205}	0.0078	0.0089	10175	16001
CGMYSA - Nelder-Mead:1	{7.5922, 28.369, 48.026, 0.5659, 0.6140, 0.7032, 2.0909, 3.3511, 5.2477}	0.0136	0.0126	4549	5278
CGMYSA - Nelder-Mead:2	{125.420, 4.0228, 7.7756, -1.8972, -1.1298, 0.3021, 25.768, 5.5866, 29.858}	0.1347	0.1717	4549	5278
CGMYSA - CMA-ES: 1	{77.209, 55.293, 105.35, 0.2880, 0.3382, 0.4702, 1.4191, 40.000, 16.303}	0.0080	0.0078	14105	80005
CGMYSA - BOBYQA: 1	{11.683, 23.964, 40.455, 0.3753, 0.4298, 0.6303, 2.1572, 5.5001, 7.4837}	0.0162	0.0159	4459	5124
CGMYSA - BOBYQA: 2	{318.51, 390.66, 45.239, -0.9227, 0.5229, 0.8226, 25.102, 20.185, 33.081}	0.1025	0.2124	4459	5124

The calibration times per objective function evaluation are only a few percent longer here compared to in the calibration to prices. Most fits are better, even arpe(prices). This is potentially because of the allowed doubling of max evaluations. Since such good optima are found, it seems that the search space is not much more complicated when calibrating to implied volatilities. For some parameters, their set lower/upper bounds are found to be optimal.

The CMA-ES algorithm once again runs for longer, using more evaluations than the other algorithms but does find the best optimum in all four models. It seems the other two algorithms do no use up all the allowed evalutions. This could be changed by tweaking the stopping criteria. The Nelder-Mead and BOBYQA algorithms perform similarly when calibrating to prices and the BOBYQA does slightly better when calibrating to vols, though for Heston and Heston Jump it is using 2 and 3 times the duration, respectively.

It is hard to pick a model winner or loser for these data. In the vol calibration, perhaps it was easier to find the best fits in the Heston Jump model, but the others are not far behind in terms of arpe. The BNS and CMGYSA could perhaps be said to do slightly worse, CGMYSA mostly because it was harder to find the best optimum. All models could fit well enough in my opinion and all algorithms find good optima.

In figure 24, I plot the vol match for one of the optima with a vol arpe < 0.9%; the Heston - BOBYQA: 1. In it, it can be seen that still it is most difficult to fit low strikes.

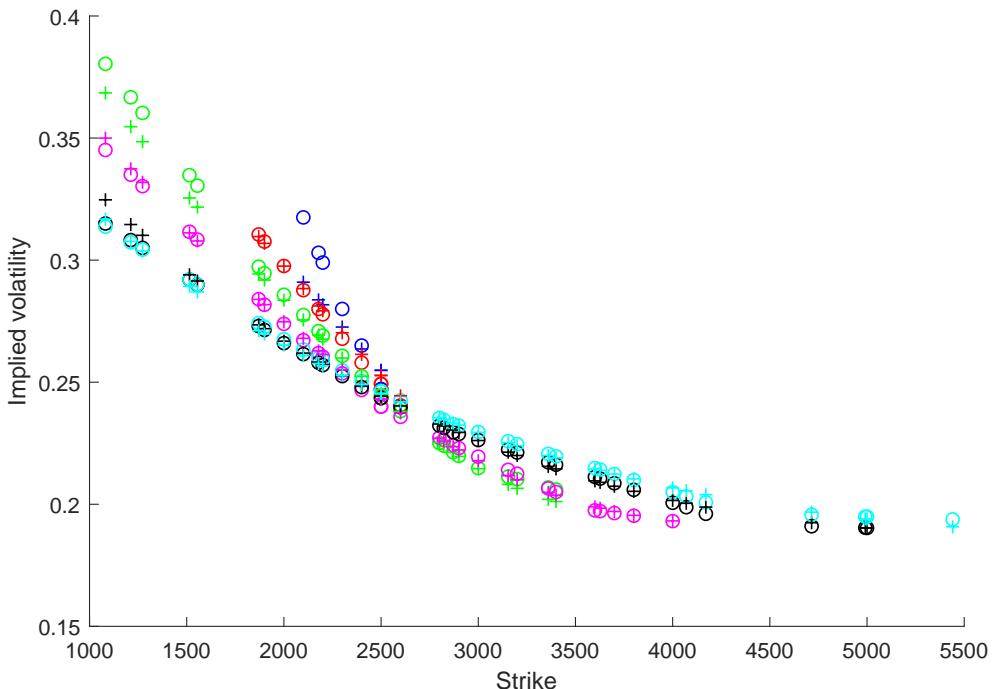


Figure 24: Market - model vol match for the Heston model; using the best parameters found by BOBYQA when calibrating to vols. Market vols are represented by circles and model vols by plus signs.

I'm curious about the differences in vol surfaces in between models. Therefore, I created figure 25 which shows vol differences in the Heston and Heston Jump models. It can be seen that for most (K,T), the differences are very small.

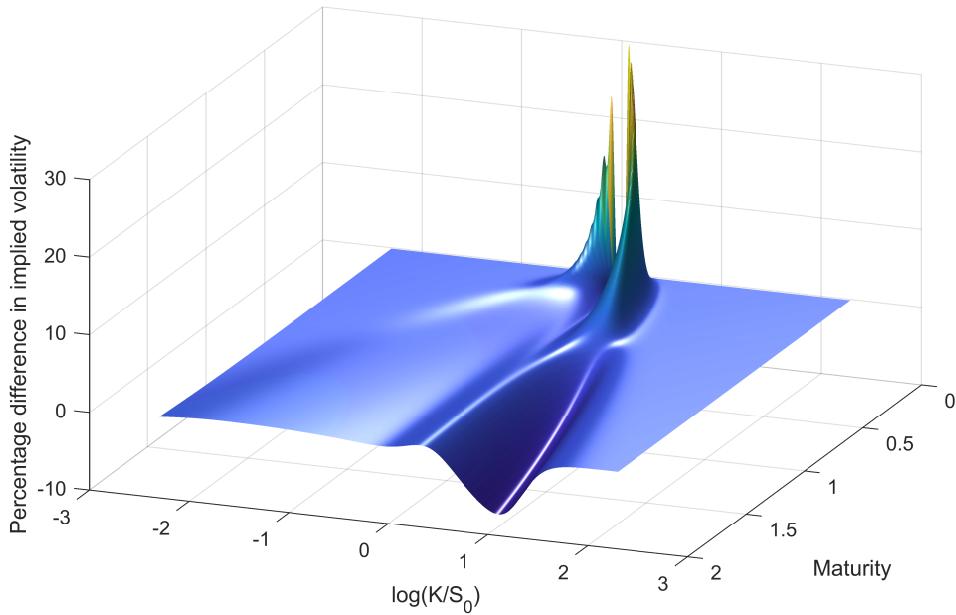


Figure 25: Showing percentage differences in vol surfaces from table 2. This is a plot of 100 times the vol surfaces of (Heston Jump - BOBYQA: 1) minus (Heston - BOBYQA: 1), both divided by (Heston - BOBYQA: 1).

I'm also curious to check whether very similar parameter sets, each being local optima in the same model, produce similar surfaces. This of course depends a lot on model and parameter choices. Nevertheless, I look at the vol differences between the Nelder-Mead and BOBYQA optima for the Heston models having practically the same arpes and the very similar parameters; $\{-0.6529, 0.3451, 0.5645, 0.0724, 0.2592\}$ and $\{-0.6520, 0.3416, 0.5371, 0.0733, 0.2587\}$, respectively. This gives figure 26.

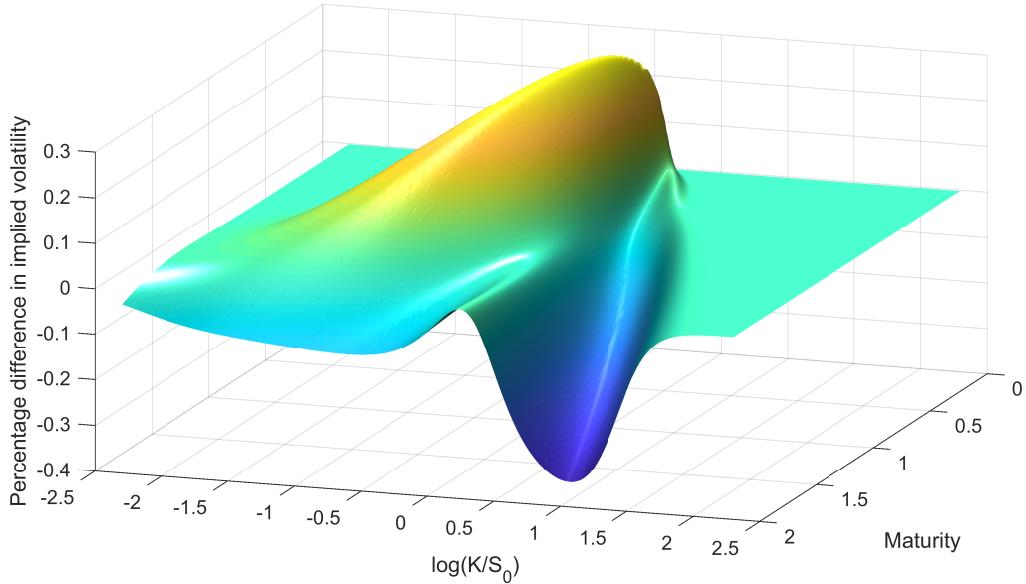


Figure 26: Differences in vol surfaces for Heston - Nelder-Mead 1 and Heston - BOBYQA: 1, of table 2.

As expected, there are no big differences between these surfaces for any (K, T) included in the

figure.

11.2 Nikkei 225 data

Next, I calibrated models to a more challenging data set. I used all European call option closing prices on Nikkei 225, traded on JPX and found in their daily report of September 27th 2017 and presented in Appendix B. I chose to set S_0 to the same value for all trades that day; the underlying's closing price of 20'267 - adjusted for dividends and splits. Moreover, I chose $r = -0.001$ as being representative of for all T's, based on a TIBOR curve. I set $q = 0$, since S_0 was adjusted for dividends. The data is illustrated in figure 27.

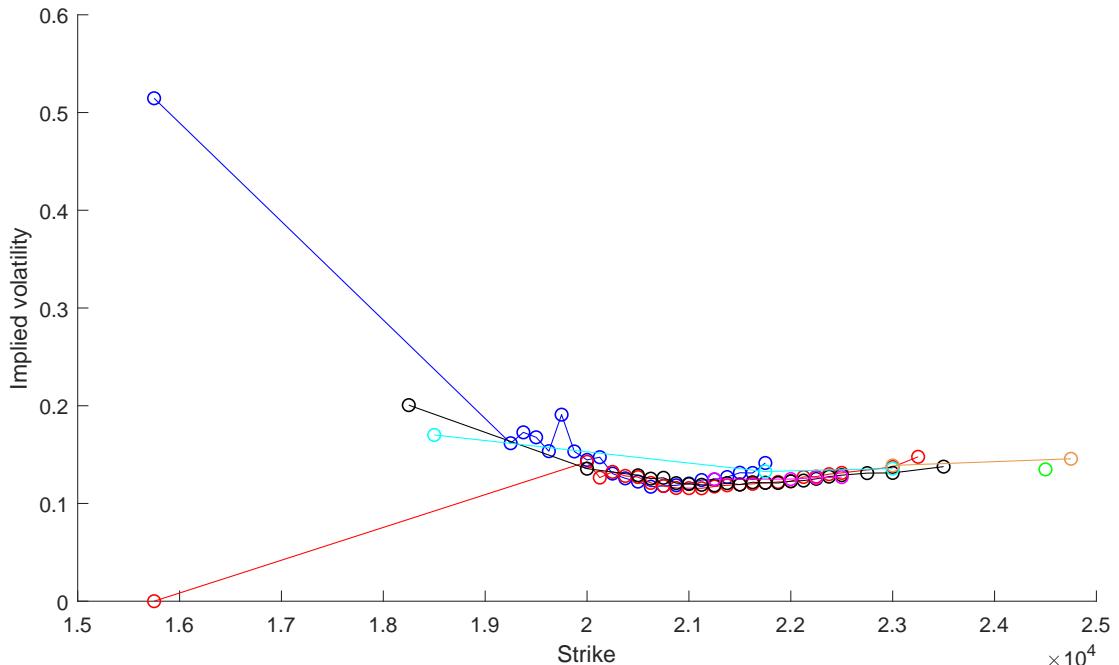


Figure 27: Market implied volatilities for the Nikkei data in Appendix B, when $S_0 = 20'267$, $r = -0.001$ and $q = 0$ are used. Implied volatilities which could not be found between 0 and 1000, are set to zero. In ascending T's (in years), the smiles are in blue = 0.0384, red = 0.115, black = 0.1918, magenta = 0.288, green = 0.3644, cyan = 0.4411 and brown = 0.69.

Running the arbitrage check on these prices leads to more arbitrage than in the previous data. There are 11 out of the 78 data points which give rise to static arbitrage according to the check. Most of them are negative butterflies where the middle price is off by about 1-3 %. All are in the range 0.2-7% off. For one point, the check; $C > S_0 - K$ is failed and this is quite serious because then one cannot solve for the volatility. Call options are at least worth their intrinsic value and I cannot believe the trade was made where this check was overruled. It speaks for the low quality of the setup, but let's find vol surfaces anyway. There is also one point-pair with calendar arbitrage: The points with strike 15'750, as can be seen in Appendix B.

Even though there are some problematic data points, I try to calibrate to the complete data. I set arpe of option prices as objective function since calibrating to vols is not straight-forward with inclusion of the point with undefined vol. I change the upper limit of integration from 250 to 350 and use a maximum number of objective function evaluations as well as algorithm iterations of 16'000. As previously, the algorithms are run 5 times with different initial guesses. The results can be seen in table 3.

Table 3: Results from calibrations to arpe of option prices for the Nikkei 225 data. When several optima were found, also the best non-degenerate optima is displayed (for instance indicated by a 3 in contrast to a 1). Time is the total time elapsed for all runs. Evaluations is the total number of objective function evaluations used for all runs. Hence time and evaluations are equal for a model-algorithm pair but differing optima. Arpe (vols) is not included since one vol cannot be found. In the Heston model, CMA-ES found the same optimum 4 times, then stopped.

Model - Method: Optimum	Model parameters	arpe (prices)	Time (s)	Evaluations
BNS - Nelder-Mead: 1	$\{-3.0275, 0.0468, 19.404, 8.8257E - 8, 0.1249\}$	0.1577	433	1573
BNS - Nelder-Mead: 2		1.0000	433	1573
BNS - CMA-ES: 1		0.0800	7776	80005
BNS - BOBYQA: 1		0.0881	311	629
BNS - BOBYQA: 2		0.1581	311	629
Heston - Nelder-Mead: 1	$\{-0.5079, 0.5854, 0.1150, 0.3099, 0.1351\}$	0.0935	1072	2056
Heston - Nelder-Mead: 3		0.1470	1072	2056
Heston - CMA-ES: 1		0.0806	10779	18426
Heston - BOBYQA: 1		0.1711	425	628
Heston - BOBYQA: 2		0.1750	425	628
Heston Jump - Nelder-Mead: 1	$\{-0.6533, 0.8617, 0.4570, 0.1749, 0.1305, 1.0999, 0.0371, 3.5250E - 4\}$	0.0539	3245	5990
Heston Jump - Nelder-Mead: 2		0.0950	3245	5990
Heston Jump - CMA-ES: 1		0.0460	10826	37391
Heston Jump - CMA-ES: 2		0.0508	10826	37391
Heston Jump - BOBYQA: 1		0.1169	1374	2006
Heston Jump - BOBYQA: 2		0.1301	1374	2006
CGMYSA - Nelder-Mead: 1	$\{3.4827, 63.670, 128.38, 0.6714, 0.6635, 0.8898, 2.8880, 5.8248, 8.1806\}$	0.0741	4329	5022
CGMYSA - Nelder-Mead: 2		0.1074	4329	5022
CGMYSA - CMA-ES: 1		0.0780	17762	80005
CGMYSA - BOBYQA: 1		0.0658	1615	1456
CGMYSA - BOBYQA: 2		0.1151	1615	1456

The fits obtained are not impressive. It is apparently hard to find an arpe(prices) below 5%. In addition, I found that most models produce some vols which cannot be found within [0,1000].

I now try to clean the data from significant arbitrage opportunities. First I remove the data point for which the implied volatility could not be found, that is for $(K,T) = (15'750,0.115)$. Removing this point also removes the one and only calendar arbitrage opportunity. In addition, I decide to remove butterfly arbitrage where the mid price is more than 5% off from a the convexity/concavity boarder and where I consider the surrounding points to be OK. There are two such points: $(K,T) = (19'750,0.0384)$ and $(22'125,0.115)$ where the mid prices are off by 6.4% and 7.1%, respectively. Left are 8 butterfly arbitrage points where the mid prices are off by 0 – 3.2%. These could be removed too but I'm unsure about whether these points are actually worse than the surrounding points. Additionally, I want to see how the models react when there is arbitrage in the data. The cleaned up data can be seen below in figure 28.

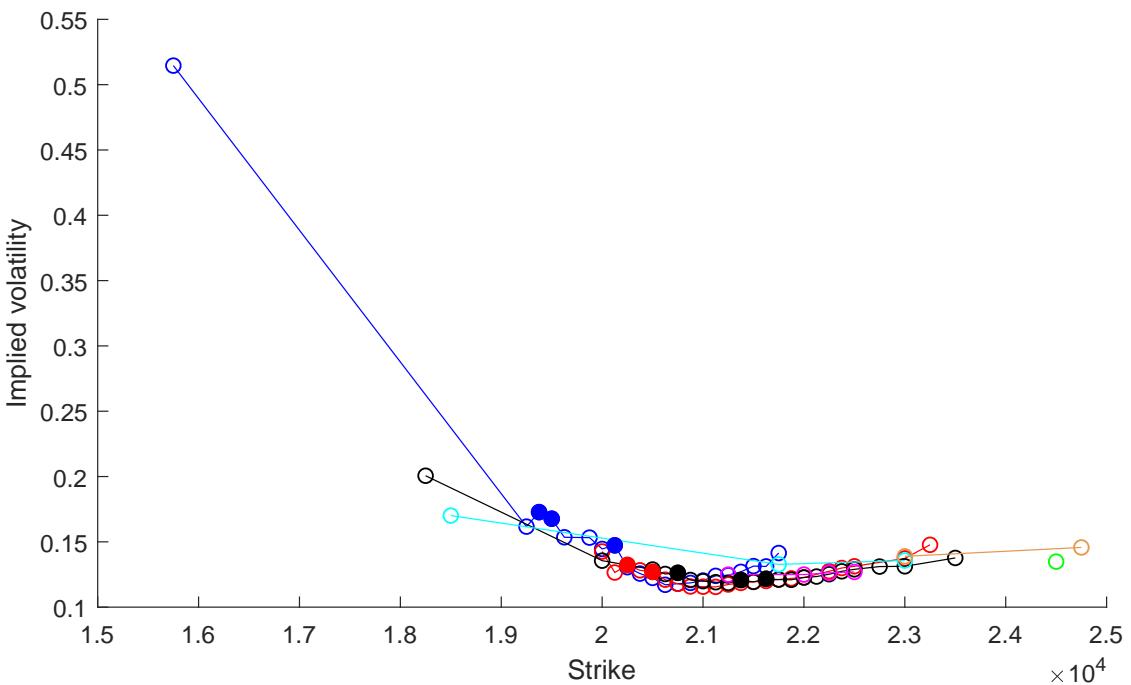


Figure 28: The Nikkei data with some arbitrage-introducing points removed. The mid points of the 8 remaining butterflies are represented by filled circles. In ascending T's (in years), the smiles are in blue = 0.0384, red = 0.115, black = 0.1918, magenta = 0.288, green = 0.3644, cyan = 0.4411 and brown = 0.69.

The first dark blue point is quite high, and will be the chief determinant of vols at low strikes. It could be ignored since it is quite uncertain, but it is also good to have data there. Note that concavity in vols does not guarantee presence of arbitrage, as can be seen by the sixth dark blue point in figure 28, for instance.

Now I calibrate to the cleaned data using arpe of implied vols and the same calibration settings as in the last calibration round. The results can be seen in table 4.

Table 4: Results from calibrations to arpe of implied volatilities for the cleaned Nikkei 225 data. When several optima were found, also the best non-degenerate optimum is displayed (for instance indicated by a 3 in contrast to a 1). Time is the total time elapsed for all runs. Evaluations is the total number of objective function evaluations used for all runs.

Model - Method: Optimum	Model parameters	arpe (prices)	arpe (vols)	Time (s)	Evaluations
BNS - Nelder-Mead: 1	{-1.2229, 25.338, 10.323, 0.0141, 0.1147}	0.1521	0.0458	764	1890
BNS - Nelder-Mead: 2	{-1.6590, 5.1256, 15.945, 0.0903, 0.1144}	0.1631	0.0477	764	1890
BNS - CMA-ES: 1	{-0.5211, 0.4266, 20.638, 10.000, 0.1169}	0.0861	0.0295	7671	20995
BNS - CMA-ES: 3	{-0.1503, 0.2350, 0.1226, 0.1763, 0.1163}	0.1179	0.0400	7671	20995
BNS - BOBYQA: 1	{-0.8966, 0.9224, 29.999, 2.5605, 0.1145}	0.1268	0.0430	705	1435
BNS - BOBYQA: 2	{-5.0803, 0.7056, 17.732, 0.2162, 0.1173}	0.1755	0.0515	705	1435
Heston - Nelder-Mead: 1	{-0.5771, 0.8271, 1.4162, 0.0692, 0.1357}	0.0897	0.0287	1719	2806
Heston - Nelder-Mead: 3	{-0.5625, 0.7132, 0.3129, 0.1940, 0.1363}	0.0977	0.0317	1719	2806
Heston - CMA-ES: 1	{-0.5722, 0.8472, 2.0000, 0.0555, 0.1350}	0.0846	0.0277	10637	24689
Heston - CMA-ES: 3	{-0.4583, 0.5209, 0.0000, 0.8342, 0.1414}	0.1181	0.0526	10637	24689
Heston - BOBYQA: 1	{-0.5609, 0.8072, 1.8311, 0.0549, 0.1352}	0.0840	0.0281	1574	2372
Heston - BOBYQA: 3	{-0.5757, 0.7531, 0.0674, 0.8438, 0.1377}	0.1039	0.0328	1574	2372
Heston Jump - Nelder-Mead: 1	{-0.6535, 0.8146, 0.3756, 0.2172, 0.1312, 0.8857, 0.0331, 0.0148}	0.0531	0.0223	3194	5700
Heston Jump - Nelder-Mead: 2	{-0.0566, 0.0954, 1.6429, 0.0200, 0.1202, 0.00107, 0.2206, 0.9994}	0.0991	0.0590	3194	5700
Heston Jump - CMA-ES: 1	{-0.6387, 0.8712, 1.0816, 0.0842, 0.1266, 3.1212, 0.0000, 0.0273}	0.0471	0.0205	10802	30851
Heston Jump - CMA-ES: 2	{-0.6454, 0.8744, 0.9568, 0.0920, 0.1257, 3.8841, 0.0000, 0.0258}	0.0475	0.0205	10802	30851
Heston Jump - BOBYQA: 1	{-0.6134, 0.7689, 1.4548, 0.0674, 0.1313, 0.2938, 0.0263, 0.0376}	0.0468	0.0219	3559	5281
Heston Jump - BOBYQA: 2	{-0.6001, 0.6701, 0.2691, 0.2229, 0.1288, 0.9144, 5.5327E - 5, 0.0394}	0.0491	0.0228	3559	5281
CGMYSA - Nelder-Mead: 1	{5.8828, 32.190, 86.207, 0.4151, 0.4326, 0.6015, 3.5633, 7.8087, 9.7477}	0.0739	0.0325	5523	5819
CGMYSA - Nelder-Mead: 2	{245.63, 31.418, 70.376, -0.4277, -1.0936, 1.1978, 19.889, 39.999, 24.452}	0.4783	0.1381	5523	5819
CGMYSA - CMA-ES: 1	{14.385, 241.32, 400.00, 0.8679, -1.9061, 0.0257, 2.2110, 0.0000, 0.0449}	0.1270	0.0764	17705	80005
CGMYSA - BOBYQA: 1	{3.5791, 19.995, 64.127, 0.3976, 0.3877, 0.6556, 3.4643, 4.5396, 6.8206}	0.0726	0.0354	2002	1787
CGMYSA - BOBYQA: 2	{236.04, 108.86, 317.89, 0.1718, 0.0630, 1.4709, 15.030, 35.855, 43.213}	0.2118	0.1013	2002	1787

The price arpes are better now and the vol arpes are even better than these in general. According to the table, the Heston jump model performs best and the BNS model worst. It is harder to find optima in CGMYSA as usually, but two of the found are quite good. Regarding algorithms, Nelder Mead and BOBYQA work almost equally well. Maybe the BOBYQA dos better since it uses slightly less time and still overall produces better results but it is hard to say since it depends a lot on their settings. CMA-ES still runs for much longer in general so it is quite hard to compare it to the others in that sense, but the results are similar. In the previous calibration to Nikkei prices, the BOBYQA runs for about 2 or 3 times shorter but arguably produces worse results than Nelder-Mead. There, CMA-ES also runs for much longer but overall finds better results.

The fits to option prices are yet again worse than the fits to vols. This could have many explanations. Pricing errors blow up due to low vegas. The mapping between vols and prices could be wrong due to usage of wrong r and q and one could of course calibrate to these if their values are uncertain. The data quality could be poor because changes in S_0 throughout the day are not taken into account.

I plot one market - model vol match for the cleaned Nikkei data: The best CGMYSA model optimum found by BOBYQA. The decent match can be seen in figure 29. My guess is that the point with much lower strikes than the rest is hard to match for all models.

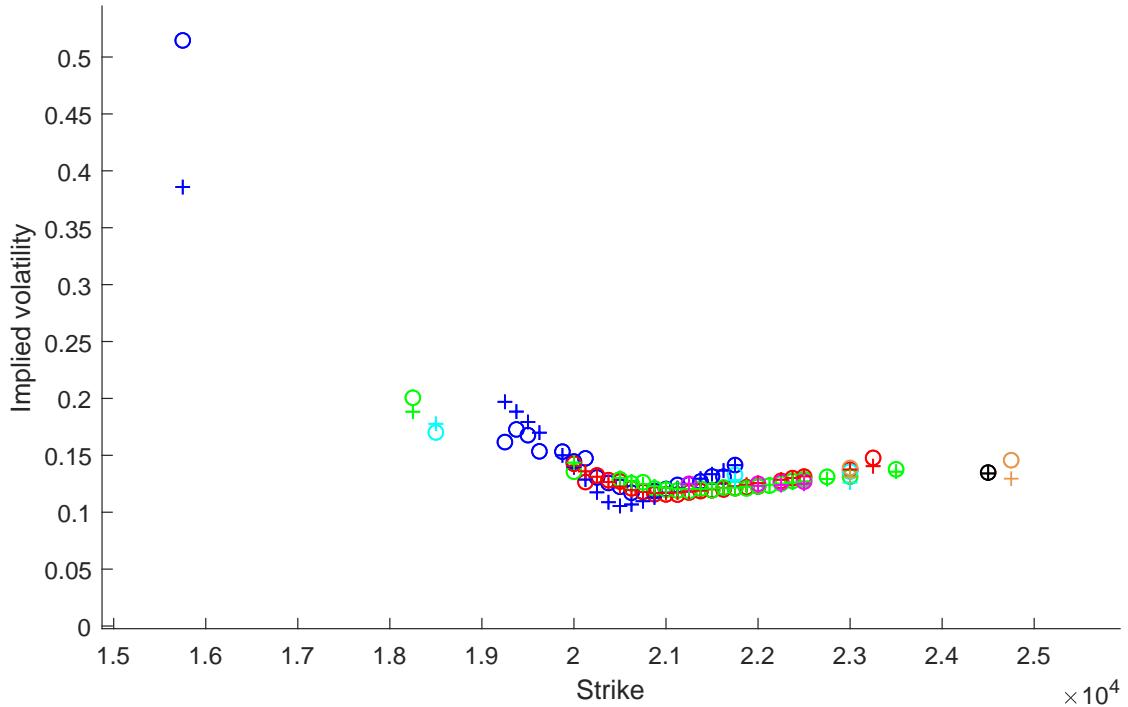


Figure 29: Market-model vol match for CGMYSA - BOBYQA: 1 in table 4. Market and model vols are displayed as circles and plus signs, respectively.

The corresponding surface can be seen in figure 30.

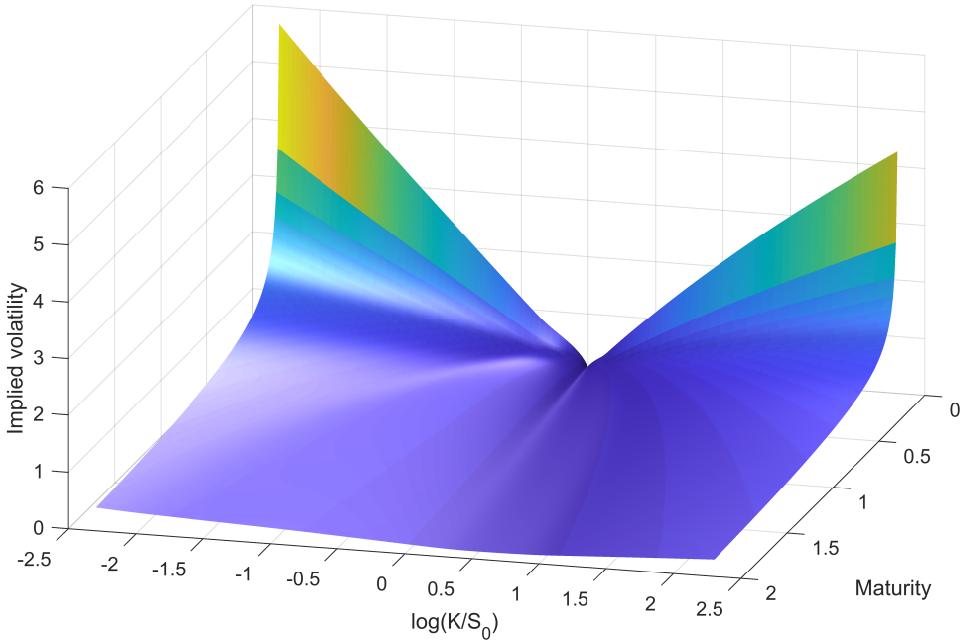


Figure 30: Vol surface from the CGMYSA model using the best parameters found by BOBYQA when calibrating to vols.

So far, there has been no surface produced by the BNS model plotted. In figure 31, such a surface is shown using the best optima found by CMA-ES of table 4.

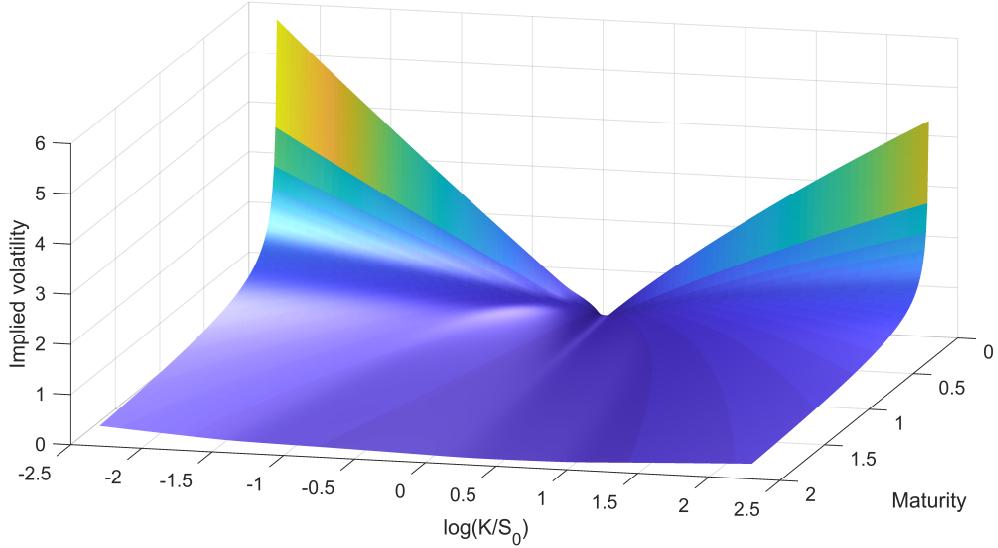


Figure 31: Vol surface from the BNS model using parameters found by CMA-ES.

The difference between these two surfaces has an interesting shape which is displayed in figure 32.

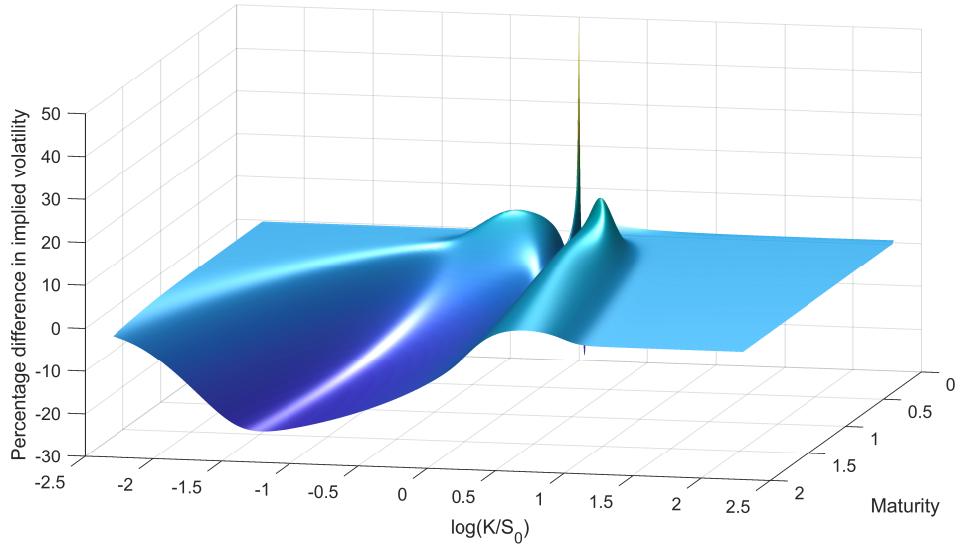


Figure 32: Percentage difference between volatility surfaces in table 4. This is a plot of the vol surfaces of (BNS - CMA-ES: 1) minus (CGMYSA - BOBYQA: 1), both divided by (CGMYSA - BOBYQA: 1) and multiplied by 100.

Figure 32 shows some significant differences. A $\log(K/S_0)$ of -1 is a strike of 37% of S_0 . Vol uncertainties of 20% are then attained for a 2 year maturity option of that strike and this is significant. High quality input data is hence important for accurate surfaces at low strikes.

It is interesting to see how much arbitrage there is in the vol surfaces built from the Nikkei data which itself has arbitrage. Figures 33 and 34 show arbitrage in the BNS and CGMYSA models, respectively.

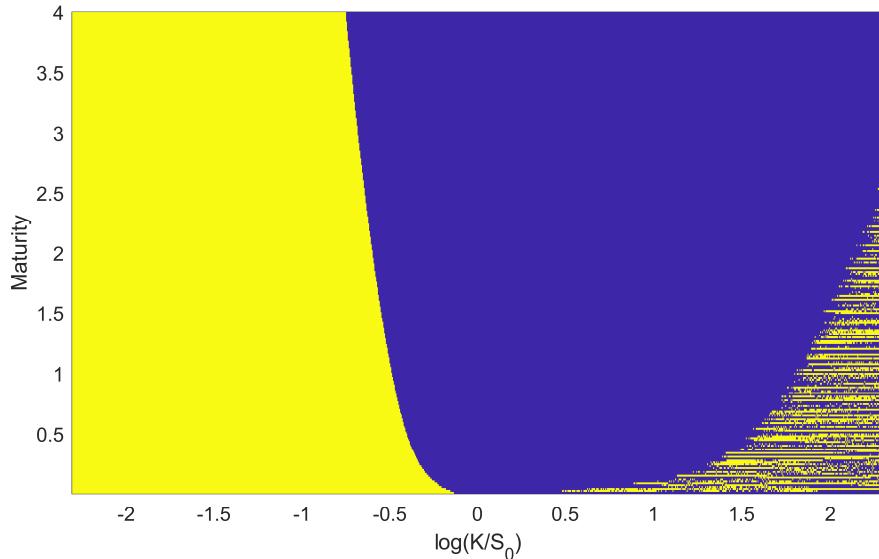


Figure 33: Arbitrage in the vol surface from the BNS model found by CMAES for the cleaned Nikkei data. Yellow points are points which fail the check and hence have static arbitrage. Indigo points are points which pass the check.

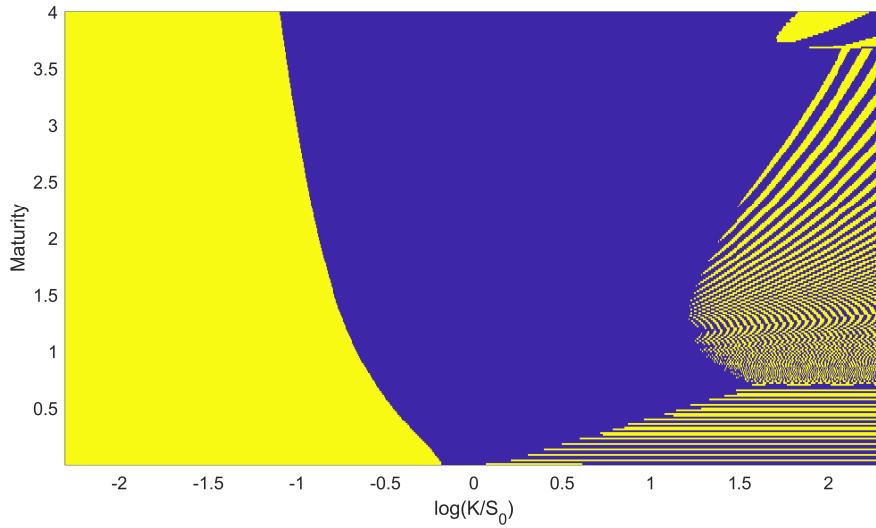


Figure 34: Arbitrage in the vol surface from the CGMYSA model found by BOBYQA for the cleaned Nikkei data. Yellow points are points which fail the check and hence have static arbitrage. Indigo points are points which pass the arbitrage check.

Figure 35 shows a more detailed view of figure 34. Here I have tried to capture which type of arbitrage there is in each point. A difficulty here is that often, if there is one type of arbitrage, there are others simultaneously. I set the arbitrage in each point to the first that was satisfied in this list from left to right: Vertical arbitrage, (A4), calendar arbitrage and finally butterfly arbitrage.

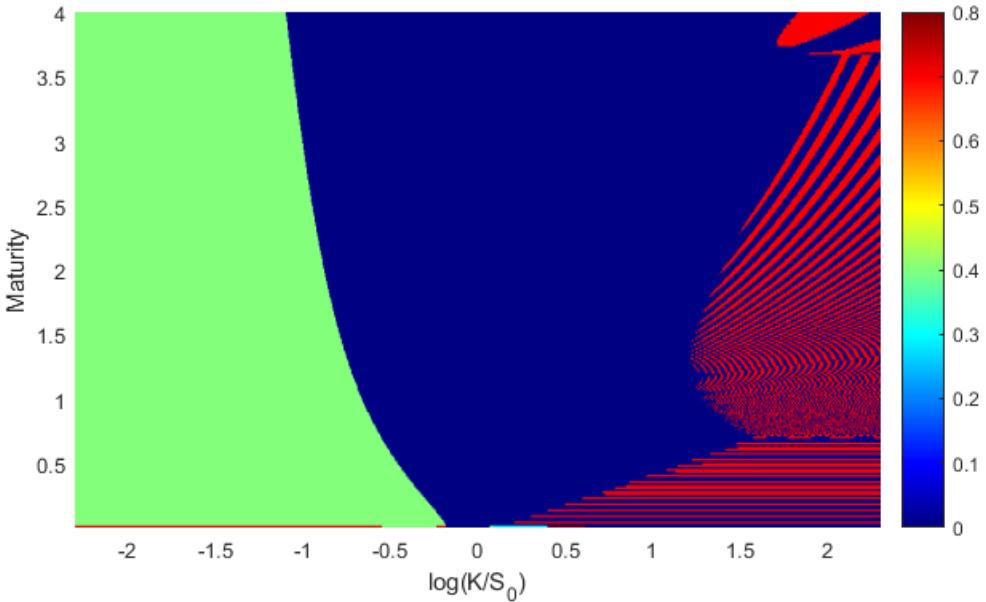


Figure 35: Figure 34 colored in a different way. Arbitrage in the vol surface from the CGMYSA model found by BOBYQA for the cleaned Nikkei data. Dark blue points have no arbitrage. Light blue (0.3) represents vertical spreads. Light green (0.4) represents that option prices are lower than $\max(S_0 - K, 0)$. Red (0.7) represents calendar arbitrage. Purple (dark 0.8) is only butterflies.

It is apparently more arbitrage in the surfaces constructed for the Nikkei data and especially for low strikes where the models try to match the odd point. The arbitrage in the right part in figure 34 of the CGMYSA model constitute a weird pattern. Could this be due to discontinuities

in the price calculations? A plot of one integrand supporting this claim is shown in figure 36.

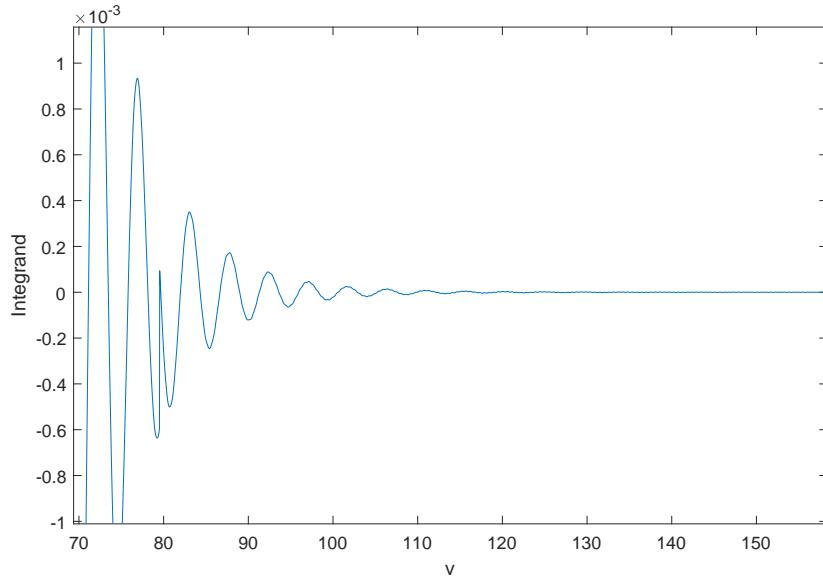


Figure 36: The integrand in the CGMYSA model when using the parameters in BOBYQA: 1 in table 4 and $(K, T) = (5S_0, 2)$. A discontinuity shows up near $v = 80$.

Finally I ran a calibration on the cleaned data with also the odd point at the low strike removed. I do not include the resulting table here, but note that the arpes were significantly reduced. The best being one Heston Jump match with a vol arpe of 0.0149. Removing the point, though, leads to an increase in uncertainties at low strikes.

11.3 On discontinuities from complex branch cut crossings

In this section I want to sort out a topic which sometimes caused disturbances in the results: Presence of discontinuities in the models' characteristic equations - leading to mispricings. Do all 4 models have this problem and how serious is it?

The problem arises occasionally when multiplying the argument of a complex number by a number in (-1,1). The reason is that in such operations, a continuous increase in the input is not matched in the output. Consider for instance the complex numbers $z(\theta) = re^{\theta i}$. These are continuous when increasing θ but after taking the square-root: $\sqrt{z(\theta)} = \sqrt{r}e^{\theta/2}$, there is a discontinuity somewhere. For instance at $\theta = \pi$ or $\theta = 2\pi$ if $\theta \in [-\pi, \pi]$ or $\theta \in [0, 2\pi]$, respectively. In other words, if the branch cut is along the negative or positive real axis. When adding, subtracting, multiplying or dividing complex numbers these problems do not arise since θ is not divided.

11.3.1 Heston and Heston Jump

In [63], they theoretically show that the version of the Heston characteristic function used in this thesis does not produce any discontinuities due to branch crossings. Looking at the characteristic equation of the Heston Jump model,

$$\begin{aligned}\phi(u, t) = & \exp(iu(\log(S_0) + (r - q)t)) \\ & \times \exp(\eta\kappa\theta^{-2}((\kappa - \rho\theta ui - d)t - 2\log((1 - ge^{-dt})/(1 - g)))) \\ & \times \exp(\sigma_0^2\theta^{-2}(\kappa - \rho\theta ui - d)(1 - e^{-dt})/(1 - ge^{-dt})) \\ & \times \exp(-\lambda\mu_J iut + \lambda t((1 + \mu_J)^{iu} \exp(\sigma_J^2(iu - 1)) - 1)),\end{aligned}\tag{73}$$

where

$$\begin{aligned}d &= \sqrt{(\rho\theta ui - \kappa)^2 - \theta^2(-iu - u^2)}, \\ g &= (\kappa - \rho\theta ui - d)/(\kappa - \rho\theta ui + d),\end{aligned}\tag{74}$$

it can be seen that no argument dividing operation is to be performed on the forth exp-term, which is the only extension of the pure Heston model. Hence the Heston Jump model should be free from discontinuities if the pure model is.

11.3.2 BNS

We have not seen any discontinuities in the BNS model thus far. Is it always free of discontinuities? Its characteristic equation is

$$\begin{aligned}\phi(u, t) = & \exp(iu(\log(S_0) + (r - q - a\lambda\rho(b - \rho)^{-1})t)) \\ & \times \exp(-\lambda^{-1}(u^2 + iu)(1 - \exp(-\lambda t))\sigma_0^2/2) \\ & \times \exp\left(\frac{a}{b - f_2(u)} \left(b \log\left(\frac{b - f_1(u)}{b - iu\rho}\right) + f_2\lambda t\right)\right)\end{aligned}\tag{75}$$

with

$$\begin{aligned}f_1(u) &= iu\rho - \lambda^{-1}(u^2 + iu)(1 - \exp(-\lambda t))/2 \\ f_2(u) &= iu\rho - \lambda^{-1}(u^2 + iu)/2\end{aligned}\tag{76}$$

There is only one argument dividing function in it and that is the complex logarithm. Remember that in the FFT-approach, the input to the characteristic equation is $u = v - (\alpha + 1)i$, where $v > 0$. Then the argument to the logarithm can be rewritten

$$\frac{[b - f_1(v - (\alpha + 1)i)][b - \rho(\alpha + 1) + iv\rho]}{(b - \rho(\alpha + 1))^2 + v^2\rho^2}.$$

In the expression above, I denote the numerator by E. The branch cut for the logarithm is set to the negative real axis for the programs I have used (Java and Matlab). Then a discontinuity arises if the imaginary part in the numerator:

$$Im(E) = Im(f_1)(\rho(\alpha + 1) - b) - \rho v Re(f_1) + b\rho v$$

crosses 0 while the real part in the numerator:

$$Re(E) = b^2 - b\rho(\alpha + 1) + (\rho(\alpha + 1) - b)Re(f_1) + \rho v Im(f_1)$$

is negative. Inserting

$$Re(f_1(v - (\alpha + 1)i)) = (\alpha + 1)\rho - (v^2 - (\alpha + 1)^2 + (\alpha + 1)) \left[\frac{1 - e^{-\lambda t}}{2\lambda} \right] \quad (77)$$

$$Im(f_1(v - (\alpha + 1)i)) = v\rho - v(1 - 2(\alpha + 1)) \left[\frac{1 - e^{-\lambda t}}{2\lambda} \right] \quad (78)$$

in the above expression gives that, apart from a positive real factor, the input in the logarithm is

$$\begin{aligned} E = & \rho v^2 \left(\rho - (1 - 2(\alpha + 1)) \left[\frac{1 - e^{-\lambda t}}{2\lambda} \right] \right) + b^2 - b\rho(\alpha + 1) \\ & + (\rho(\alpha + 1) - b) \left((\alpha + 1)\rho - (v^2 - (\alpha + 1)^2 + (\alpha + 1)) \left[\frac{1 - e^{-\lambda t}}{2\lambda} \right] \right) \\ & + iv \left((\rho(\alpha + 1) - b) \left(\rho - (1 - 2(\alpha + 1)) \left[\frac{1 - e^{-\lambda t}}{2\lambda} \right] \right) + b\rho \right) \\ & - i\rho v \left((\alpha + 1)\rho - (v^2 - (\alpha + 1)^2 + (\alpha + 1)) \left[\frac{1 - e^{-\lambda t}}{2\lambda} \right] \right) \end{aligned}$$

Let's now evaluate if this crosses the negative real axis. In such a case, $Im(E)$ has to cross 0. This imaginary part can be expressed in terms of orders of v :

$$\begin{aligned} Im(E) = & v^1 \left((\rho(\alpha + 1)^2 + b - 2b(\alpha + 1)) \left[\frac{1 - e^{-\lambda t}}{2\lambda} \right] \right) + \\ & + v^3 \left(\rho \left[\frac{1 - e^{-\lambda t}}{2\lambda} \right] \right) \end{aligned}$$

The (one and only) crossing point of 0 is then given by the solution to the equation:

$$\rho v^2 + (\rho(\alpha + 1)^2 + b - 2b(\alpha + 1)) = 0 \quad (79)$$

which is solved by

$$v_{crossing} = \sqrt{\frac{b}{\rho}(2\alpha + 1) - (\alpha + 1)^2} \quad (80)$$

In equation 80 it can be seen that for there to be a crossing, $b > \rho$ and $b\rho > 0$. This combined with that from the BNS model, $b > 0$ gives $b > \rho > 0$. The question now is whether this sole chance of a discontinuity could hold when simultaneously, $Re(E) < 0$. $Re(E)$ is:

$$\begin{aligned} Re(E) = & v^0 \left(\rho^2(\alpha + 1)^2 - 2b\rho(\alpha + 1) + b^2 + [\rho(\alpha + 1) - b][(\alpha + 1)^2 - (\alpha + 1)] \left[\frac{1 - e^{-\lambda t}}{2\lambda} \right] \right) + \\ & + v^2 \left(\rho^2 + (\rho\alpha + b) \left[\frac{1 - e^{-\lambda t}}{2\lambda} \right] \right) \end{aligned}$$

Setting $v = v_{crossing}$ in $Re(E)$ gives

$$Re(E_{crossing}) = \left[\frac{1 - e^{-\lambda t}}{2\lambda} \right] \left(b \left(\frac{b}{\rho}(2\alpha + 1) - (2\alpha + 1) \right) \right) + b(b - \rho) \quad (81)$$

The existence of a crossing point requirement, $b > \rho > 0$, forces both terms in the real part to be positive. Hence, crossing the negative part of the real line is not possible. Therefore there are no discontinuities for a branch cut along that line.

11.3.3 CGMYSA

We have seen that discontinuities can occur in the CGMYSA model. I here hint at how it can be seen from the characteristic equation of the model, which is the following,

$$\phi_{CGMYSA}(u) = \exp(iu(\log(S_0)) + (r-q)t) \times \frac{\phi_{CIR}(-i\phi_{CGMY}(u; 1, G, M, Y_p, Y_n, \zeta), t; \kappa, \eta, \lambda, C)}{[\phi_{CIR}(-i\phi_{CGMY}(-i; 1, G, M, Y_p, Y_n, \zeta), t; \kappa, \eta, \lambda, C)]^{iu}}$$

with

$$\begin{aligned}\phi_{CGMY}(u; C_p, G, M, Y_p, Y_n, \zeta) &= C_p [\Gamma(-Y_p) ((M - iu)^{Y_p} - M^{Y_p}) + \zeta \Gamma(-Y_n) ((G + iu)^{Y_n} - G^{Y_n})] \\ \phi_{CIR}(u, t; \kappa, \eta, \lambda, y_0) &= \frac{\exp(\kappa^2 \eta t / \lambda^2) \exp(2y_0 iu / (\kappa + \gamma \coth(\gamma t/2)))}{(\cosh(\gamma t/2) + \kappa \sinh(\gamma t/2) / \gamma)^{2\kappa\eta/\lambda^2}} \\ \gamma &= \sqrt{\kappa^2 - 2\lambda^2 iu}.\end{aligned}$$

The exponent $2\kappa\eta/\lambda^2$ in the denominator of ϕ_{CIR} could be in (0,1). If then $\cosh(\gamma t/2) + \kappa \sinh(\gamma t/2) / \gamma$ crosses the negative real axis, there will be discontinuities. I skip doing a theoretical exhaustion of all cases here and note instead that for one such parameter set; $C_p = 1$, $G = 3$, $M = 3$, $Y_p = -2$, $Y_n = -2$, $\zeta = 0.001$, $\kappa = 1$, $\eta = 2$, $\lambda = 8$, $t = 2$ and $\alpha = 1$, there are problems. Using these parameters, figure 37 shows an example of discontinuities arising. It plots the base of the denominator of the numerator of $\phi_{CGMYSA}(u)$ to the left and to the right; the denominator of the numerator of $\phi_{CGMYSA}(u)$. In other words, to the left: $\cosh(\gamma t/2) + \kappa \sinh(\gamma t/2) / \gamma$ and to the right: This raised to the power of $2\kappa\eta/\lambda^2$.

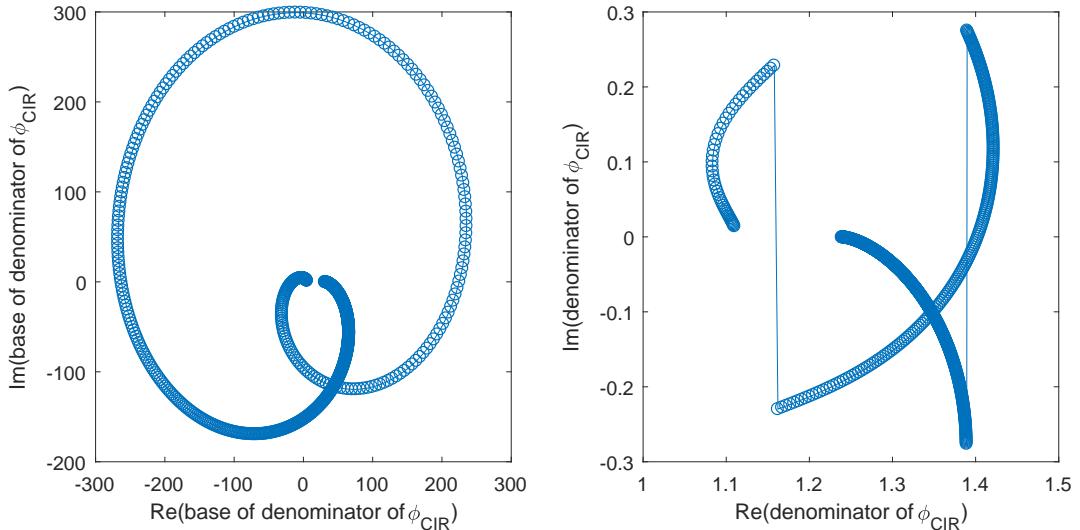


Figure 37: The base of the exponential term is shown to the left. Its imaginary part crosses 0 twice in the figure when the real part is negative. The exponentiated base is shown to the right. It displays two discontinuities corresponding to the two crossings in the left subplot.

12 Concluding discussion

I want to end with a concluding discussion on the results and the choices made in the thesis.

The implemented methods have produced surfaces with good fits to data of high quality and decent fits to data of lesser quality. In my opinion, fits are good for high-quality data if they have vol arps of less than 1%, because we have seen that if arps are above that, option prices can differ significantly within the input data region. Worth noting is that a perfect fit is perhaps not the goal for data of lesser quality, but instead a smooth surface which is not so dependent on the calibration input and which is free from arbitrage opportunities.

Regarding models, the CGMYSA was slightly more difficult to calibrate compared to the others and it had problems with discontinuities but the found best matches often produced fits which were just as good as the fits of the other models. I would say that the BNS overall had a slightly worse performance than the Heston and Heston Jump models.

There were no particular problems with calibrating to implied volatilities and it was found to not extend the required computational time significantly. The relative differences of vol surfaces produced from the FFT-approach, both within a model using differing parameter sets and between models, are substantial at times. Even (and actually often in particular) for non-extreme strikes and maturities; that is for options which are likely traded. The surfaces were found to be free of static arbitrage for most normal strikes and maturities, if the input data was good. If it was not, there could be a lot of arbitrage for not so far out-of-the-money options. Hence the models do not generate arbitrage-free surfaces for all parameter sets.

The obtained fits combined with the differences in between calibrated surfaces, the presence of arbitrage and potential problems with discontinuous pricing integrands leads me to conclude that the FFT-pricing method can be used successfully but it should be done with care and meticulousness to circumvent the pitfalls. Even if the calibrations are performed perfectly on clean data I suspect that there will be many well-fitting optima with significantly dissimilar implied S_t -predictions and vol surfaces. It seems simpler to just go and model the surface directly but I'm sure this has its disadvantages too. I wonder how the FFT-approach would perform compared to someone familiar with the subject drawing the smiles by hand. All in all, there is a lot of uncertainty in vol surface construction so as much reliable information as possible is preferably utilized.

A side note is that I'm quite impressed by that the implied vols found from the Nikkei data were not less uniform. The tick size for the options is one and some of the prices for high-strike options were of a comparable size; 1,1,3,6, for instance. This had me expect highly non-linear smiles. The tick size combined with the low vega could potentially explain that some low arpe(vols) give high arpe(prices) for the Nikkei data: A slight vol increase there would double the prices.

I would like to point out some of the things I did not investigate fully in this thesis, either because of time constraints or lack of access:

- What if puts are included and used for calibration?
- What if I had gotten the corresponding S_0 for each trade?
- What if I had exotic option prices to calibrate to?
- What if r and q are wrong? These could be included as unknown parameters in the calibration.
- What if some pricing integrands did not converge before the upper integration limit was reached?
- What if the algorithm settings are inappropriate?
- What if $\psi_T(v)$ and $\gamma_T(v)$ produce very different results for some (K,T) .

I did ponder these but I was mostly looking for the overall qualities of the approach and the models which I think I successfully uncovered.

References

- [1] Bank for International Settlements - Monetary and Economic Department. "Statistical release: OTC derivatives statistics at end-June 2016", Annex A. (2016). Available at https://www.bis.org/publ/otc_hy1611.pdf.
- [2] World Development Indicators database, World Bank. "Gross domestic product 2016". 17 April 2017. (2017). Available at <http://databank.worldbank.org/data/download/GDP.pdf>.
- [3] R. Cont, J. da Fonseca, "Dynamics of implied volatility surfaces". Quantitative finance, vol. 2. (2002). IOP Publishing Ltd. DOI: <https://doi.org/10.1088/1469-7688/2/1/304>.
- [4] R. Cont, J. da Fonseca, "Deformation of implied volatility surfaces: an empirical analysis". Takayasu H. (eds) Empirical Science of Financial Fluctuations. (2002). Springer, Tokyo. DOI: https://doi.org/10.1007/978-4-431-66993-7_25.
- [5] F. A. Longstaff, "Option Pricing and the Martingale Restriction ". The Review of Financial Studies, vol. 8, issue 4, pp. 1091–1124. 1 October (1995). DOI: <https://doi.org/10.1093/rfs/8.4.1091>.
- [6] R. Hafner, M. Wallmeier, "The Dynamics of DAX Implied Volatilities". International Quarterly Journal of Finance, vol. 1, pp. 1-27. (2000). Available at http://www.math.uni-frankfurt.de/~stoch/hafner_wallmeier.pdf.
- [7] M. R. Fengler, "Arbitrage-Free Smoothing of the Implied Volatility Surface". SFB 649 Discussion Paper (2005)-019. Available at <https://core.ac.uk/download/pdf/6978470.pdf>.
- [8] M. Fengler, W. Härdle, E. Mammen, "A Dynamic Semiparametric Factor Model for Implied Volatility String Dynamics". SFB 649 Discussion Paper (2005)-020. Available at <http://sfb649.wiwi.hu-berlin.de/papers/pdf/SFB649DP2005-020.pdf>.
- [9] I. Pena, G. Rubio, G. Serna, "Why do we smile? On the determinants of the implied volatility function". Journal of Banking & Finance, vol. 23, pp. 1151-1179. (1999). Available at https://e-archivo.uc3m.es/bitstream/handle/10016/7085/why_peña_JBF_1999_ps.pdf;jsessionid=E1E90CE29CB8FD3C75B33FEB64076C29?sequence=1.
- [10] T. Daglish, J. Hull, W. Suo, "Volatility Surface: Theory, Rules of Thumb, and Empirical Evidence". Quantitative Finance, vol 7, pp. 507-524. (2007). DOI: <https://doi.org/10.1080/14697680601087883>.
- [11] J. Gatheral, "Rough volatility: An overview". Global Derivatives, Barcelona (May 2017). Available at <http://mfe.baruch.cuny.edu/wp-content/uploads/2017/05/RoughVolatilityBarcelona2017.pdf>. Same as: "Presentation at the LSE Risk and Stochastics Conference 2017 by Jim Gatheral", Baruch College.
- [12] J. C. Hull. Options, Futures, and Other Derivatives. no. 8, pp. 218-231. Pearson Education, (2012).
- [13] R. C. Merton, "Theory of Rational Option Pricing". The Bell Journal of Economics and Management Science, vol. 4, no. 1, pp. 141-183. The RAND Corporation. (Spring, 1973). Stable url: <http://www.jstor.org/stable/3003143>.
- [14] D. Reiswich, "A Note on Standard No-Arbitrage Conditions". SSRN Electronic Journal, (2010). DOI: <https://doi.org/10.2139/ssrn.1619844>.
- [15] P. Carr, "Implied Vol Constraints". Bloomberg, working paper. (2004). Available at <http://docplayer.net/28019290-Implied-vol-constraints.html>.
- [16] R. W. Lee, "Implied volatility: Statics, dynamics, and probabilistic interpretation". Recent Advances in Applied Probability, pp. 241–268. (2005). Springer, Boston, MA. DOI: https://doi.org/10.1007/0-387-23394-6_11.
- [17] R. W. Lee, "The Moment Formula for Implied Volatility at Extreme Strikes". Mathematical Finance, vol. 14, no. 3, pp. 469-480. (July 2004). Available at <https://ssrn.com/abstract=558885>.

- [18] P. Carr, D. B. Madan, "A note on sufficient conditions for no arbitrage". Finance Research Letters, vol. 2, issue 3, pp. 125-130. (2005). Available at: <http://engineering.nyu.edu/files/FRLarticle.pdf>.
- [19] H. G. Kellerer, "Markov-Komposition und eine Anwendung auf Martingale". Mathematische Annalen, vol. 198, pp. 99–122. (1972). Available at <https://eudml.org/doc/162296>.
- [20] M. Roper, "Arbitrage free implied volatility surfaces", preprint, (2010). School of Mathematics and Statistics, The University of Sydney, Sydney, New South Wales, Australia. Available at <http://talus.maths.usyd.edu.au/u/pubs/publist/preprints/2010/roper-9.pdf>.
- [21] F. Mercurio, "No-Arbitrage Conditions for a Finite Options System". Banca IMI. Available at <http://www.fabiomercurio.it/NoArbitrage.pdf>.
- [22] M. H. A. Davis, D. G. Hobson, "The range of traded option prices". Mathematical Finance, vol. 17, no. 1, pp. 1-14. (2007). DOI: <https://doi.org/10.1111/j.1467-9965.2007.00291.x>.
- [23] E. Derman, M. B. Miller, D. Park. The Volatility Smile, p. 8. (2016).
DOI: <https://doi.org/10.1002/9781119289258>.
- [24] ManWo Ng. Option Pricing via the FFT and its Application to Calibration. M.Sc. Thesis. Delft University of Technology, Delft, The Netherlands. (2005). Available at <http://ta.twi.tudelft.nl/mf/users/oosterle/oosterlee/ng.pdf>.
- [25] P. Carr, D. B. Madan, "Option valuation using the fast Fourier transform". The Journal of Computational Finance, vol. 2. (1999). Available at <http://engineering.nyu.edu/files/jcfpub.pdf>.
- [26] R. Cont, J. Da Fonseca, V. Durrleman, "Stochastic models of implied volatility surfaces". Economic notes by Banca Monte dei Paschi di Siena SpA, vol. 3.1. no. 2, pp.361-377. (2002). Available at <http://www.cmap.polytechnique.fr/~rama/papers/verona.pdf>.
- [27] E. Ayache, P. Henrotte, S. Nassar, X. Wang, "Can anyone solve the smile problem?". Wilmott Magazine 2004(1), pp. 78-96. (2004). Available at <http://citeserx.ist.psu.edu/viewdoc/summary?doi=10.1.1.232.7835>.
- [28] E. Ayache, "Can Anyone Solve The Smile Problem? A Post-Scriptum". (2005). Available at http://www.ito33.com/sites/default/files/articles/0501_ayache.pdf.
- [29] N. Kahalé, "An arbitrage-free interpolation of volatilities". RISK 17(5), pp. 102–106. (2004). Available at https://www.researchgate.net/publication/228872089_An_Arbitrage-free_Interpolation_of_Volatilities.
- [30] D.A. Bloch, C.A Coello, "Smiling at evolution". Applied Soft Computing, vol. 11, issue 8, pp. 5724-5734. (2011). DOI: <https://doi.org/10.1016/j.asoc.2011.03.016>.
- [31] L. B. G. Andersen, R. Brotherton-Ratcliffe, "The equity option volatility smile: an implicit finite-difference approach". Journal of Computational Finance, vol. 1, no. 2, pp. 5-37. (1997). DOI: <https://doi.org/10.21314/JCF.1997.009>.
- [32] Lorenzo Bergomi, "Smile Dynamics I". SSRN Electronic Journal. (2004). DOI: <https://dx.doi.org/10.2139/ssrn.1493294>.
- [33] W. Schoutens, E. Simons, J. Tistaert, "A perfect calibration! Now what?". Wilmott Magazine, March 2004: pp 66–78. (2004). Available at <https://perswww.kuleuven.be/~u0009713/ScSiTi03.pdf>.
- [34] C. Homescu, "Implied volatility surface: construction methodologies and characteristics". SSRN Electronic Journal. (2011). DOI: <https://doi.org/10.2139/ssrn.1882567>.
- [35] S. Benaim, M. Dodgson, D Kainth, "An arbitrage-free method for smile extrapolation". Technical report, Royal Bank of Scotland. (2008). Available at http://www.quarchome.org/risktailspaper_v5.pdf.

- [36] R. Crisóstomo, "An Analysis of the Heston Stochastic Volatility Model: Implementation and Calibration using Matlab". CNMV Working papers, no. 58. (2014). Available at <https://arxiv.org/pdf/1502.02963.pdf>.
- [37] A. Medvedev, O. Scaillet, "Approximation and Calibration of Short-Term Implied Volatilities Under Jump-Diffusion Stochastic Volatility". Swiss Finance Institute Research Paper, no. 06-8. (2006). DOI: <http://dx.doi.org/10.2139/ssrn.910212>.
- [38] D. Bates, "Jumps and Stochastic Volatility: Exchange Rates Processes Implicit in Deutsche Mark Options". Review of Financial Studies, vol. 9, pp. 69-107. (1996). Available at https://deriscope.com/docs/Bates_1996.pdf.
- [39] G. Bakshi, C. Cao, Z. Chen (1997), "Empirical Performance of Alternative Option Pricing models". The Journal of Finance, vol. 52, no. 5, pp. 2003-2049. (1997). DOI: <http://dx.doi.org/10.1111/j.1540-6261.1997.tb02749.x>.
- [40] M. Chernov, R. Gallant, E. Ghysels, G. Tauchen, "Alternative Models for Stock Price Dynamics". Journal of Econometrics, vol. 116, issue 1-2, pp. 225-257. (2003). Available at <https://pdfs.semanticscholar.org/4b53/093bdab542f6f1fd5ef2fbe7efab87f5e20b.pdf>.
- [41] T. Andersen, L. Benzoni, J. Lund, "An Empirical Investigation of continuous-time equity return models". Journal of Finance, vol. 57, issue 3, pp. 1239-1284. (2002). Available at <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.555.5472&rep=rep1&type=pdf>.
- [42] J. Pan, "The Jump Risk Premia Implicit in Options: Evidence from an Integrated Time Series Study". Journal of Financial Economics, vol. 63, issue 1, pp. 3-50. (2002). Available at <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.22.5088&rep=rep1&type=pdf>.
- [43] D. Bates, "The crash of 87: was it expected? The evidence from option markets". Journal of Finance, vol. 46, pp. 1009-1044. (2000). Available at http://pages.stern.nyu.edu/~dbackus/GE_asset_pricing/disasters/Bates%20crash%20JF%2091.PDF.
- [44] S. Galluccio, Y. Le Cam, "Implied Calibration and Moments Asymptotics in Stochastic Volatility Jump Diffusion Models". Preprints from the Probability and Financial Mathematics team, Evry University. (2007). Available at <https://www.maths.univ-evry.fr/prepubli/251.pdf>.
- [45] B. Eraker, M. Johannes, N. Polson, "The Impact of Jumps in Volatility and Returns". Journal of Finance, vol. 58, issue 3, pp. 1269-1300. (2003). Available at <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.409.4349&rep=rep1&type=pdf>.
- [46] P. Carr, H. Geman, D. B. Yor, M. Yor, "The fine structure of asset returns: an empirical investigation". The Journal of Business, vol. 75, pp. 305-332. Available at <http://eprints.bbk.ac.uk/archive/00000411>.
- [47] A. Lipton, "The vol smile problem". Masterclass with Deutsche Bank. Risk Magazine, vol. 15, issue. 2, pp. 61-66. (2002). Available at http://www.math.ku.dk/~rolf/Lipton_VolSmileProblem.pdf.
- [48] O. E. Barndorff-Nielsen, E. Nicolato, N. Shephard, "Some recent developments in stochastic volatility modelling". Quantitative finance, vol. 2, issue 1, pp. 11-23. (2002). Available at <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.583.7335&rep=rep1&type=pdf>.
- [49] M. Jex, R. Henderson, D. Wang, "Pricing Exotics under the Smile". Risk Magazine, November 1999. (1999). J.P.Morgan Securities Inc. JPMorgan London Derivatives Research. Available at <http://www.smartquant.com/references/OptionPricing/option14.pdf>.
- [50] J. Guyon, "Path-Dependent Volatility". Risk magazine, October 2014. (2014). DOI: <http://dx.doi.org/10.2139/ssrn.2425048>.
- [51] C. Corrado, T. Su, "Skewness and kurtosis in S&P 500 index returns implied by option prices". Journal of Financial Research, vol. 19, no. 2, pp. 175-192. (1996). Available at <https://moya.bus.miami.edu/~tsu/jfr1996.pdf>.

- [52] S. Gurrieri, "A class of term structures for SVI Implied Volatility". SSRN Electronic Journal. (2010). DOI: <https://dx.doi.org/10.2139/ssrn.1779463>.
- [53] J. Gatheral, A. Jacquier, "Arbitrage-free SVI volatility surfaces". Quantitative Finance, vol. 14, no. 1, pp. 59–71. (2014). DOI: <https://dx.doi.org/10.2139/ssrn.2033323>.
- [54] G. Guo, A. Jacquier, C. Martini, L. Neufcourt, "Generalised Arbitrage-Free SVI Volatility Surfaces". SIAM Journal on Financial Mathematics, vol. 7, issue 1, pp. 619–641. (2016). DOI: <http://dx.doi.org/10.2139/ssrn.2167263>.
- [55] J. Gatheral, T. Jaisson, M. Rosenbaum, "Volatility is rough". Quantitative Finance, forthcoming. (2014). Available at <https://arxiv.org/abs/1410.3394>.
- [56] C. Bayer, P. Friz and J. Gatheral, "Pricing under rough volatility". Quantitative Finance, vol. 16, issue 6. (2016). DOI: <https://doi.org/10.1080/14697688.2015.1099717>.
- [57] P. Carr, L. Wu, "Analyzing Volatility Risk and Risk Premium in Option Contracts: A New Theory". Journal of Financial Economics, vol. 120, pp. 1–20. (2016). Available at <http://engineering.nyu.edu/files/JFE2016.pdf>.
- [58] M. Andersson. Models for the Dynamics of Implied Volatility Surfaces. M. sc. thesis. Chalmers University of Technology. (2014). Available at <http://www.math.chalmers.se/~palbin/MartinAndersson.pdf>.
- [59] P. Carr, H. Geman, D. B. Madan, M. Yor, "Stochastic Volatility for Lévy Processes". Mathematical finance, vol. 13, issue 3, pp. 345–382. (2003). Available at <https://pdfs.semanticscholar.org/7edb/0a320ee89e4319eb47d8ee2ba32686241b0b.pdf>.
- [60] J. C. Lagarias, J. A. Reeds, M. H. Wright, P. E. Wright, "Convergence Properties of the Nelder-Mead Algorithm in Low Dimensions". SIAM J. Optim., vol. 9, no. 1, pp. 112–147. (1998). Available at <http://www.jasoncantarella.com/downloads/SJE000112.pdf>.
- [61] M. J. D. Powell, "The BOBYQA algorithm for bound constrained optimization without derivatives". Technical Report, DAMTP 2009/NA06. Centre for Mathematical Sciences, University of Cambridge. (2009). Available at http://www.damtp.cam.ac.uk/user/.na/NA_papers/NA2009_06.pdf.
- [62] N. Hansen, "The CMA Evolution Strategy: A Tutorial". (2009). Available at <https://www.lri.fr/~hansen/cmatutorial.pdf>.
- [63] H. Albrecher, P. Mayer, W. Schoutens, J. Tistaert, "The little Heston trap". Wilmott magazine, january 2007. (2007). Available at <https://perswww.kuleuven.be/~u0009713/HestonTrap.pdf>.
- [64] Japan Exchange Group, Statistics (Derivatives), Daily report on September 27th, 2017. (2017). Available at: <http://www.jpx.co.jp/english/markets/statistics-derivatives/daily/>.

A EURO STOXX 50 call option data

Table 5: Implied Volatility Surface of EURO STOXX 50 on Oct 7th 2003, [33]. Maturities are expressed as fractions of a year.

Maturity Strike \ \diagdown	0.0361	0.2000	1.1944	2.1916	4.2056	5.1639
1081.82			0.3804	0.3451	0.3150	0.3137
1212.12			0.3667	0.3350	0.3082	0.3073
1272.73			0.3603	0.3303	0.3050	0.3043
1514.24			0.3348	0.3116	0.2920	0.2921
1555.15			0.3305	0.3084	0.2899	0.2901
1870.30		0.3105	0.2973	0.2840	0.2730	0.2742
1900.00		0.3076	0.2946	0.2817	0.2714	0.2727
2000.00		0.2976	0.2858	0.2739	0.2660	0.2676
2100.00	0.3175	0.2877	0.2775	0.2672	0.2615	0.2634
2178.18	0.3030	0.2800	0.2709	0.2619	0.2580	0.2600
2200.00	0.2990	0.2778	0.2691	0.2604	0.2570	0.2591
2300.00	0.2800	0.2678	0.2608	0.2536	0.2525	0.2548
2400.00	0.2650	0.2580	0.2524	0.2468	0.2480	0.2505
2499.76	0.2472	0.2493	0.2446	0.2400	0.2435	0.2463
2500.00	0.2471	0.2493	0.2446	0.2400	0.2435	0.2463
2600.00		0.2405	0.2381	0.2358	0.2397	0.2426
2800.00			0.2251	0.2273	0.2322	0.2354
2822.73			0.2240	0.2263	0.2313	0.2346
2870.83			0.2213	0.2242	0.2295	0.2328
2900.00			0.2198	0.2230	0.2288	0.2321
3000.00			0.2148	0.2195	0.2263	0.2296
3153.64			0.2113	0.2141	0.2224	0.2258
3200.00			0.2103	0.2125	0.2212	0.2246
3360.00			0.2069	0.2065	0.2172	0.2206
3400.00			0.2060	0.2050	0.2162	0.2196
3600.00				0.1975	0.2112	0.2148
3626.79				0.1972	0.2105	0.2142
3700.00				0.1964	0.2086	0.2124
3800.00				0.1953	0.2059	0.2099
4000.00				0.1931	0.2006	0.2050
4070.00					0.1988	0.2032
4170.81					0.1961	0.2008
4714.83			1		0.1910	0.1957
4990.91					0.1904	0.1949
5000.00					0.1903	0.1949
5440.18						0.1938

B Nikkei 225 call option data

Table 6: Closing prices for all call options traded on Nikkei 225 on September 27th, 2017, [64]. The index closed at 20267 on this day. Maturities are expressed as fractions of a year.

Maturity Strike	0.0384	0.115	0.1918	0.288	0.3644	0.4411	0.69
15750	4520	4510					
18250			2110				
18500						2020	
19250		1030					
19375		920					
19500		805					
19625		685					
19750		625					
19875		485					
20000	385	535	620				
20125	310	420					
20250	215	370					
20375	150	300					
20500	100	245	350				
20625	60	185	290				
20750	38	140	250				
20875	23	105	195				
21000	14	80	160				
21125	9	60	130				
21250	5	47	105	195			
21375	3	36	90				
21500	2	27	70				
21625	1	20	60				
21750	1	15	47		220		
21875		11	37				
22000		8	31	75			
22125		7	25				
22250		5	21	55			
22375		4	18				
22500		3	15	39			
22750			10				
23000		1	6		70	170	
23250			1				
23500				3			
24500					6		
24750						55	