Calibrating an option pricing model under regime-switching volatility

Johan Kilander*
Master Thesis
Stockholm School of Economics
Department of Finance

27th February 2007

Abstract

A Black-Scholes market is considered in which asset prices are modelled by a geometric Brownian motion with regime-switching volatility. The regime-switching allows the volatility to jump randomly amongst a finite number of volatility states. Pricing equations of European options are derived and the equations are solved numerically. The model is calibrated for two, three and four volatility states to observed market prices of call options on the OMXS30 index during the period August 2003 to August 2006. The findings show that two volatility states are sufficient to replicate market prices with a high degree of accuracy. Mispricings are found to be considerably smaller under a regime-switching model than under the traditional Black-Scholes model.

Acknowledgements I would like to thank my tutor Professor Tomas Björk at the Stockholm School of Economics for valuable advice, comments and criticism on a variety of issues. I am also very grateful to Ola Hammarlid, Head of Quantitative Analytics - Equities at Swedbank Markets for suggestions on how to improve the final code.

*19753@student.hhs.se

Contents

1 Introduction **1**

2 Motivation for stochastic volatility **2**

3 The model **5**
3.1 Regime-switching volatility defined 5
3.2 Continuous time regime-switching volatility 7
3.3 Rationale for regime-switching volatility 10

4 Option pricing **11**
4.1 Option pricing in general 11
4.2 Option pricing under regime-switching volatility 12
4.2.1 Selecting the risk adjusted measure 12
4.2.2 Derivation of the pricing equations 14

5 Data and method **18**

6 Empirical Results **20**
6.1 Parameters 20
6.2 In-sample fit 22
6.3 Out-of-sample performance 24
6.4 Stability 25

7 Conclusions **28**

1 Introduction

In 1973, Black and Scholes (1973) [6] published a path-breaking work on the pricing of options. To derive their famous option pricing formula Black and Scholes assumed that the underlying security S_t followed a geometric Brownian motion with a constant drift μ and volatility σ

$$dS_t = \mu S_t dt + \sigma S_t dW_t.$$

It is widely recognized that the Black-Scholes model does not provide a realistic description of real-world stock price dynamics and it is not consistent with how options are priced in the market.[1] In particular, asset returns are much more heavy-tailed than suggested by the normal distribution and volatility tends to vary quite a lot over time. To obtain more realistic descriptions of asset price dynamics, emerging interest have focused on so called stochastic volatility models, i.e. models where the volatility of the risky asset changes from time to time in a random fashion. In this report the volatility will be modelled by a *regime-switching* model in which the volatility is allowed to jump randomly amongst a finite number of volatility states. The issue of pricing European options in this framework will be studied in detail.

The purpose of this thesis is *to calibrate a continuous time version of a regime-switching volatility model to observed option prices.* The findings will contribute to previous research in two central aspects. First, I show how the pricing equations for European options can be solved by numerical means and prices obtained quickly and accurately without any need for simulation. Second, I intend to explore the issue of determining the appropriate number of states in the regime-switching model by studying the pricing performance for an increasing number of volatility states. To the best of my knowledge no previous studies exist where this type of model is calibrated to observed option prices. Most research on regime-switching volatility only considers two-state discrete time models. By calibrating the model for an increasing number of volatility states we can investigate whether two states are indeed sufficient to replicate market prices or if more states are needed. Considering call options on the OMXS30 index, I will try to answer the following questions: *Can a regime-switching model replicate observed call option prices? How many regimes are needed to obtain a good fit to data?* The reason for focusing on call options on the OMXS30 index is beacuse these are liquid contracts traded for a large number of strikes and time to maturities. Also, from a structured products desk operating in the Swedish market it could be interesting to consider Swedish data.[2] Potential applications of the model could be in pricing of OTC-contracts or hedging.

This thesis is organized in the following way. Section 2 discusses the relevance of using stochastic volatility models in option pricing. In section 3, the concept of regime-switching volatility is defined. The equations for pricing options under regime-switching volatility are derived in section 4. Section 5 discusses the method employed and the empirical results are presented in section 6. Finally, section 7 summarises the conclusions.

[1] See Alexander (2001) [1] and Hull (1999) [31].

[2] This assigment was given to me by Ola Hammarlid at Swedbank Markets.

2 Motivation for stochastic volatility

As noted above, the assumption of constant volatility made in the Black-Scholes model is inconsistent with the behaviour of real-world financial markets. Contrary to what is assumed in the Black-Scholes model, much empirical research has shown that volatility in financial markets tends to vary quite a lot over time.[3] Convincing evidence that this is indeed the case can be found from stock market crashes like the one in October 1987 or financial crises such as the Mexican crisis in 1994, the Asian crisis in 1997 or the Russian crisis in 1998. The top panel of Figure 1 displays the daily log returns of the OMXS30[4] index over the period August 2003 to August 2006. The bottom panel shows the 30-day historic standard deviation over the same period.[5] During this three year period

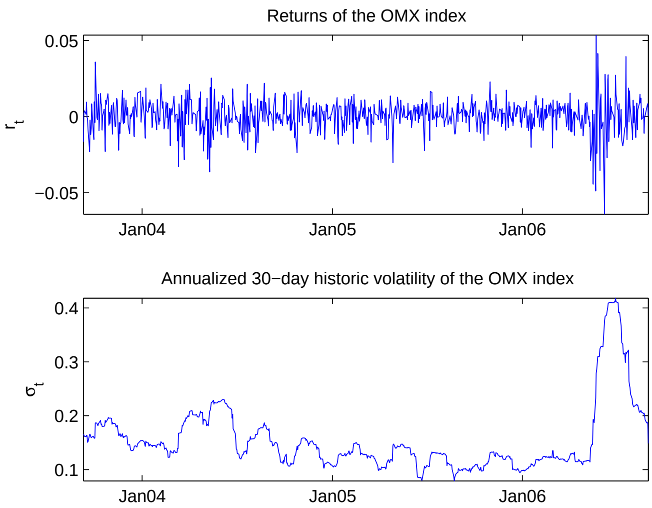

Figure 1: The top panel show the daily log returns of the OMXS30 over the period August 2003 to August 2006. The bottom panel shows the annualized 30-day historic volatility of the log returns.

the historic annualized volatility has varied between 8% and 42%. This indicates that there is indeed a substantial variation of volatility in the Swedish equity market. Furthermore, periods of high and low volatility tend to occur together. This phenomenon is commonly denoted *volatility clustering.*[6] During periods of high volatility, returns can be of a magnitude far greater than returns of more normally volatile periods. One could think of several possible explanations to what causes volatility to vary over time.

[3]See Alexander (2001) [1] for an overwiev of time-varying volatility.

[4]The OMXS30 index is a value weighted index of the thirty largest companies on the Swedish stock exchange.

[5]The 30-day standard deviation has been scaled to yearly standard deviation by multiplication with $\sqrt{12}$, since there are approximately 12 30-day periods in a year.

[6]See Alexander (2001) [1] or Tsay (2002) [34] for a discussion of volatility clustering.

- *Arrival of new information is not evenly distributed in time.* New information arrives to the market lumpwise at discrete points in time. If the information is perceived as very important by the investor community this could have a substantial impact on the level of market activity. Examples of things that might cause market sentiments to change abruptly are corporate events such as mergers or bankruptcies, natural disasters or wars.

- *Fundamental factors of the underlying economy changes.* Apart from arrival of new information, fundamental factors of the underlying economy can also change which in turn causes market activity to change. The most obvious example of this is if we consider the transition from a fixed to a floating exchange rate regime. Under the fixed exchange rate policy volatility in the foreign exchange market should be zero. However, when the exchange rate is floating the volatility is obviously greater since it is allowed to change with supply and demand. Yet another example is changes in the tax legislation making it easier or harder for investors to trade, thereby affecting the volatility on the market.

- *Herd behaviour of market participants.* Sometimes there might be no fundamental causes to abrupt changes in volatility. Such fluctuations could arise when investors try to imitate the behavior of others, thereby magnifying movements already observed on the market.[7] In the literature this goes under the term *herd behaviour*. One explanation to herd behaviour proposed by Avery and Zemsky (1998) [2] is that in financial markets in which information is asymmetrically distributed, some investors will try to mimic the behaviour of other investors in the belief that those are better informed.

Time-varying volatility and volatility clustering cause empirical asset returns to be heavy-tailed. Table 1 shows some descriptive statistics and the Jarquebera statistic for the OMXS30 log returns for the period August 2003 to August 2006. One can see that the standard deviation of the log returns is more than ten times greater than the mean. This is common in most equity markets and makes it practically impossible to estimate the mean of the returns with any reasonable accuracy. The kurtosis is more than twice as great as the kurtosis

Statistic	Mean	Variance	Kurtosis	Jarquebera
Value	0.0007	0.0103	7.5405	699.5519

Table 1: Statistics for OMXS30 log returns.

of a normal distribution and indicates that the empirical distribution of the log returns is heavy-tailed.[8] To test the assumption of normally distributed returns more formally I have employed the Jarquebera test.[9] The Jarquebera statistic was computed to 699.55 which is much greater than the critical value of 9.21 at the 1% significance level. Therefore, we can safely reject the hypothesis of

[7] See Bikhchandani and Sharma (2000) [4].

[8] The kurtosis of a normal distribution is 3.

[9] See Gujarati (2003) [22] for a description of the Jarquebera test.

normally distributed returns at all reasonable significance levels.
From this discussion it should be clear that the assumptions made in the Black-Scholes model of constant volatility and normally distributed log returns are inadequate to explain stock market dynamics. However, market participants are well aware of these anomalies and they take this into account in their option pricing. The market corrects for the heavy-taildness by over-pricing in- and out-of-the-money put and calls and under-pricing at-the-money options relative to the Black-Scholes. This corresponds to implied volatilities[10] being high for in- and out-of-the-money options and relatively lower for at-the-money options. Therefore, a plot of implied volatilities as a function of strike price is often referred to as a volatility smile. The reason for the smile is that the market expects large deviations to occur more frequently than predicted by the Black-Scholes model.[11] Figure 2 shows implied volatilities of a call option on the OMXS30 index on the 30th of November 2005 with maturity in December when the index was at 912. To make prices consistent with a volatility smile and to obtain more

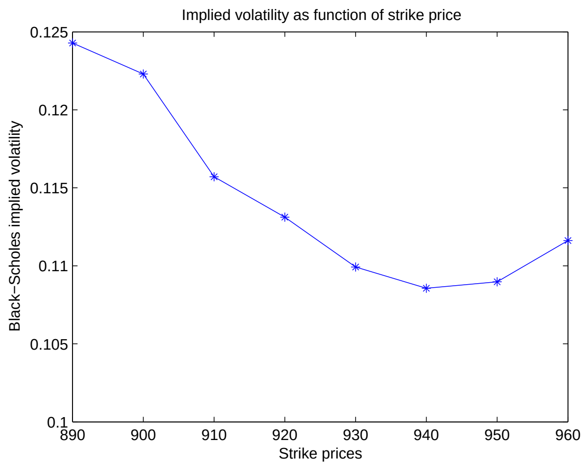

Figure 2: Implied volatilities for a call on the OMXS30.

realistic models of asset price dynamics, the assumption of constant volatility has to be dropped. The natural extension of the Black-Scholes model pursued in the literature is to make the volatility stochastic.[12] By allowing the volatility to follow a stochastic process some of the discrepancies in the Black-Scholes model can be efficiently mitigated. However, option pricing under stochastic volatility constitutes a particularly challenging problem for two reasons. First, the volatility can not be directly observed which makes calibration difficult. Second, when volatility is stochastic the market is incomplete, implying that we can not construct a replicating portfolio in terms of the underlying securities and thus there is no unique arbitrage price of the option. Therefore, option pricing and calibration in the context of a stochastic volatility model becomes a very delicate task.

[10]Implied volatilities make theoretical Black-Scholes prices equal to market prices.
[11]For more on the causes of volatility smiles see Hull (1999) [31].
[12]See Foque, Papanicolaou and Sircar (2000) [21].

3 The model

In section 3.1 the concept of regime-switching volatility is defined. Section 3.2 deepens the discussion and explains the mathematical framework of regime-switching in more detail. Finally, section 3.3 discusses the economical content and the benefits of the model.

3.1 Regime-switching volatility defined

The objective of a regime-switching model is to obtain a model that allows a given variable to follow different time series processes over different subsamples.[13] The intuition of this model is that as conditions in the market change, so does the data generating process of the variable of interest. As argued in section 2, for pricing purposes, we would like to have a model where the volatility of the underlying security is allowed to change randomly in time. Thus, we want to extend the classical Black-Scholes model for the underlying asset S_t as follows

$$dS_t = \mu S_t dt + \sigma(X_t) S_t dW_t,$$

where μ is a constant drift term, $\sigma(X_t)$ is the volatility and dW_t is an increment of a Brownian motion W_t[14]. In this setup volatility depends on the state variable X_t.[15] The crucial assumption of a regime-switching model is that X_t follows a so called markov chain. A markov chain is a stochastic process defined as follows.

Definition 1 *A stochastic process $X_t; t \geq 0$ follows a markov chain if the following two conditions are satisfied.*

- *X_t satisfies the markov property.*
- *X_t takes values from amongst a finite set of states.*

Next, we state the definition of the markov property.[16]

Definition 2 *A stochastic process $X_t : t \geq 0$ is said to fulfil the markov property if at any time $s > t > 0$ the conditional distribution of X_s given the whole history of the process up to and including time t, depends only on the state of the process at time t.*

Loosely speaking, the markov property implies that only the present is relevant for determining the future. Due to this property a markov process is said to lack memory. Thus, a regime-switching volatility model is a model where the volatility jumps randomly amongst a finite number of volatility states and past states can not be used to predict future states. To see what this means in practice we can consider a model with only two volatility states.

$$\text{High volatility state: } \sigma = 70\%$$
$$\text{Low volatility state: } \sigma = 30\%$$

[13] See Hamilton (1989) [25] for a description of regime-switching models.

[14] Recall that a Brownian motion is a stochastic process W_t such that $dW_t \sim N(0, dt)$ and dW_s and dW_t, for $s \neq t$, are independent. For more on the properties of the Brownian motion see Öksendal (1985) [37].

[15] In principle we could let the drift term also depend on X_t. However, as the drift term does not affect option prices it is more convenient to model this as a simple constant.

[16] See Enger and Grandell (2003) [19].

Depending on whether X_t is in the high volatility state or the low volatility state, and for the moment assuming $\mu = 0$ we have the two models as

$$\text{High volatility state: } dS_t = 0.7 S_t dW_t$$
$$\text{Low volatility state: } dS_t = 0.3 S_t dW_t$$

As X_t is stochastic and jumps randomly between the high and the low volatility state, the model of asset price dynamics will also change randomly according to the prevailing volatility state. Figure 3 displays a realization of a simulated path from this model. We see that periods of high volatility occur when the

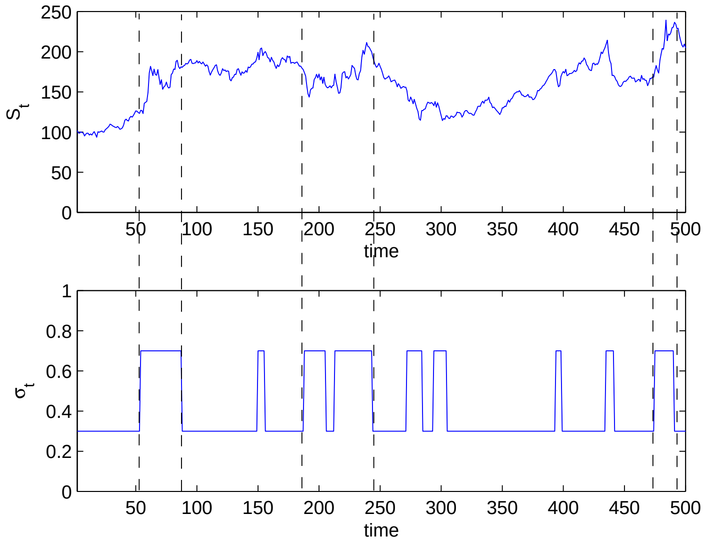

Figure 3: The top panel displays a realization of S_t when the price process is modelled by a geometric Brownian motion with zero drift and a volatility driven by a two-state regime-switching model, shown in the bottom panel.

underlying markov chain makes a transition from the low-volatility regime to the high-volatility regime. Regime-switching models was first introduced by Hamilton (1989) [25] who applied the model to the US GDP growth. Since then, regime-switching has been applied to a multitude of financial and economic variables including stock prices in Hamilton and Susmel (1994) [26], exchange rates in Engle and Hamilton (1990) [20] and interest rates in Wu and Zeng (2004) [35]. Regime-switching volatility models in the context of option pricing was first addressed in Guo (2001) [24] where option pricing in a discrete time framework was considered. In this thesis we will draw upon the results of Buffington and Elliott (2002) [9] and Mamon and Rodrigo (2004) [33] where the volatility is modelled by a continuous time regime-switching model.

3.2 Continuous time regime-switching volatility

To begin the description of the regime-switching model we consider a model with an arbitrary number of states $K \geq 2$. As in Buffington and Elliott (2002) [9] and Mamon and Rodrigo (2004) [33] we take the state space of X_t as the set of unit vectors $e_1, e_2, ..., e_N$, where e_i is a $K \times 1$ column vector with a one at the i:th position and zeros everywhere else

$$e_i = (\underbrace{0, 0, ..., 0, 1}_{i}, \underbrace{0, ..., 0}_{K-i})'.$$

The interpretion is that if the state variable X_t is in volatility state i at time t we have $X_t = e_i$. Thus, a *one at position i in the vector denotes that the volatility is currently in state i*. Conditioned on this information the volatility is a known positive real constant

$$\sigma(X_t | X_t = e_i) = \sigma_i.$$

Here, we consider a regime-switching model in continuous time where $t \in [0, \infty)$. In empirical research, discrete time regime-switching is more frequently employed than continuous time regime-switching.[17] When the regime-switching is used in conjunction with time series data, a discrete time model is preferable since it is more consistent with how real-world time series data are sampled, e.g. hourly, daily, weekly etc. However, a continuous time model is more beneficial to the application considered here since it enables us to derive analytical option pricing formulas. In addition, we need not sample any time series data since the model will be calibrated directly to observed option prices at a given point in time.

We begin by denoting distribution of the random time T_{ij} the markov chain stays in state i before it jumps to state j by F_{ij}. This distribution has to be consistent with the memory-less property of a markov process. The memory-less property implies that conditioned on that T_{ij} is greater than t, the probability that the markov chain will stay in state i for a time $t+h$, for $h \geq 0$, is the same as that the markov chain will stay in that state for a time h

$$P(T_{ij} > t + h | T_{ij} > t) = P(T_{ij} > h). \tag{1}$$

As in Enger and Grandell (2003) [19] one can show that this property is satisfied if and only if T_{ij} is exponentially distributed

$$F_{ij}(t) = 1 - e^{-\lambda_{ij} t}, \text{for } t \geq 0,$$

where λ_{ij} is the intensity with which the markov chain jumps from state i to state j. To see that this choice of distribution indeed fulfils the markov property (1) we compute

$$\begin{aligned} P(T_{ij} > t + h | T_{ij} > t) &= \frac{P(T_{ij} \geq t + h \cap T_{ij} \geq t)}{P(T_{ij} \geq t)} = \frac{1 - P(T_{ij} \leq t + h)}{1 - P(T_{ij} \leq t)} \\ &= \frac{1 - F_{ij}(t + h)}{1 - F_{ij}(t)} = \frac{1 - (1 - e^{-\lambda_{ij}(t+h)})}{1 - (1 - e^{-\lambda_{ij} t})} = \frac{e^{-\lambda_{ij}(t+h)}}{e^{-\lambda_{ij} t}} \\ &= e^{-\lambda_{ij}(t+h) + \lambda_{ij} t} = e^{-\lambda_{ij} h} = 1 - F_{ij}(h) = P(T_{ij} > h). \end{aligned}$$

[17] See Enger Grandell (2003) [19].

Now, considering a very short period of time $h \geq 0$ and using the memory-less property of the markov chain, the probability $p_{ij}(h)$ that the process jumps from state i to state j during this short period of time conditioned on that it stays in state i for a time $t \geq 0$, can be approximated as

$$\begin{aligned} p_{ij}(h) &= P(T_{ij} < t + h | T_{ij} > t) = 1 - P(T_{ij} > t + h | T_{ij} > t) \\ &= \{(1)\} = 1 - P(T_{ij} > h) = P(T_{ij} < h) \\ &= F_{ij}(h) = 1 - e^{-\lambda_{ij} h} \approx \lambda_{ij} h. \end{aligned} \tag{2}$$

In the last step I have used that h is very small so the linearization around $h = 0$ is a rather good approximation. It follows that the probability that the markov chain stays in state i, i.e. does not jump to any of the other $K - 1$ states, during this short period of time is one minus the sum of the probabilities that the process does jump to any other state

$$p_{ii}(h) \approx 1 - \sum_{i \neq j}^{K} \lambda_{ij} h. \tag{3}$$

If we for the moment confine ourselves to two states and put the transition probability $p_{ij}(h)$ as the (i, j):th element[18] in a matrix $\mathbf{P}(h)$ and use equation (2) and (3) we get

$$\mathbf{P}(h) = \begin{pmatrix} p_{11}(h) & p_{12}(h) \\ p_{21}(h) & p_{22}(h) \end{pmatrix} = \begin{pmatrix} 1 - \lambda_{12} h & \lambda_{12} h \\ \lambda_{21} h & 1 - \lambda_{21} h \end{pmatrix}.$$

The matrix $\mathbf{P}(h)$ is known as the transition probability matrix. One can show that the markov property implies that the transition probability matrix must satisfy the so called Chapman-Kolmogorov equation

$$\mathbf{P}(t + h) = \mathbf{P}(t)\mathbf{P}(h), \tag{4}$$

for each $t, h \geq 0$.[19] Subtracting $\mathbf{P}(t)$ from both sides of equation (4) we get

$$\begin{aligned} \mathbf{P}(t + h) - \mathbf{P}(t) &= \mathbf{P}(h)\mathbf{P}(t) - \mathbf{P}(t) = (\mathbf{P}(h) - \mathbf{I})\,\mathbf{P}(t) \\ &= \begin{pmatrix} -\lambda_{12} & \lambda_{12} \\ \lambda_{21} & -\lambda_{21} \end{pmatrix} \mathbf{P}(t) h \\ &= \underbrace{\begin{pmatrix} -\lambda_{11} & \lambda_{12} \\ \lambda_{21} & -\lambda_{22} \end{pmatrix}}_{\mathbf{Q}} \mathbf{P}(t) h \\ &= \mathbf{Q}\mathbf{P}(t) h, \end{aligned}$$

where $\lambda_{11} = \lambda_{12}$, $\lambda_{22} = \lambda_{21}$ and $\mathbf{I}$ is the identity matrix[20]. Dividing by h on both sides in this last expression and letting $h \to 0$ we get

$$\frac{d\mathbf{P}}{dt}(t) = \mathbf{Q}\mathbf{P}(t). \tag{5}$$

[18]The (i, j):th element in a matrix refers to the element on the i:th row and the j:th column in that matrix.

[19]See Enger and Grandell (2003) [19].

[20]$\mathbf{I}$ is a 3×3 matrix with ones at the main diagonal and zeros every where else.

This system is referred to as the Kolmogorow equations and can be solved explicitly using standard techniques.[21] As we can see in equation (5) the dynamics of the transition probabilities are controlled by the intensity matrix $\mathbf{Q}$. In the general K-state case this matrix has the following appearance

$$\mathbf{Q} = \begin{pmatrix} -\lambda_{11} & \lambda_{12} & \cdots & \lambda_{1K} \\ \lambda_{21} & -\lambda_{22} & \cdots & \lambda_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{K1} & \lambda_{K2} & \cdots & -\lambda_{KK} \end{pmatrix},$$

where $\lambda_{ii} = \sum_{j \neq i}^{K} \lambda_{ij}$. The parameter λ_{ii} can be given the interpretation as the *total* intensity with which the markov chain jumps out of state i and the expected time the process stays in state i is $1/\lambda_{ii}$. Thus, as we measured time in number of years, a λ_{ii} of 12 can be interpreted as that the expected time the markov chains stays in state i is one month. In a K-state model the matrix contains $K(K-1)$ unknown parameters since the diagonal elements are given by the off-diagonal elements. For calibration purposes we therefore need to estimate $K(K-1)$ intensities. Figure 4 gives a schematic illustration of the concepts developed in this section for a two-state model. The transition intensities and the arrows show which parameter controls which transition. The state with high volatility represents turbulent periods with high market activity and the state with low volatility represents more quiet periods with smaller price movements.

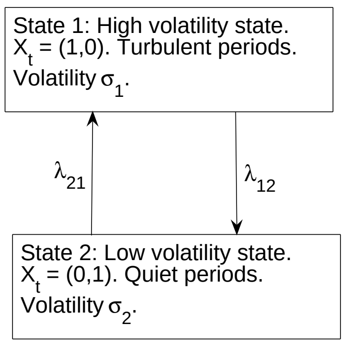

Figure 4: Illustration of a two-state regime-switching model.

[21] For techniques how to solve this problem see Boyce and DiPrima (2000) [8].

3.3 Rationale for regime-switching volatility

Above, the mathematics behind the regime-switching volatility model have been described in detail. However, so far not much have been said about the economical content of the states. The states were first given an economical interpretation in Guo (2001) [24] where they represent different degrees of asymmetric information in the investor community. When the volatility is in the lower state, price movements are moderate and people believe that they are well informed in a seemingly complete market. When some investors have more information than other investors, this could cause larger fluctuations on the market as the better informed investors exploit the perceived mispricings by heavy trading. Even though the trades of the better informed investors do not have any substantial influence on market prices, the market could become more volatile if the less informed investors suspect that some groups or individuals have exclusive information. This could give rise to herding behaviour as the less informed investors react fiercely on small movements in the market.[22] Thus, when information is asymmetrically distributed among investors we are in a high volatility state. Although the states could be given some economical content this will not be crucial for pricing purposes. Rather, the choice of using a regime-switching model for the volatility can be motivated on more pragmatic grounds. I consider the regime-switching model to be very beneficial of the following reasons.

- *Generates volatility clustering.* Once a transition has occurred, the volatility tends to stay in that state for some time. Thus, periods of high and low volatility will be somewhat clustered.

- *Generates heavy-tailed distributions.* High volatility regimes can produce returns of a magnitude that by far exceeds the returns of more quiet periods and thereby increase the probability of extreme movements.

- *Replicates various volatility structures.* As noted in Section 2, implied Black-Scholes volatilities often vary across strike and time to maturity. As we will see, a regime-switching volatility model can generate prices corresponding to various volatility structures and thus pricing will be facilitated.

- *Simple and easy to understand.* The concept of regime-switching volatility is intuitively appealing and easy to understand. This enhances the value of the model since it can readily be communicated to traders and sales people.

- *Consistent with modern financial theory.* As described in section 3.1, a regime switching-model satisfies the markov property so that the history has no predicitve power on future outcomes. Thus, the model is consistent with the Efficient Market Hypothesis as this implies that all information, past or present, should already be incorporated into market prices.[23]

[22] This is in line with the arguments in Avery and Zemsky (1998) [2], see section 2.

[23] The argument behind the Efficient Market Hypothesis goes like this. If the past do indeed has some predictive power on the future, rational investors will realize this and exploit these dependences. Consequently, supply and demand will cause prices to adjust until all information has been incorporated into market prices. If the market is efficient this process of sequential adjustments is instantaneous and past information has no relevance for determining future prices. (Grinblatt and Titman, 2002)

4 Option pricing

Section 4.1 gives brief introduction to option pricing and derives the general pricing formula. Section 4.2 discusses some of the problems faced when pricing options under stochastic volatility and the pricing equations for a regime-switching model are derived.

4.1 Option pricing in general

The main problem in derivatives pricing is to find a price process

$$\Big\{\Pi(t;\Phi); t \in [0,\infty)\Big\}$$

for a contingent claim which pays off the amount

$$\Phi(S_T)$$

on a future date T. Since the payoff is known at time T the value of the contingent claim at time T is

$$\Pi(T;\Phi) = \Phi(S_T).$$

However, since we are standing at time $t < T$, the value of the risky asset at time T is not known and so is the value of the derivative. The starting point to this problem is to formulate a model for the dynamics of the underlying asset S_t. Once we have formulated a model for the stock price dynamics, the objective is to price the derivative in a way that is consistent with the chosen price process. More specifically, we want to price the derivative so that we do not introduce arbitrage on the market. Harrison and Kreps (1979) [28] and Harrison and Pliska (1983) [29] have shown that absence of arbitrage is equivalent to the existence of a risk adjusted measure $\mathcal{Q}$ under which the discounted price processes of all traded risky securities are martingales. Since this statement will be very important for the derivation below we state it formally as a theorem.

Theorem 1 *The market is free of arbitrage if and only if there exists a risk adjusted measure $\mathcal{Q}$ such that the discounted price processes of the traded risky assets $S^1(t), S^2(t), ..., S^N(t)$*

$$\frac{S^1(t)}{B(t)}, \frac{S^2(t)}{B(t)}, ..., \frac{S^N(t)}{B(t)},$$

are martingales under $\mathcal{Q}$.

As usual, B_t is the money-market account with dynamics given by

$$dB_t = rB_t dt, \quad B_0 = 1,$$

which implies that

$$B_t = e^{rt}.$$

That the discounted price process of all traded risky asset are martingales under $\mathcal{Q}$ implies that

$$\frac{S_t^i}{B(t)} = E_t^{\mathcal{Q}}\Big[\frac{S_T^i}{B(T)}\Big], \quad \text{for } i = 1,..,N, \tag{6}$$

where $E_t^{\mathcal{Q}}[\cdot]$ denotes that the expectation is taken under $\mathcal{Q}$ and conditioned on all information generated until time t.[24] Since equation (6) also holds for the price process of the derivative, given that it is traded, and using that $\Pi(T;\Phi) = \Phi(S_T)$ we obtain the familiar pricing formula as

$$\Pi(t;\Phi) = e^{-r(T-t)} E_t^{\mathcal{Q}}\Big[\Phi(S_T)\Big]. \tag{7}$$

Thus, the problem of pricing the derivative boils down to finding a risk adjusted measure $\mathcal{Q}$ under which the discounted price processes of all traded assets are martingales. If we can find a *unique* risk adjusted measure that satisfies relation (6) then our job is done and we can obtain unique prices of all derivatives. However, if the risk adjusted measure is not unique then there exists a whole set of prices which are all consistent with the no-arbitrage assumption. As will be explained below, when volatility is stochastic the martingale measure is not unique and prices can not be uniquely determined. Thus, we need to make some additional assumptions to be able to price options under regime switching-volatility.

4.2 Option pricing under regime-switching volatility

In section 4.2.1 various approaches to selecting the risk adjusted measure are discussed and in section 4.2.2 the mathematical derivation of the pricing equations is outlined.

4.2.1 Selecting the risk adjusted measure

The problem of pricing options under stochastic volatility is that the risk adjusted measure $\mathcal{Q}$ can not be uniquely determined. So, why is the risk adjusted measure not unique? One can show that the risk adjusted measure is unique if and only if the market is complete. In general, a necessary condition for completeness is that the number of traded securities are at least as great as the number of random sources in the economy.[25] In the model considered here, we have effectively introduced $1 + K(K-1)$ random sources in the market, one random source corresponding to the Wiener process W_t and $K(K-1)$ random sources corresponding to the transitions between the volatility states. Since there is only one risky security S_t, the market is not complete and therefore the price of the derivative is not unique. In more economical terms the incompleteness means that we can not construct a replicating, self-financing portfolio in terms of underlying securities which generates exactly the same payoff as the derivative. Thus, to obtain unique prices we need to impose some additional assumptions on the market model. Assumptions that might resolve the problem are:

1. Assume a specific form of the utility function of the investors.
2. Assume how the market prices volatility risk.

[24] More formally we should write that the expectation is taken under the filtration $\mathcal{F}_t$ generated by the σ-field $\sigma\{S_u, X_u : u < t\}$, which can be interpreted as all the information generated by S and X up until time t. See Öksendal (1985) [37] for a introduction to σ-fields.

[25] See Björk (2004) [5] for a discussion of the relation between completeness of the market and uniqueness of the martingale measure.

3. Assume that the risk adjusted measure is chosen to be by some measure as close as possible to the objective probability measure.

4. Assume that there exists so called change-of-state securities which payoffs are linked to the change of volatility state.

5. Assume what the dynamics look like under the risk adjusted measure $\mathcal{Q}$.

In the literature, all approaches have been considered.[26] In Chourdakis (2002) [11] and Duan (2002) [16] the first approach is used on a two-state discrete time model and assumptions on the utility function of the investors are made. Hardy (2001) [27] adopts the second approach and assumes that volatility risk is not priced in the market.[27] Another method is to assume that the market chooses the risk adjusted measure which is by some measure as close as possible to the objective probability measure $\mathcal{P}$. One common technique is to choose the $\mathcal{Q}$ that maximizes the information the two measures have in common.[28] This approach has been pursued in Chan, Elliott and Siu (2005) [10] and Di Masi, Kabanov and Runggaldier (1993) [14]. As argued by Guo (2001) [24] one could also find unique option prices by to completing the market by introducing so called change-of-state securities. These are securities which pay one unit of account when the volatility changes states and then become worthless. Then, a new change of-state security is issued that pays off at the next change of state. One can realise that this will indeed complete the market and we can obtain prices in terms of the underlying and the change-of-state securities. The last approach is considered in Buffington and Elliott (2002) [9] and Mamon and Rodrigo (2004) [33] where the dynamics of X_t is modelled directly under the risk adjusted measure $\mathcal{Q}$. Thus, we do not care about what the dynamics look like in the real-world since it is only the dynamics under the risk adjusted measure that matters for pricing. The drawback of this method is that we do not know any of the parameters in the model since the $\mathcal{Q}$-dynamics is unobservable to us. The only way to choose the appropriate parameters is to back them out from observed prices so that the model produces the same prices as those observed on the market. In bond option pricing, where the market is also incomplete, this technique is commonly known as martingale modelling where it is the dynamics for the short rate that is modelled directly under the risk adjusted measure.[29]

The standpoint taken here is that there is no convincing evidence why a specific utility function should be preferred to another or how the market should price the extra risk introduced by the regime-switching volatility. It is also very hard to find any solid economical arguments why the market should choose the risk adjusted measure to be as close as possible to the objective probability measure. Neither is pricing in terms of change-of-state securities feasible since such instruments do not exist on the market. Therefore, I consider the last approach,

[26] Here, we will not go into details about how the different approaches resolve the problem in practice. For more information please refer to the cited articles.

[27] Bollen (1998) [7] relies on the same assumption.

[28] This means that the relative entropy between the two measures is minimized, see Chan, Elliott and Siu (2005) [10].

[29] Models for the short-term interest rate under $\mathcal{Q}$ are known as no-arbitrage models and the technique of backing the parameters out from observed bond prices is known as inverting the yield curve. See Björk (2004) [5] and Hull (1999) [31].

where the dynamics of X_t is modelled directly under $\mathcal{Q}$, to be the most appealing as this does not require us to assume anything about the utility or the risk preferences of the market. The issue of choosing the appropriate risk adjusted measure is resolved since we intend to calibrate the model to observed market prices which contain information about the specific risk adjusted measure used by the market. In fact, parameters obtained in this manner will reflect all information available on the market including the degree of risk aversion and various factors affecting supply and demand including liquidity constraints, among others. Below we will show that the model is tractable and can be used to derive equations which can be solved for the price of the derivative. In addition, the model is guaranteed to be free of arbitrage.

4.2.2 Derivation of the pricing equations

To begin the derivation of the pricing equations we first apply Theorem 1 and Itô's formula to S_t/B_t to obtain the $\mathcal{Q}$-dynamics of S_t as

$$dS_t = rS_t dt + \sigma(X_t)S_t dW_t, \tag{8}$$

where as before X_t follows a K-state continuous time markov chain and dW_t is a Wiener increment under $\mathcal{Q}$.[30] Equation (8) always holds regardless of choice of models for the volatility or the risk-free interest rate. For the moment we assume that the risky asset S_t does not pay any dividends. As in the standard Black-Scholes model we consider a market which fulfils the following set of assumptions.

Assumption 1 *All assets are infinitely divisible*[31]*.*

Assumption 2 *The risk-free interest rate is a deterministic constant.*

Assumption 3 *Trades can be made continuously in time.*

Assumption 4 *The market is free of arbitrage.*

Assumption 5 *The price process for the derivative asset is of the form*

$$\Pi(t; \Phi) = F(t, S_t, X_t),$$

where $\Phi(S_T)$ is the final payoff and F is some smooth function.

In addition to these assumptions we also state the following crucial assumption.

Assumption 6 *X_t follows a continuous regime-switching model under $\mathcal{Q}$.*

Now, from (7) the price of an option with final payoff $\Phi(S_T)$ is given by

$$\Pi(t; \Phi) = e^{-r(T-t)} E_t^{\mathcal{Q}}\Big[\Phi(S_T)\Big].$$

According to assumption 5 we can price this derivative using the function F as

$$F(t, S_t, X_t) = \Pi(t; \Phi).$$

[30]The derivation of the dynamics of S_t under $\mathcal{Q}$ can be found in Björk (2004) [5].

[31]This means that we can sell or buy any number or fraction of an asset

Our task is to determine what the function $F(t,s,x)$ might look like, or at least to derive some relation which F must satisfy. To derive the pricing equation satisfied by $F(t,s,x)$ we will use relation (6) which states that the discounted price processes of all traded securities are martingales under $\mathcal{Q}$. This implies that the discounted price process

$$V(t,S_t,X_t) = e^{-rt}F(t,S_t,X_t),$$

should be a martingale under $\mathcal{Q}$. That V is a martingale means that the future expected value of V is equal to today's value. How can we use this to price the option? To illustrate the conceptual idea, let us consider a random variable Y_t with the stochastic differential

$$dY_t = \mu_t dt + \sigma_t dM_t,$$

where M_t is a martingale. Then, modulo some technicalities, a necessary condition for Y_t to be a martingale is that *the drift term is zero*

$$\mu_t = 0.$$

See Öksendal (1985) [37] for a more rigorous argument. From this example the mechanics should be fairly clear. If we can derive a stochastic differential for V, we can impose the condition that the drift term must equal zero to derive some relation which V must satisfy. Consequently, we need to go through the following steps.
Step 1. Derive the stochastic differential for V.
Step 2. Set the drift term of V_t to zero and derive the pricing equations.
Below we will carry out both steps in turn.

Step 1 - Derive the stochastic differential of V

To start the derivation of the stochastic differential of V we begin by setting $V(t,S_t,X_t)$ conditioned on that the markov chain is in state i as

$$V_i(t,S_t) = V(t,S_t,X_t|X_t = e_i),$$

for $i = 1,...K$. We now set

$$\mathbf{V}(t,S_t) = \left(V_1(t,S_t),...,V_K(t,S_t)\right),$$

so that $\mathbf{V}(t,S_t)$ a $1\times K$ row vector. Since X_t is a $K\times 1$ column vector, $V(t,S_t,X_t)$ can be succinctly expressed as a matrix multiplication

$$V(t,S_t,X_t) = \mathbf{V}(t,S_t)X_t = \sum_{i=1}^{K} V_i(t,S_t)X_t^i,$$

where X_t^i is the i:th element of the vector X_t. Of course, as the state space of X_t is the set of unit vectors, see section 3.2, only one of the terms in this sum will be non-zero. Applying the multidimensional Itô rule[32] to $V(t,S_t,X_t)$ we

[32] See Öksendal (1985) [37].

obtain the stochastic differential[33] as

$$\begin{aligned} dV_t &= \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial s}dS_t + \frac{1}{2}\frac{\partial^2 V}{\partial s^2}(dS_t)^2 + \mathbf{V}dX_t \\ &= \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial s}(rS_t dt + \sigma_t S_t dW_t) + \frac{1}{2}\frac{\partial^2 V}{\partial s^2}\sigma_t^2 S_t^2 dt + \mathbf{V}dX_t. \end{aligned}$$

The stochastic differential of V involves the differential of X_t. As in Elliott, Aggoun and Moore (1994) [18] one can show that the differential of X_t has a representation of the form

$$dX_t = \mathbf{Q}X_t dt + dM_t,$$

where M_t is a K-dimensional martingale and $\mathbf{Q}$ is a before the intensity matrix of the markov chain. Inserting this into the Itô expansion and collecting terms we obtain the stochastic differential of V as

$$\begin{aligned} dV_t &= \Big(\frac{\partial V}{\partial t} + rS_t\frac{\partial V}{\partial s} + \mathbf{VQ}X_t + \frac{1}{2}\sigma_t^2 S_t^2\frac{\partial^2 V}{\partial s^2}\Big)dt \\ &+ S_t\sigma_t\frac{\partial V}{\partial s}dW_t + \mathbf{V}dM_t. \end{aligned} \tag{9}$$

Step 2: Derive the pricing equations

By definition V is a martingale and thus the drift terms in (9) must sum to zero. That is

$$\frac{\partial V}{\partial t} + rS_t\frac{\partial V}{\partial s} + \frac{1}{2}\sigma_t^2 S_t^2\frac{\partial^2 V}{\partial s^2} + \mathbf{VQ}X_t = 0. \tag{10}$$

Since $V_i(t, S_t) = e^{-rt}F_i(t, S_t)$ and by setting

$$\mathbf{F}(t, S_t) = (F_i(t, S_t), ..., F_K(t, S_t))$$

equation (10) transforms into

$$-rF + \frac{\partial F}{\partial t} + rS_t\frac{\partial F}{\partial s} + \frac{1}{2}\sigma_t^2 S_t^2\frac{\partial^2 F}{\partial s^2} + \mathbf{FQ}X_t = 0.$$

Also, we must have that

$$\Pi(T; \Phi) = \Phi(S_T).$$

These two equations have to hold with probability 1 for each fixed t.[34] Under some weak assumptions one can show that this holds for any S_t, so $F(t, s, x)$ must satisfy the completely deterministic equation

$$-rF + \frac{\partial F}{\partial t} + rs\frac{\partial F}{\partial s} + \frac{1}{2}\sigma_t^2 s^2\frac{\partial^2 F}{\partial s^2} + \mathbf{FQ}X_t = 0$$

with the boundary condition

$$F(T, s, x) = \Phi(s).$$

[33] In the differential the notation σ_t is used to denote $\sigma(X_t)$.
[34] See Björk (2004) [5].

Since this holds for every choice of volatility state X_t we must have that

$$-rF_i + \frac{\partial F_i}{\partial t} + rs\frac{\partial F_i}{\partial s} + \frac{1}{2}\sigma_i^2 s^2 \frac{\partial^2 F_i}{\partial s^2}\mathbf{FQ}e_i = 0 \qquad (11)$$

with boundary condition

$$F_i(T,s) = \Phi(s),$$

for $i = 1, ..., K$. Carrying out the matrix multiplication in the last term of equation (11) we obtain K equations of the following kind

$$-rF_i + \frac{\partial F_i}{\partial t} + rs\frac{\partial F_i}{\partial s} + \frac{1}{2}\sigma_i^2 s^2 \frac{\partial^2 F_i}{\partial s^2} - \lambda_{ii}F_i + \sum_{j\neq i}^{K} \lambda_{ji}F_j = 0, \qquad (12)$$

for $i = 1, ..., K$. This completes the derivation of the pricing equations.

As an example we consider a call option and a model with only two volatility states. In this simple case the pricing equations become

$$\begin{aligned} &-rF_1 + \frac{\partial F_1}{\partial t} + rs\frac{\partial F_1}{\partial s} + \frac{1}{2}\sigma_1^2 s^2 \frac{\partial^2 F_1}{\partial s^2} - \lambda_{11}F_1 + \lambda_{21}F_2 = 0 \\ &-rF_2 + \frac{\partial F_2}{\partial t} + rs\frac{\partial F_2}{\partial s} + \frac{1}{2}\sigma_2^2 s^2 \frac{\partial^2 F_2}{\partial s^2} - \lambda_{22}F_2 + \lambda_{12}F_1 = 0 \end{aligned}$$

and the boundary conditions

$$\begin{aligned} F_1(T,s) &= \max(s-K,0), \\ F_2(T,s) &= \max(s-K,0). \end{aligned}$$

Solving these equations yield two prices $F_1(t,s_t)$ and $F_2(t,s_t)$, one for each regime. Since the price of the option depends on the current regime, the investor has to decide which one of these prices is the true price. Since there is no easy way to determine the true state one often has to rely on some form of expert judgement. As discussed in Appendix II, the selection of the current regime is not a severe problem since we determine this within the calibration. In the above derivation we have assumed that the risky asset does not pay dividends. If the underlying pays a continuous dividend stream d we can follow the argument presented in Hull (1999) [31] and correct for this by replacing S_t with $S_t e^{-d(T-t)}$.

Equation (12) resembles very much the standard Black-Scholes PDE. Indeed, if all transition intensities are zero the two models are identical. In the case of non-zero intensities, the price in each regime is coupled with the prices in all the other regimes and we are therefore required to solve the system of equations simultaneously. This makes pricing in a regime-switching model more computationally demanding than in the Black-Scholes model since no analytical methods exist which enable us to derive a closed-form solution for the price of the option. However, there are powerful techniques available that can be used to derive approximate solutions to these type of problems. A detailed discussion of such techniques is outlined in Appendix I.

5 Data and method

The pricing equations derived in section 4.2.2 can be solved by numerical methods. Once we know how to price options in this framework, our goal is to calibrate the model to observed option prices. That is, we want to back out the parameters from the model by making the theoretical prices as close as possible to prices observed on the market. Here, I will not further discuss how the pricing equations have been solved or how the model have been calibrated to observed market prices. The interested reader may refer to Appendix I and Appendix II for a detailed discussion of the techniques employed.

When calibrating the model it is important that the number of contracts used is sufficiently large to ensure that the algorithm converges quickly to the optimal solution. To avoid problems of instability and over-fitting we have to use at least as many contracts in the estimation as there are unknown parameters. For a K-state model we have $K(K-1)$ transition intensities and K volatilities so in total we have K^2 unknown parameters. Thus, for a two-state model we need at least four contracts, for a three-state model we need nine contracts and so on.

Another important issue is that the options used for calibrating the model are relatively liquid. Illiquid contracts are potential outliers which could distort the calibration. To minimize the impact of such problems I have focused on call options on the OMXS30 index.[35] The OMXS30 index options are European call and put options on the OMXS30 index with maturities within the next 36 calendar months. Expiration takes place at the fourth Friday in each month. These options are very liquid and there exists a relatively large amount of traded options of different strikes and maturities. I have excluded put options since this would introduce a systematic error in the results corresponding to the spread between put and call options. In addition, put and call options are usually quoted for the same set of prices and therefore adding put options would not increase the total number of contracts substantially. The price data covering the period from August 2003 to August 2006 was obtained from the newspaper Dagens Industri.[36] During this period I have on the fourth Friday in each month collected the prices at which the last trade took place of all call options with expiration in the next two calendar months. This is the same day as expiration took place and new options were introduced.[37] The reason for only considering options with expiration within two months is due to the significantly better liquidity on those contracts. To reduce the risk of having illiquid contracts affecting the calibration I have excluded those contracts where the bid-ask spread was greater than 20%. At the time this study was initiated, August 2006 was the latest month available containing price data on the fourth Friday. In the three year period preceding this month, the volatility has been medium high during August 2003 to December 2004, low during January 2005 to April 2006, and high from May 2006, see Figure 1. Thus, this period enable us to study how the parameters have varied under different levels of market

[35] The OMXS30 index is a basket of the 30 largest companies on the Swedish stock exchange weighted by their market capitalizations.

[36] The newspapers were obtained from the archive at the Stockholm School of Economics.

[37] The prices for specific date are quoted in the newspaper the following weekday.

activity.[38] As the risk-free interest rate I have taken the three month Swedish government bond rate and the dividend yield has been approximated with the average dividend yield of all companies included in the OMXS30 index. This data was also obtained from Dagens Industri. Since there are 36 months in the time period studied, a total number of 36 data sets were obtained. Given these data sets I have proceeded as follows.

1. On each estimation date collect all call options with maturity within the next two months and bid-ask spreads less than 20%.

2. Take out the at-the-money option with shortest time to maturity. We denote this as the benchmark option.

3. Calibrate the model using all contracts, except the benchmark option.

4. Price the benchmark option with the calibrated model and compute the percentage deviation.

The steps 1 to 4 were performed for models with two, three and four volatility states. In this way both the in-sample and the out-of-sample pricing performance for an increasing number of states can be studied. The reason for choosing an at-the-money option with a short time to maturity as the benchmark option on which to evalute the out-of-sample pricing performance, is that this is usually the most liquid of all traded options. Thus, deviations of the theoretical prices from the benchmark prices should not be to due to insufficient liquidity. Also, using an at-the-money option to evaluate pricing performance is interesting from the viewpoint of a structured products desk since new OTC-contracts are usually written at-the-money.[39] Therefore, it is important that the model performs well for at-the-money options.

I have considered a maximum of four states since five states or more require a substantially larger number of quoted options. A five state model contains 25 unknown parameters and we therefore require at least 26 quoted options, including the benchmark option. Since there rarely exists 26 liquid options in the market this effectively rules out the five-state model. A four-state model contains 16 unknown parameters and thus requires only 17 contracts, including the benchmark option. Since almost all data sets contained 17 contracts I limited the study to a maximum number of four states. However, after applying the liquidity criteria some of the data sets had an insufficient number of contracts. Those dates were the 25th of February 2005, the 25th of March 2005, the 27th of January 2006, the 24th of February 2006 and the 24th of June 2006. The model was not calibrated on those dates. Thus, after removing the dates with an insufficient number of liquid contracts we got a total of 31 data sets.

[38]Considering the limited scope of this thesis and other time constraints, a three year period was deemed sufficient.

[39]This was pointed out to me by Ola Hammarlid at Swedbank Markets.

6 Empirical Results

In section 6.1 the parameters of the calibration are shown. The results of the in-sample pricing performance are presented in section 6.2 and in section 6.3 the out-of-sample pricing performance of the different models is discussed. Stability of the parameters is analysed in section 6.4.

6.1 Parameters

As described in section 5, I have on the last day in each month calibrated the model for two, three and four states to observed option prices with maturity within the next two months. The calibrated model is therefore fitted to a volatility surface since we consider variation across both moneyness and remaining time to maturity. In Table 2 the average values over the whole sample of the estimated volatility parameters are shown.[40] For the two-state model, the

Parameter	Two states	Three states	Four states
$\hat{\sigma}_1$	19.9%	18.0%	30.2%
$\hat{\bar{\sigma}}_2$	10.9%	13.7%	15.1%
$\hat{\bar{\sigma}}_3$	-	9.7%	10.7%
$\hat{\bar{\sigma}}_4$	-	-	7.3%

Table 2: The average level of the estimated volatilities in each state.

average volatility in the high volatility state and the low volatility state were bout 20% and 11%, respectively. This can be compared with the average historic volatility during this time period of 16%. That the historic volatility lies somewhere between the high and the low volatility states should be reasonable since we do not expect the market to either underestimate or overestimate the average volatility in the long-run. For all models there seems to be a relatively large gap between the highest and the second highest volatility state whereas the less volatile states are closer to each other. As states with similar volatilities also generate similar prices, not much is gained by including more states in the model. On the contrary, several states with similar volatilities could result in unstable estimates since this makes the parameter surface flatter and thereby harder for the calibration algorithm to find an optimal solution.

In Table 3 the averages of the intensity parameters for each of the models are shown. For a two-state model the volatility jumps on average 6.1 times per year from the high volatility state to the low volatility state and jump 6.6 times per year from the low volatility state to the high volatility state. On average, the expected time the volatility stays in either state is about 2 months.[41] The transition intensities could also be interpreted as a measure of speed of mean reversion of the volatility process (Foque, Papanicolaou and Sircar, 2000). To explain this concept more clearly we can think of the two-state volatility process as switching between a high and a low state with the expected volatility level lying somewhere in between.[42] The intensities are a measure of how fast the

[40] Averages are denoted by hats on the estimated parameters.

[41] Recall that the expected time the volatility stays in regime i is $1/\lambda_{ii}$.

[42] In fact, we can find the expected value explicitly by solving equation (5) and letting

Parameter	Two-states	Three-states	Four-states
$\bar{\hat{\lambda}}_{12}$	6.1	4.4	5.4
$\bar{\hat{\lambda}}_{13}$	-	7.2	4.2
$\bar{\hat{\lambda}}_{14}$	-	-	4.3
$\bar{\hat{\lambda}}_{21}$	6.6	10.5	4.7
$\bar{\hat{\lambda}}_{23}$	-	8.8	4.8
$\bar{\hat{\lambda}}_{24}$	-	-	5.1
$\bar{\hat{\lambda}}_{31}$	-	7.3	5.0
$\bar{\hat{\lambda}}_{32}$	-	9.0	4.7
$\bar{\hat{\lambda}}_{34}$	-	-	4.8
$\bar{\hat{\lambda}}_{41}$	-	-	4.6
$\bar{\hat{\lambda}}_{42}$	-	-	5.0
$\bar{\hat{\lambda}}_{43}$	-	-	4.2

Table 3: The average level of the estimated volatilities in each state .

volatility mean-reverts to this expected level since a very high intensity will make the volatility jump from the high volatility state to the low volatility state more often and vice versa. Foque, Papanicolau and Sircar (2000) [21] develop a framework for estimating a stochastic volatility model based on the idea that the speed of mean reversion of the volatility in the market is very high. They argue that if the volatility runs very fast, i.e. the volatility returns to the mean-value in a few days, then the market would behave almost as in a constant volatility Black-Scholes model, as the volatility on average will be more or less equal to the expected value. Thus, one should be able to treat the market as a small perturbation from a Black-Scholes world. The findings in this study question this reasoning since the average λ of the two-state model was about 6, which can not be considered a very high speed of mean-reversion. What we have found is instead that a slowly varying volatility process is more consistent with observed market prices. As the volatility varies slowly we also suspect that there will be a substantial difference in prices between the low and the high volatility state. This is because the volatility tends to stay some time in that state before switching to another state. Thus, the level of the volatility in the current state will have a larger impact on the price compared to a situation where the volatility jumps more frequently back and forth. To see what the difference in volatilities means in terms of prices for the two-state model I have computed the price of a two month call option on an underlying that is currently worth 100 SEK with a strike of 100, a volatility of 20% in the high volatility state and 11 % in the low volatility state. The risk-free interest rate has been set to zero and the transition intensities have been set to 6 in each state. The price in the high volatility and the low volatility state becomes 2.9 and 2.3, respectively, a difference of 26%.

As noted in section 3.2, the total intensity for leaving a state is the sum

$t \to \infty$ to obtain the time-invariant distribution which can then be used to compute the expected value.

of all intensities for jumping between that state and all other states. The intensity for leaving state one is 11.6 (4.4+7.2) in the three-state model and 13.9 (5.4+4.2+4.3) in the four-state model. Thus, in a model with more states, transitions tend to occur a lot more frequently than in a model with fewer states. This is expected as a model with two states can only discern whether the volatility is high or low, not if it is medium-high or medium-low. Transitions between e.g. a low and a medium-low volatility state will therefore go undetected in the two-state model. As a consequence, transitions intensities in a model with fewer states will be lower.

In this report I have calibrated the model under the assumption that the volatility follows a regime-switching volatility model under the risk adjusted measure $\mathcal{Q}$. It is important to stress that parameters obtained in this way are *not* equal to the parameters under the actual probability measure $\mathcal{P}$, which describes real-world asset price dynamics. The estimates obtained here also reflect the degree of risk aversion in the investor community, liquidity constraints and various supply and demand factors. Therefore, we can not directly compare the model developed above with a model calibrated directly to historic asset returns. It is most likely that a model calibrated on historic asset returns will yield quite different estimates of the parameters compared to a model estimated from observed option prices. As previously noted, one reason to this is that parameters calibrated to observed option prices take into account the degree of risk aversion of the market. Furthermore, parameters that have been implied out from observed option prices reflect the markets *future* expectation on the volatility process whilst a model calibrated on historic data will provide a description of the volatility model that has prevailed in the *past*. Since the models differ substantially with respect to underlying assumptions I argue that it is not a meaningful exercise to use one model to validate or invalidate the other. A more relevant criterion on which to value a model should be to study its performance in the actual context it will be used. Since the main application of the model will be pricing, this should also be the relevant operational performance criterion.

6.2 In-sample fit

To evaluate the in-sample pricing performance I have calculated one R^2 for the options with approximately one month to maturity and another R^2 for the options with approximately two months to maturity. The mean R^2 values for different number of regimes are displayed in Table 4. From Table 4 we make

Number of states	$\bar{R}^2_{1\text{ month}}$	$\bar{R}^2_{2\text{ months}}$
Two states	0.9941 (0.8738)	0.9935 (0.9190)
Three states	0.9970 (0.9621)	0.9956 (0.9566)
Four states	0.9981 (0.9941)	0.9968 (0.9837)

Table 4: Mean R^2 values for different number of states and minimum and maximum R^2 values. Minimum values in parentheses.

the following observations. First, the overall in-sample fit is very good since the mean R^2 values are very close to one for all states and maturities. Even if the models did provide a perfect description of the data, a slightly smaller value seems natural since all traded options are to some extent affected by liquidity factors causing prices to deviate somewhat from the true model. Second, the in-sample-fit to options with one month to maturity is higher than for options with two months to maturity. This is probably due to the fact that a larger number of options with one month to maturity have been used in the calibration thus giving larger weight on obtaining a good fit for those contracts. It is reasonable to assume that a larger number of liquid contracts with two months to maturity would improve the R^2 value for longer maturities. Third, the in-sample fit is higher the more states are included in the model. This is expected since the degrees of freedom increase with the number of states, thereby enabling a better fit to the data. However, I argue that this increase in the R^2 has to be weighted against the larger number of parameters used. It is harder to obtain good estimates if the number of parameters is large and the time it takes for estimating the model is substantially longer. Since the two-state model provides a very good fit to data and not much is gained by increasing the number of parameters, I argue that a two-state model is sufficient. In Figure 5 I have plotted the observed prices (stars) and the theoretical prices (rings) for a two-state model at the 27th of December 2005. As discussed above, the stars in

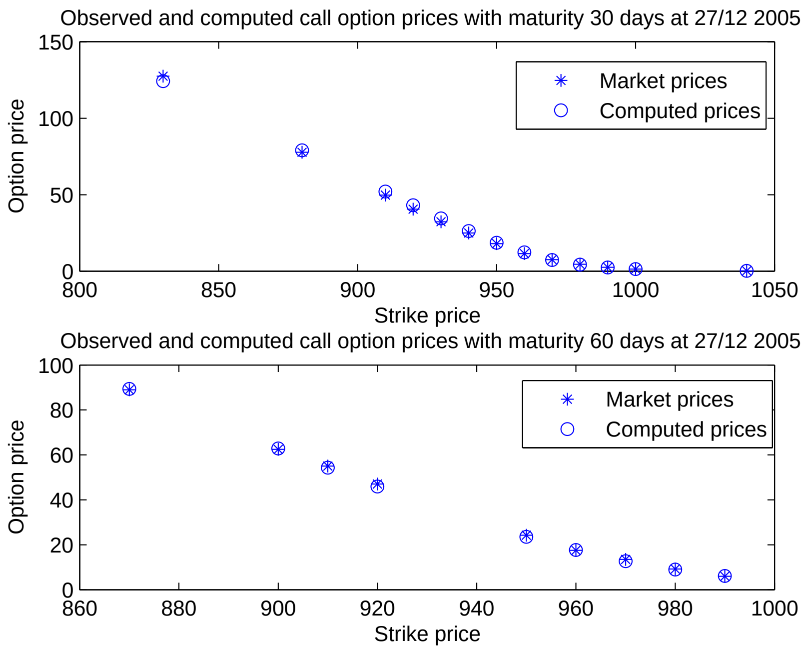

Figure 5: Theoretical prices from a two-state model (rings) and observed market prices (stars) for call options on OMXS30 index on the 27th of December 2005. The top panel show the prices for the contracts with maturity in one month and the bottom panels show the prices for the contracts with maturity in two months.

Figure 5 represents the prices of the contracts used for calibrating the model at this date. We see that the calibrated model seems to replicate the observed market prices with a high degree of accuracy. Plots of the other estimated models display similar appearances. From the results obtained here I conclude that the in-sample pricing performance of all models seems to be very good and taking into account the number of parameters used, the two-state model should be favoured.

6.3 Out-of-sample performance

The out-of-sample pricing performance has been evaluated as follows. For each calibrated model the theoretical price of the selected benchmark at-the-money option with one month to maturity have been computed. In the next step, the percentage deviation of the theoretical price from the observed option price has been calculated and recorded. This procedure has been repeated for all estimation dates.[43] As a comparison we have also computed a Black-Scholes price of the benchmark option. This was done similarly to the regime-switching model by backing out the implied volatility from the prices of the in-sample options.[44] However, in the Black-Scholes case we only have to back out a single volatility parameter. More formally we have solved

$$\{\hat{\sigma}_t^*\} = arg \min_{\sigma_t^*} \sum_{i=1}^{N_t} \left(\Pi^i_{B\&S}(t, s_t | \sigma_t^*) - \Pi^{i*}(t) \right)^2,$$

where N_t denotes the total number of contracts used for estimating the model at day t, $\Pi^i_{B\&S}(t, s_t)$ is the standard Black-Scholes call price of option i and $\Pi^{i*}(t)$ is the observed market price of option i at date t. Thus, the value $\hat{\sigma}_t^*$ is the constant volatility which should make the theoretical Black-Scholes price of all contracts as close as possible to observed prices. Using this volatility as input in the Black-Scholes formula I have then computed the theoretical Black-Scholes price of the benchmark option and recorded the percentage deviation from the actual price. If we believe that we live in a Black-Scholes world where the volatility is constant across moneyness and time to maturity this approach should price the benchmark option correctly.

In Figure 6 the average absolute percentage deviations from the price of the benchmark option are shown for each model: Black-Scholes denotes the Black-Scholes model, 2S denotes the two state model, 3S denotes the three state model and 4S denotes the four state model. From this Figure we can distinguish two important observations. First, the Black-Scholes model produces considerably larger mispricings than any of the regime-switching models. This indicates that the regime-switching volatility is indeed more consistent with observed market prices than a constant volatility. Second, the pricing performances of the regime-switching models are more or less equal. Thus, adding more states will not increase the pricing performance of the model. On the contrary, we see that the average deviation of theoretical prices from observed prices is slightly larger in the four-state model than in the three-state model or the two-state model. This

[43] See Section 5 for a more detailed description of the out-of-sample evaluation of the pricing performance.

[44] See Appendix II.

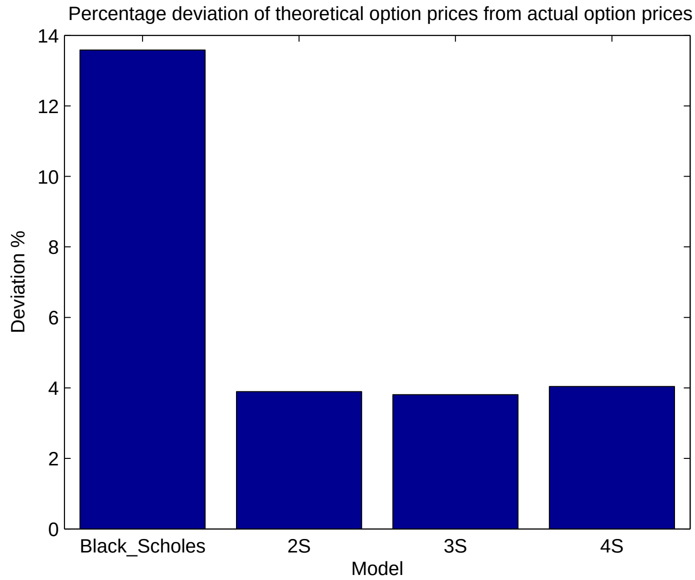

Figure 6: Average percentage deviations of theoretical prices from observed prices of the benchmark options.

is most likely due to instability and numerical inaccuracies in the calibration algorithm caused by the relatively larger number of parameters in this model. As discussed above, a large number of parameters makes the parameter surface[45] flatter and therefore it is much more difficult for the calibration algorithm to find an optimal solution. Often, when calibrating the four-state models, the calibration did not end until the maximum number of iteration was reached, indicating that an optimal solution could not be found.

6.4 Stability

In this study we have calibrated the model to observed option prices by solving a nonlinear quadratic optimization problem.[46] This approach of calibrating the parameters differs from other estimation techniques such as regression or maximum-likelihood methods in the respect that the underlying data, the option prices, are not considered to be a realization of some random data generating process. Instead, at a given point in time option prices are deterministically determined by the current interest rate, the stock price and other contract parameters. Consequently, we can not say anything about the statistical significance of the parameters in the usual probabilistic sense. Rather, in this setup the relevant question is whether the parameters are stable or not. Two types of stability are of particular interest.

[45]The parameter surface for a K-state model is the K-dimensional hyper-surface showing how the value of the objective function depends on the choice of the input parameters. Here, the objective function is the sum of squared deviations of theoretical prices from observed prices and the input parameters are the volatilities and the intensities, see Appendix II.

[46]See Appendix II.

1. *Stability of the parameters over time.*

2. *Stability of the optimal parameters at a given point in time.*

The first point focuses on how much the parameters fluctuate over time. If the parameters vary quite a lot over time, this indicates that the model does not provide a true description of the $\mathcal{Q}$-dynamics since the model relies on the assumption that the parameters are constant in time. Since a two-state model was found to be sufficient to replicate market prices in- and out-of-the sample, I have choosed to focus on the two-state model only. Plots of the parameters of the two-state model are shown in Figure 7. We see that the parameters

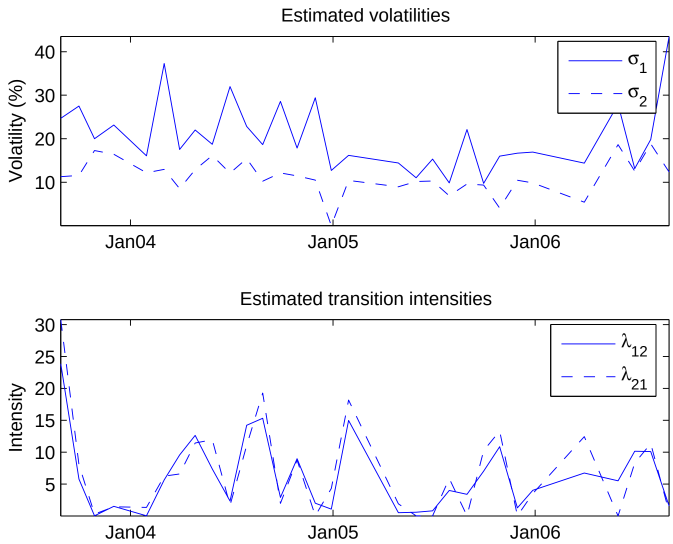

Figure 7: The upper panel shows the volatilities in the two states for the two-state model over the time period studied. The bottom panel shows the corresponding plot for the transition intensities.

have varied quite a lot over time. This indicates that the model is not true in the more fundamental sense. However, the model might still be used to price non-traded contracts as that price would be consistent with the prices of other contracts traded on the market at a given point in time.

The second point focuses on whether the obtained parameters are the only choices of parameter values which attain a reasonable in-sample fit. It could be the case that there exist several choices of parameter values, all providing a reasonable fit to data. In other words, since the calibrated parameters are obtained by solving a nonlinear optimization problem we are not guaranteed that the obtained solution is indeed a unique and global solution. This issue can be investigated by varying the start values of the parameters used to initialize the calibration algorithm.[47] If the optimal parameters are sensitive to

[47]See Appendix II.

changes in the start values this indicates that there are several local optima and thus the optimal parameters will be unstable. I have confined the analysis for the two-state model to a single date, the 27th of December 2005, which was selected randomly amongst the total number of calibration dates. Since the models at different dates are conceptually the same they should not differ from each other in terms of their stability characteristics. A single date should be enough to analyse the stability of the solution. At the 27th of December 2005 the start values of the volatilities were set to 17% and 10% in the high volatility regime and the low volatility regime respectively, and the start values of the intensities were set to 12, in accordance with the method described in Section 7. The optimal parameters with these start values were obtained as $\hat{\sigma}_1 = 15.5\%$, $\hat{\sigma}_2 = 10.4\%$, $\hat{\lambda}_{12} = 4.7$ and $\hat{\lambda}_{21} = 4.2$. To analyse the stability for various choices of start values I have increased and decreased the volatilities and the intensities by five units in each direction as indicated in Figure 8 below.

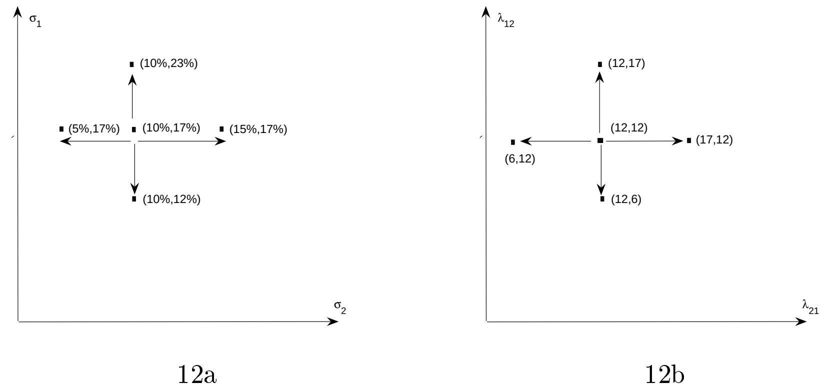

Figure 8: Figure 12a shows the start values for the volatilities and Figure 12b shows the start values for the intensities.

To analyse the effect of changing the start values for the volatilities and the intensities separately, I kept the intensities unchanged while changing the volatilities and vice versa. Calibrating the model for these start values generated 8 additional solutions corresponding to the 8 branches in Figure 12a and 12b. These were all found to be equal to the solution for the original start values. Thus, the results indicate that the optimal parameters at a given point in time is rather robust to moderate changes in the start values.

7 Conclusions

The objective of this thesis was to show how options can be priced under regime-switching volatility and examine the issue of calibrating this model to observed option prices. A numerical solution to the pricing equations for European options was derived and the calibration algorithm was formulated in terms of an optimization problem. The model was calibrated for two, three and four states to call options on the OMXS30 index with maturity within two months during the period August 2003 to August 2006. The findings show that a regime-switching model can replicate observed market prices with a high degree of accuracy. The model produces smaller mispricings than the Black-Scholes model. This indicates that regime-switching volatility is indeed more consistent with observed option prices than a constant volatility. Furthermore, the pricing performance does not seem to improve as the number of regimes is increased. In particular, a two-state model generates almost identical prices as a model with three or four states. Since the two-state model is more parsimonious than the other models, stable estimates are more easily obtained. Therefore, I argue that the two-state model is more advantageous. The parameters in this model were also found to be stable for moderate changes in the start values. However, the parameters vary substantially over time thus violating the assumption of constant parameters. This indicates that the model does not provide a true description of the $\mathcal{Q}$-dynamics for the volatility process.

The framework for calibrating the regime-switching model proposed in this thesis should be considered as a reference for further research. In particular, one could explore the stability of the estimates in the two-state model in more detail. Moreover, the pricing performance to other types of payoffs, e.g. digital options, or underlyings could be studied in more depth. Yet another interesting aspect is to compare the parameters implied out of market prices to parameters estimated from asset return data using some maximum-likelihood technique. If the two models yield similar parameters this should indicate that the market base future expectations on the volatility process on past behaviour. Overall, more back testing and validation of the model should be carried out before implementation. However, once this has been done the model can run quickly and efficiently on a daily basis.

References

[1] Alexander, C. (2001) *Market Models: A guide to financial Data Analysis.* West Sussex: Wiley & Sons.

[2] Avery, C. and Zemsky, P. (1998) Multidimensional Uncertainty and Herd Behavior in Financial Markets. *American Economic Review*, Vol 88, pp. 724-748.

[3] Bakshi, G., Cao, C. and Chen, Z. (1997) Empirical performance of alternative option pricing models. *Journal of Finance*, Vol 53, pp. 499-547.

[4] Bikhchandani, S. and Sharma, S. (2000) Herd Behavior in Financial Markets: A Review. *IMF Working paper. IMF Institute.*

[5] Björk, T. (2004) *Arbitrage Theory in Continuous Time.* 2nd Ed. Oxford: Oxford University Press.

[6] Black, F. and Scholes, M. (1973) The pricing of options and corporate liabilities. *Journal of Political Economy*, Vol 81, pp. 637-654.

[7] Bollen, N. (1998) American options with regime switching. *Journal of Derivatives*, Vol 5, pp. 497-514.

[8] Boyce, W. & DiPrima, R. (2000) *Elemantary Differentail Equations and Boundary Value Problems.* 7th Ed. New York: Wiley& Sons.

[9] Buffington, J. and Elliott J. (2002) American options with regime switching. *International Journal of Theoretical and Applied Finance*, Vol 5, pp. 497-514.

[10] Chan, L., Elliott, R. & Siu T. (2005) Option pricing and Esscher transform under regime switching. *Annals of Finance*, Vol 1, pp. 423-432.

[11] Chourdakis, K. (2002) Continuous time regime switching models and applications in estimating processes with stochastic volatility and jumps. *Working Paper. University of London.*

[12] Coleman, T.F. and Li, Y. (1996) An Interior Trust Region Approach for Nonlinear Minimization Subject to Bounds. *SIAM Journal of Optimization*, Vol 6, pp. 418-445.

[13] Courtadon, G. (1982) A More Accurate Finite Difference Approximation for the Valuation of Options. *The Journal of Financial and Quantitative Analysis*, Vol. 17, pp. 697-703.

[14] Di Masi, G., Kabanov, Y. and Runggaldier, W. (1993) Mean-variance hedging of options on stocks with markov volatilities. *Working Paper. Univesity of Padova.*

[15] Dibda, B. and Grossman, H. (1988) The Theory of Rational Bubbles in Stock Prices. *The Economic Journal*, Vol 98, pp. 746-754.

[16] Duan, J., Popova, I. and Ritchken, P. (2002) Option pricing under regime switching. *Quantitative Finance*, Vol 2, pp. 116-132.

[17] Dumas, B., Fleming, J. and Whaley, B. (1998) Implied Volatility Functions: Empirical Tests. *Journal of Finance*, Vol 111, pp. 2059-2106.

[18] Elliott, R., Aggoun, L. and Moore, J. (1994) *Hidden Markov Models - Estimation and Control*. Berlin: Springer Verlag.

[19] Enger, J. and Grandell, J. (2003) *Markovprocesser och köteori.* Lecture notes. Royal Institute of Technology, Stockholm.

[20] Engle, C. and Hamilton, J. (1990) Long svings in the dollar: are they in the data and do the market know it? *American Economic Review*, vol. 80, 689-713.

[21] Foque. J-P., Papanicolaou, G. and Sircar, K. (2000) *Derivatives in Financial Markets with Stochastic Volatility*. New York: Cambridge University Press.

[22] Gujarati, D. (2003) *Basic Econometrics*, 4th Ed. New York: McGraw-Hill.

[23] Grinblatt, M. and Titman, S. (2002) *Financial Markets and Corporate Strategy*, 2nd Ed. New-York: McGraw-Hill.

[24] Guo, X. (2001) Information and option pricing. *Quantitative Finance*, Vol 1, pp. 38-44.

[25] Hamilton, J. (1989) A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, Vol 57, pp. 357-384.

[26] Hamilton, J. and Susmel, R. (1994) Autoregressive conditional heteroscedasticity and changes in regime. *Journal of Empirical Finance*, Vol 11, pp. 279-289.

[27] Hardy, M. (2001) A regime-switching model of long-term stock returns, *North American Actuarial Journal*, Vol 5, pp. 41-53.

[28] Harrison, J. and Kreps, M. (1979) Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory*, Vol. 20, pp. 381-408.

[29] Harrison, J. and Pliska, R. (1983) A stochastic calculus model of continuous trading: complete markets. *Stochastic Processes and Their Appplications*, Vol 15, pp. 313-316.

[30] Heath, M. (2002) *Scientific Computing - An Introductory Survey*. New York: McGraw-Hill.

[31] Hull, J. (1999) *Options, Futures & Other Derivatives*. 4th Ed. New York: Prenctice-Hall.

[32] Iserles, A. (2003) *A First Course in the Numerical Analysis of Differential Equations*, Cambridge: Cambrige University Press.

[33] Mamon, R. and Rodrigo, R. (2004) Explicit solutions to European options in a regime-switching economy. *Operation Research Letter 00/06*, Vol 33, pp. 581-586.

[34] Tsay, R. (2002) *Analysis of Financial Time Series.* New York: Wiley& Sons.

[35] Wu, S. and Zeng, Y. (2004) Affine regime-switching models for interest rate term structure. *Mathematics of Finance*, Vol 351, pp. 375-386.

[36] Wilmott, P. (1995) *The Mathematics of Financial Derivatives - A Student Introduction.* Cambridge: Cambridge University Press.

[37] Öksendal, B., (1985) *Stochastic differential equations: an introduction with applications.* New York: Springer-Verlag.

Appendix I - Solving the pricing equations

Below we show how the pricing equations can be solved by numerical means. The general concept underlying numerical solution techniques are discussed and the method employed is described in some detail. The quintessence of this Appendix is that the pricing equations can be solved using numerical approximations and the solution technique employed is called the fully-implicit method. I will primarily focus on solving the pricing equations for European call options on the underlying.

Discretization of the pricing equations

To price the derivative in a K-state model we have to solve the coupled system of partial differential equations

$$-rF_k + \frac{\partial F_k}{\partial t} + rs\frac{\partial F_k}{\partial s} + \frac{1}{2}\sigma_i^2 s^2 \frac{\partial^2 F_k}{\partial s^2} - \lambda_{kk}F_k + \sum_{m \neq k} \lambda_{mk}F_m = 0, \quad k = 1, ..., K. \tag{13}$$

Since this system can not be solved analytically, as in the Black-Scholes case, we have to solve the equations numerically. To solve the equations numerically means that we convert the partial differential equations into a set of difference equations and then solve the difference equations iteratively. By working with difference equations instead of differential equations we only compute the solution of the equations at a finite number of nodes in time and space. However, if the number of nodes are sufficiently large the accuracy of the solution will be very good.[48] Before attempting to solve the system of equations (12) we first transform them into a more computationally efficient form by setting $x = \ln(s)$ and $u^k(t,x) = F_k(t,s)$. Taking partial derivatives of $u^k(t,x)$ we get

$$\begin{aligned} \frac{\partial F_k}{\partial t} &= \frac{\partial u^k}{\partial t}, \\ \frac{\partial F_k}{\partial s} &= \frac{\partial u^k}{\partial x}\frac{1}{s}, \\ \frac{\partial^2 F_k}{\partial s^2} &= \frac{\partial^2 u^k}{\partial x^2}\frac{1}{s^2} - \frac{\partial u^k}{\partial x}\frac{1}{s^2}. \end{aligned}$$

Inserting this into equation (13) we get

$$-ru^k + \frac{\partial u^k}{\partial t} + rs\frac{\partial u^k}{\partial x}\frac{1}{s} + \frac{1}{2}\sigma_k^2 s^2 \Big(\frac{\partial^2 u^k}{\partial x^2}\frac{1}{s^2} - \frac{\partial u^k}{\partial x}\frac{1}{s^2}\Big) - \lambda_{kk}u^k + \sum_{m \neq k}\lambda_{mk}u^m = 0.$$

Rearranging and cancelling terms the transformed equation becomes

$$-ru^k + \frac{\partial u^k}{\partial t} + \Big(r - \frac{1}{2}\sigma_k^2\Big)\frac{\partial u^k}{\partial x} + \frac{1}{2}\sigma_k^2\frac{\partial^2 u^k}{\partial x^2} - \lambda_{kk}u^k + \sum_{m \neq k}\lambda_{mk}u^m = 0, \tag{14}$$

[48] For an introduction to numerical solutions to partial differential equations see Heath (2002) [30] and Iserles (2003) [32] and for applications to financial problems see e.g Wilmott (1995) [36], Courtadon (1982) [13] or Hull (1999) [31].

for $k = 1, ..., K$. The first step in a finite-difference approximation is to define a mesh of points in time and space where the equations are to be evaluated. We begin by denoting the time to maturity of the derivative in fraction of years by T. The time to maturity is divided into N equally spaced intervals of length $\delta t = T/N$ so we obtain a partition in time as $0, \delta t, 2\delta t, ..., T$. Next, suppose that we can take a sufficiently large stock price S_{max} that when the stock price reaches this level we are more or less certain that the call option will expire in-the-money and the value of the option equals that of a forward[49]

$$u(X_{max}, t) = S_{max} - Ke^{-r(T-t)},$$

where $X_{max} = \ln(S_{max})$. The lowest possible price that the stock price can take is denoted S_{min} where the value of the call option is zero

$$u(X_{min}, t) = 0,$$

where $X_{min} = \ln(S_{min})$. Of convenience I let $X_{min} = 0$ so the lowest possible stock price is 1.[50] We divide the range $[0, X_{max}]$ into M equally spaced intervals by setting $\delta x = X_{max}/M$ and the partition in space becomes $0, \delta x, 2\delta x, ..., X_{max}$. Consequently, the stock is only allowed to take on values from the set of prices $\{e^0, e^{\delta x}, e^{2\delta x}, ..., e^{X_{max}}\}$. This highlights the benefit of the transformation carried out above since the attainable prices are closer together when the price is low and more sparsely distributed for higher prices. This is a desirable feature since real world stock price movements tend to be a lot smaller in absolute terms when the price of the underlying is low as compared to when the price of the underlying is high. To sum up, the partition in time and space divides the (x, t) plane into a mesh, where the mesh points have the form $(i\delta t, j\delta x)$. We concern ourselves only with the values of $u(x, t)$ at the mesh points $(i\delta t, j\delta x)$, for $i = 0, 1, ..., N$ and $j = 0, 1, ..., M$. The value at node $(i\delta t, j\delta x)$ is denoted

$$u_{i,j}^k = u^k(i\delta t, j\delta x).$$

Figure 9 displays a schematic picture of the mesh and the node (i, j).

[49] We can think of this as that the value of choice of the option is effectively zero. (Hull, 2000) [31].

[50] In principle, the price of the option could of course be zero but this choice of lower boundary on X has no influence on the results presented in this thesis since we consider options on the OMXS30 index where the underlying is about 1000.

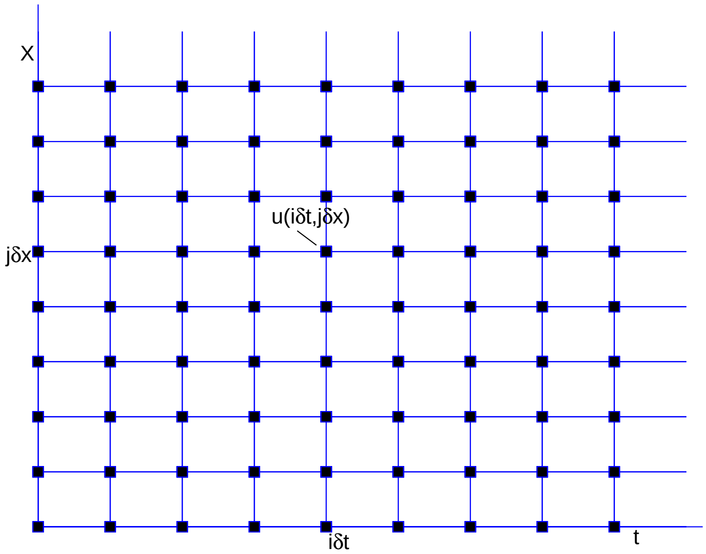

Figure 9: The mesh for a finite-difference approximation.

Numerical methods

The idea underlying finite-difference methods is to replace the partial derivatives occurring in the partial differential equations by approximations of the function $u^k(t,x)$ near the point of interest. For example one can take the partial derivative $\partial u/\partial x$ to be

$$\frac{\partial u}{\partial x}(t,x) \approx \frac{u(t,x+\delta x)-u(t,x)}{\delta x},$$

which is called a forward-difference since we take a small step δx forward. Another choice is

$$\frac{\partial u}{\partial x}(t,x) \approx \frac{u(t,x)-u(t,x-\delta x)}{\delta x},$$

which is called a backward-difference. We can also define central-differences by noting that

$$\frac{\partial u}{\partial x}(t,x) \approx \frac{u(t,x+\delta x)-u(t,x-\delta x)}{2\delta x}.$$

Figure 10 shows a geometric interpretation of these approximations. The two most common finite-difference approximations are the explicit method and the fully implicit method. Both the explicit and the fully implicit method uses central-differences for $\partial u/\partial x$ and $\partial^2 u/\partial x^2$. However, the methods differ regarding the difference of $\partial u/\partial t$ where the explicit method uses the forward-difference and the fully implicit method uses the backward-difference. The benefit of the explicit method is that one can compute the value of the function at time t explicitly by using values at time $t+\delta t$. Since we know the value of the option at

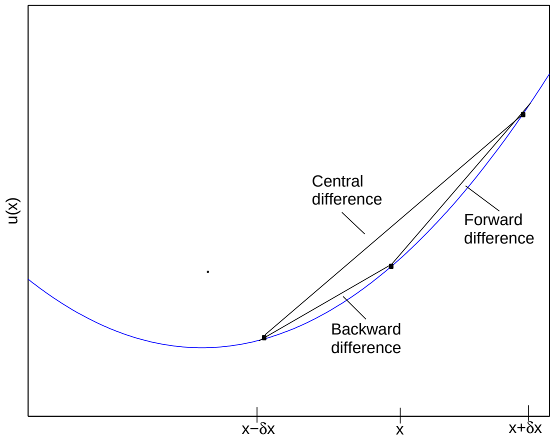

Figure 10: Forward-, backward- and central-difference approximations.

expiration we can work our way back in time by successively computing values at the previous time points. As a consequence, the explicit method is very fast. However, the drawback is that the method is unstable if[51]

$$\frac{\delta t}{\delta x^2} < \frac{1}{2}$$

This puts a severe constraint on the size of the time-steps. The fully implicit method has the advantage that it is always stable for any choices of δx and δt. In the implicit method the value at the node (i, j) is defined in terms of the values at the neighbouring points $(i, j+1)$, $(i, j-1)$ and $(i+1, j)$. As we will see below this will require us to solve a large system of linear equations to obtain the solution at each time point. Thus, the drawback of the fully implicit method is that it requires more computations in each step which slows up the solution algorithm. However, the implicit method allows us to take larger time-steps than we can when using the explicit method and the need for fewer time-steps in the implicit method more than compensates for the larger number of computations that has to be performed in each step. In addition, when calibrating a model to observed prices, stability as well as numerical accuracy should be two very important criteria on which to evaluate a solution algorithm.[52] Thus, we argue that the fully implicit method should be the proper choice for the application considered here.[53]

[51] See Heath (2002) [30] or Wilmott (1995) [36].

[52] This is because the calibration will require us to minimize some distance measure between computed and observed prices and when the number of parameters that are to be estimated are large, the optimization may not converge properly if the objective function is not accurate enough.

[53] There exists higher order methods, e.g. Crank-Nicolson, which converge faster than the fully implicit method. However, when comparing the Crank-Nicolson method with the fully implicit method significant difference in accuracy of the methods could be discerned. I argue that the simplicity of the fully implicit method makes it more advantageous.

The fully implicit method

In the fully implicit method the $\partial u/\partial t$-term is approximated with a backward-difference and $\partial u/\partial x$ and $\partial^2 u/\partial x^2$ are discretized with central-differences [54]

$$\begin{aligned}
\frac{\partial u_k}{\partial t} &= \frac{u^k_{i+1,j} - u^k_{i,j}}{\delta t},\\
\frac{\partial u_k}{\partial x} &= \frac{u^k_{i,j+1} - u^k_{i,j-1}}{2\delta x},\\
\frac{\partial^2 u_k}{\partial x^2} &= \frac{u^k_{i,j+1} + u^k_{i,j-1} - 2u_{i,j}}{\delta x^2}.
\end{aligned}$$

Inserting these approximations into the partial differential equation we obtain

$$\begin{aligned}
&-ru^k_{i,j} + \frac{u^k_{i+1,j} - u^k_{i,j}}{\delta t} + \Big(r - \frac{1}{2}\sigma_k^2\Big)\Big(\frac{u^k_{i,j+1} - u^k_{i,j-1}}{2\delta x}\Big)\\
&+ \frac{1}{2}\sigma_k^2\Big(\frac{u^k_{i,j+1} + u^k_{i,j-1} - 2u_{i,j}}{\delta x^2}\Big) - \lambda_{kk}u^k_{i,j} + \sum_{m\neq k}\lambda_{mk}u^m_{i,j} = 0.
\end{aligned}$$

Rearranging yields

$$\begin{aligned}
&u^k_{i,j-1}\Big(-\frac{r-\sigma_k^2/2}{2\delta x} + \frac{\sigma_k^2}{2(\delta x)^2}\Big) + u^k_{i,j}\Big(-\frac{1}{\delta t} - \frac{\sigma_k^2}{(\delta x)^2} - \lambda_{kk} - r\Big)\\
&+ \sum_{m\neq k} u^m_{i,j}\lambda_{mk} + u^k_{i,j+1}\Big(\frac{r-\sigma_k^2/2}{2\delta x} + \frac{\sigma_k^2}{2(\delta x)^2}\Big) + u^k_{i+1,j}\frac{1}{\delta t} = 0.
\end{aligned}$$

Solving for $u^k_{i+1,j}$ gives

$$u^k_{i+1,j} = \alpha^k u^k_{i,j-1} + \beta^k u^k_{i,j} + \sum_{m\neq k}\gamma^{mk}u^m_{i,j} + \delta^k u^k_{i,j+1}, \tag{15}$$

for $i = 1, ..., N-1$ and $j = 1, .., M-1$, where

$$\begin{aligned}
\alpha^k &= \delta t\Big(\frac{r-\sigma_k^2/2}{2\delta x} - \frac{1}{2}\frac{\sigma_k^2}{\delta x^2}\Big),\\
\beta^k &= \Big(1 + \frac{\delta t}{\delta x^2}\sigma_k^2 + (r+\lambda_{kk})\delta t\Big),\\
\gamma^{mk} &= -\lambda_{mk}\delta t,\\
\delta^k &= \delta t\Big(-\frac{(r-\sigma_k^2)}{2\delta x} - \frac{1}{2}\frac{\sigma_k^2}{\delta x^2}\Big).
\end{aligned}$$

The interior points of the mesh are computed with equation (15). The values at the boundary points $x = M\delta x$, $x = 0$ and $T = N\delta t$ are computed using the appropriate values of the option on these points. For a call option we have

$$\begin{aligned}
u^k_{i,M} &= S_{max} - Ke^{-r(T-i\delta t)}, && i = 1, ..., N \text{ and } k = 1, ..., K,\\
u^k_{i,0} &= 0, && i = 1, ..., N \text{ and } k = 1, ..., K,\\
u^k_{N,j} &= \max(e^{j\delta x} - K, 0), && j = 1, ..., M \text{ and } k = 1, ..., K.
\end{aligned}$$

[54] The central difference-method for the $\partial^2 u/\partial x^2$ term implies taking central differences of the $\partial u/\partial x$ term.

From equation (15) one can see that the values at the nodes $u_{i,j-1}$, $u_{i,j}$, $u_{i,j+1}$ and $u_{i+1,j}$ are *implicitly* defined in terms of each other, thereby the name of the method. To solve for the values at time i we have to reformulate equation (15) into a system of linear equations. To make the exposition easier to follow we confine ourselves for the moment to two states only and define a vector $\mathbf{u}_i$ by

$$\mathbf{u}_i = (u^1_{i,M-1}, u^2_{i,M-1}, u^1_{i,M-2}, u^2_{i,M-2}, ..., u^1_{i,1}, u^2_{i,1})'.$$

The vector contains the value of the option in the different regimes at time i ordered by regime number and in descending order of spatial nodes. For a general K-state model, the first K positions in the vector contain the values of the option in the K regimes corresponding to a stock price of $e^{(M-1)\delta x}$. The final M positions contain the prices in the second bottom node corresponding to a stock price of $e^{\delta x}$. Going back to the two-state case we can now write the system of equations in (15) as

$$\underbrace{\begin{pmatrix}
\beta^1 & \gamma^{21} & \alpha^1 & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\
\gamma^{12} & \beta^2 & 0 & \alpha^2 & 0 & 0 & 0 & 0 & 0 & \dots \\
\delta^1 & 0 & \beta^1 & \gamma^{21} & \alpha^1 & 0 & 0 & 0 & 0 & \dots \\
0 & \delta^2 & \gamma^{12} & \beta^2 & 0 & \alpha^2 & 0 & 0 & 0 & \dots \\
0 & 0 & \delta^1 & 0 & \beta^1 & \gamma^{21} & \alpha_1 & 0 & 0 & \dots \\
0 & 0 & 0 & \delta^2 & \gamma^{12} & \beta^2 & 0 & \alpha^2 & 0 & \dots \\
\vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\
0 & 0 & 0 & 0 & \dots & 0 & \delta^1 & 0 & \beta^1 & \gamma^{21} \\
0 & 0 & 0 & 0 & \dots & 0 & 0 & \delta^2 & \gamma^{12} & \beta^2
\end{pmatrix}}_{=\mathbf{A}}
\underbrace{\begin{pmatrix}
u^1_{i,M-1} \\ u^2_{i,M-1} \\ u^1_{i,M-2} \\ u^2_{i,M-2} \\ u^1_{i,M-3} \\ u^2_{i,M-3} \\ \vdots \\ u^1_{i,1} \\ u^2_{i,1}
\end{pmatrix}}_{=\mathbf{u}_i}$$

$$= \underbrace{\begin{pmatrix}
u^1_{i+1,M-1} \\ u^2_{i+1,M-1} \\ u^1_{i+1,M-2} \\ u^2_{i+1,M-2} \\ u^1_{i+1,M-3} \\ u^2_{i+1,M-3} \\ \vdots \\ u^1_{i+1,1} \\ u^2_{i+1,2}
\end{pmatrix}}_{=\mathbf{u}_{i+1}}
+ \underbrace{\begin{pmatrix}
-\delta^1 u^1_{i+1,M} \\ -\delta^2 u^2_{i+1,M} \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ -\alpha^1 u^1_{i+1,0} \\ -\alpha^2 u^2_{i+1,0}
\end{pmatrix}}_{=\mathbf{b}}$$

The vector $\mathbf{b}$ is a constant vector that depends on value of the option at the upper and lower boundary points. The values of the option at all spatial nodes at specific time i can be obtained if we know the values at time $i+1$ by solving the system of equations

$$\mathbf{A}\mathbf{u}_i = \mathbf{u}_{i+1} + \mathbf{b}.$$

The solution to this one-step-ahead problem is

$$\mathbf{u}_i = \mathbf{A}^{-1}(\mathbf{u}_{i+1} + \mathbf{b}). \tag{16}$$

Thus, the values at time i are defined in terms of the values at time $i+1$. Since we know that the value of the option is equal to the payoff function at expiration

$i = N$, we can use equation (16) to obtain the values of the option at time $i-1$. Then, applying equation (16) once more we can obtain the solution at time $i-2$ and so on. In this way we can work backwards in time until we get the value of the option at time $i = 0$ as

$$\mathbf{u}_0 = \mathbf{A}^{-N}\mathbf{u}_N + \sum_{i=1}^{N} \mathbf{A}^{-i}\mathbf{b}. \tag{17}$$

To pick out the right price of the vector $\mathbf{u}_0$ we first sort the prices according to regime number in two different vectors $\mathbf{u}_0^1$ and $\mathbf{u}_0^2$, corresponding to regime 1 and 2 respectively,

$$\begin{aligned}\mathbf{u}_0^1 &= (u_{0,M-1}^1, u_{0,M-2}^1, ..., u_{0,1}^1),\\ \mathbf{u}_0^2 &= (u_{0,M-1}^2, u_{0,M-2}^2, ..., u_{0,1}^2).\end{aligned}$$

If the current price of the underlying is S_0 we can compute the position in each vector corresponding to this price as $M_0 = \ln(S_0)/\delta x$. Thus if the volatility is in state k we get the price of the option as

$$F_k(t, s_t) = u_{0,M-M_0}^k.$$

The solution for an arbitrary number of K states is very much similar to the two-state case.

To determine the appropriate values for δx and δt to use when calibrating the model, I have priced a standard call option with zero intensities for different values of δx and δt and compared the price with the price obtained from the standard Black-Scholes formula. Since the intensities have been set to zero the model is equivalent with the Black-Scholes model and therefore the numerically computed prices should be equal to the prices computed with the Black-Scholes formula. If the algorithm is accurate then the deviation from the Black-Scholes price should be small. To examine the convergence of the numerically computed prices to the Black-Scholes price I have considered a call option on an underlying with a current price of $S_0 = 100$, a strike of $K = 95$, time to expiration of $T = 0.5$ years, a risk-free interest rate of $r = 10\%$ and a volatility of $\sigma = 50\%$. The Black-Scholes price of this option is 18.71. The maximum price in the numerical solution was set to $S_{max} = 5S_0$. Figure 11 displays the absolute deviations between the numerically computed prices with the implicit method $Call_{impl}$ and the value computed with the Black-Scholes formula $Call_{bs}$ for different values of δx and δt. The top panel show the convergence to the Black-Scholes price when δt is decreased. The different lines in this graph represents the price deviations for different choices of δx when varied between 0.0031 and 0.031. We can see that for each value of δx in this intervall is the pricing error with the implicit method less than 0.015 SEK. The bottom panel shows the convergence to the Black-Scholes price when δx is decreased. The different lines in this graph represents the price deviations for different choices of δt when varied between 0.0005 and 0.01. We see that for each value of δt in this intervall is the price the pricing error less than 0.015 SEK. Furthermore, for δx less than 0.02 and δt less than 0.005, the error is smaller than 0.005 SEK and we have accuracy in the second decimal. Since prices are not quoted in more than two

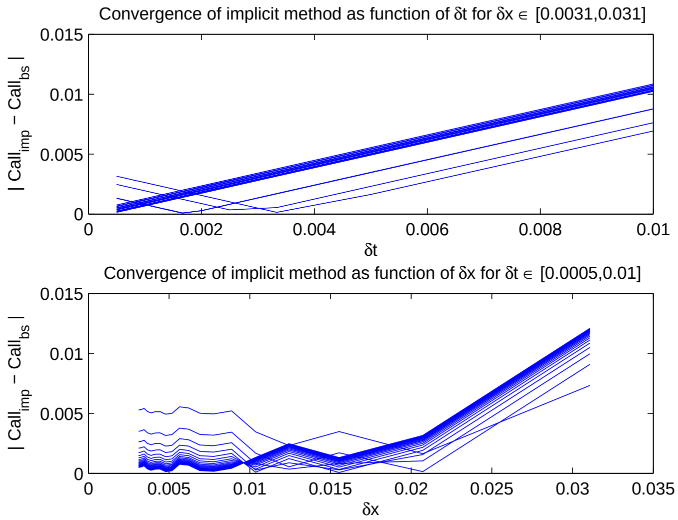

Figure 11: The top panel show convergence of the fully implicit method to the Black-Scholes price as function of δt when $\delta x \in [0.0031, 0.031]$. The bottom panel show the convergence of the fully implicit method to the Black-Scholes price as function of δx when $\delta t \in [0.0005, 0.01]$.

decimals there is no reason to consider smaller values of δx and δt. Thus, we argue that choosing $\delta t = 0.005$ and $\delta x = 0.01$ should be sufficient to obtain accurate results.

Appendix II - Calibration

The objective of the calibration is to determine the volatilities and the transition intensities so that the theoretical prices are, by some measure, as close as possible to observed market prices. The criterion adopted here was to minimize the sum of squared errors between actual and theoretical prices. This criterion has been employed in several studies to calibrate various option pricing models.[55] To facilitate the exposition we express the parameters for a K-state model in matrix notation as

$$\boldsymbol{\sigma} = (\sigma_1, ..., \sigma_K),$$
$$\mathbf{Q} = \begin{pmatrix} -\lambda_{11} & \lambda_{12} & \dots & \lambda_{1K} \\ \lambda_{21} & -\lambda_{22} & \dots & \lambda_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{K1} & \lambda_{K2} & \dots & -\lambda_{KK} \end{pmatrix}.$$

To calibrate the model at a given day t we collect a number of option prices $\Pi^{i*}, ..., \Pi^{N*}$ with strikes $K_i, ..., K_N$ and maturities $T_1, ..., T_n$. Given the prices of the market contracts, the value of the underlying s_t and the current regime e_c, the solution to the following problem gives the optimal values of the parameters.

$$\{\hat{\boldsymbol{\sigma}}, \hat{\mathbf{Q}}\} = arg \min_{\boldsymbol{\sigma},\mathbf{Q}} \sum_{i=1}^{S} \left(\Pi_c^i(t, s_t | \boldsymbol{\sigma}, \mathbf{Q}) - \Pi^{i*} \right)^2,$$

where $\Pi_c^i(t, s_t)$ is the theoretical price of option i, and subject to the restrictions

$$\begin{aligned} \lambda_{ii} &= \sum_{j \neq i}^{K} \lambda_{ij} \quad \text{for } i = 1, ..., K, \\ \lambda_{ij} &\geq 0 \quad \text{for } i \neq j, i, j = 1, ..., K, \\ \sigma_i &> 0 \quad \text{for } i = 1, ..., K. \end{aligned}$$

In other words, we try to find the transition intensities and the volatility levels by making the in-sample fit as large as possible. The positivity constraints on the intensities and the volatilities are however necessary since the volatilities is not allowed to be negative and the exponential distribution is not defined for intensities smaller or equal to zero. To avoid imposing these constraints directly on the parameters we make the following transformations

$$\begin{aligned} \sigma_i &= e^{\sigma_i^*} \quad \text{for } i = 1, ..., K \\ \lambda_{ij} &= e^{\lambda_{ij}^*} \quad \text{for } i \neq j, i, j = 1, ..., K. \end{aligned}$$

Instead of maximizing directly over the volatilities σ_i and the intensities λ_{ij} which are only allowed to be greater or equal to zero we maximize over the transformed parameters σ_i^* and λ_{ij}^* which are defined everywhere on the real axis. Since the exponential function will ensure that the values of the true parameters are kept above zero no additional constraints are needed. The transformed

[55] See for example Duan (2002) [16], Bakshi (1997) [3] and Dumas (1998) [17].

problem becomes

$$\{\hat{\boldsymbol{\sigma}^*}, \hat{\mathbf{Q}^*}\} = arg \min_{\boldsymbol{\sigma}^*, \mathbf{Q}^*} \sum_{i=1}^{S} \left(\Pi_c^i(t, s_t | \boldsymbol{\sigma}^*, \mathbf{Q}^*) - \Pi^{i*} \right)^2,$$

where the elements in the matrices $\mathbf{Q}^*$ and $\boldsymbol{\sigma}^*$ have been replaced with λ_{ij}^* and σ_i^*, respectively. This makes implementation a lot easier since we efficiently avoid imposing any constraints on the optimization algorithm. The optimization problem was implemented and solved in Matlab which is a program specially devised for handling large vectors and performing matrix computations. I used the built-in function fminunc to solve the problem. The fminunc function take start values of the parameters as inputs and then solves the problem iteratively by updating the parameters in the direction where the decline in the target function is the greatest.[56] The calibration algorithm can be summarised as follows.

1. Set the current state of the regime to $e_c = e_1$ and guess the start values $\boldsymbol{\sigma}^*$ and $\mathbf{Q}^*$.
2. Compute the theoretical prices $\Pi_1^i(t, s_t | \boldsymbol{\sigma}^*, \mathbf{Q}^*)$ using the fully implicit method, for $i = 1, ..., N$.
3. Compute $\sum_{i=1}^{N} \left(\Pi_1^i(t, s_t | \boldsymbol{\sigma}^*, \mathbf{Q}^*) - \Pi^{i*} \right)^2$.
4. Use the fminunc function to find a new set of values $\boldsymbol{\sigma}^*$, $\mathbf{Q}^*$.
5. Repeat step 2 - 4 until the decrease in the sum of squared deviations fall below a given cut-off value and take the final values of $\boldsymbol{\sigma}^*$ and $\mathbf{Q}^*$ as the optimal estimates.[57]

The current regime is set equal to the first regime e_1 in step one. It is important to stress that this does not mean that the current volatility state will always be equal to the state with the highest volatility. Since we have not imposed any restrictions on how the volatilities in the different regimes are related to each other (i.e. bigger or smaller) these parameters are allowed to change freely. It could be the case that the current regime of the calibrated model is equal to the regime with the lowest volatility, i.e. that the state e_1 has the lowest volatility. By reordering the states in descending order of volatility levels we get the model on its standard form where the first regime corresponds to the highest volatility, the second regime corresponds to the second highest volatility and so on. By not imposing any constraints on how the volatilities are related to each other we efficiently avoid the problem of setting the correct current volatility state since this is determined within the calibration. Figure 12 gives a schematic illustration of the calibration algorithm.

To initialize the algorithm, we have to guess the start values of $\boldsymbol{\sigma}^*$ and $\mathbf{Q}^*$. I have set the start value of the volatility in the most volatile regime equal to the maximum of all the Black-Scholes implied volatilities for the contracts used for calibrating the model and the start value of the volatility in the least volatile

[56]The fminunc function uses a subspace trust region method and is based on the interior-reflective Newton method described in Coleman and Li (1996) [12].

[57]I have chosen the cut-off value to 10^{-6} and the maximum number of iterations to 500.

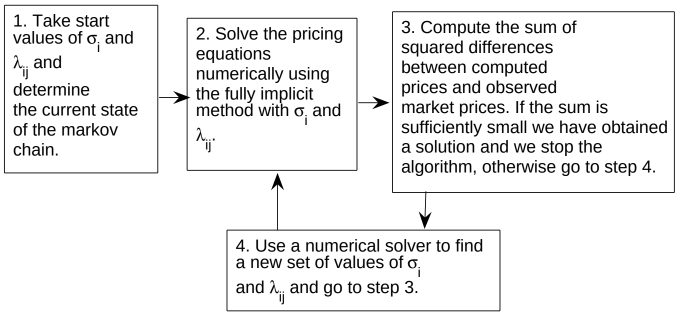

Figure 12: A schematic picture of the calibration method.

regime equal to the minimum value of the implied volatilities for the same set of contracts. The start values of the volatility in the intermediate regimes are placed equidistantly between the volatility in the highest and the lowest regime. More formally, I have chosen the start values $\sigma_1, ..., \sigma_K$ as

$$\begin{aligned}\sigma_1 &= \max(\sigma_1^{imp}, .., \sigma_N^{imp})\\ \sigma_K &= \min(\sigma_1^{imp}, .., \sigma_N^{imp}))\\ \sigma_i &= (\sigma_1^{imp} - \sigma_K^{imp})/(K-1), \quad \text{for } i = 2, ..., K-1,\end{aligned}$$

where $\sigma_1^{imp}, .., \sigma_N^{imp}$ denote the Black-Scholes implied volatilities corresponding to the observed market prices $\Pi^{1*}, ..., \Pi^{N*}$. This choice of start values should be reasonable since the volatilities in the regime-switching model and the Black-Scholes implied volatilities should be of the same magnitude.[58] Thus, one might expect that this choice of start values should result in small or moderate deviations of theoretical prices from observed prices and the calibration algorithm is therefore likely to converge quickly to the optimal solution. Of course, the speed of convergence also depends on the start values of the intensity parameters. Since we have no apriori good reason to prefer a particular choice of start values for the intensities parameters to another, I chose to set the value of all transition intensities equal. This approach can be considered to be unbiased as jumps between any two states are considered to be equally likely. The start value of the intensities have been set to 12. In Section 6.4 the stability of the calibrate parameters is investigated for moderate perturbations in the start values. If the parameters are found to be stable, picking the right start values is less of an issue.

[58] In Section 4.2 we saw that the prices in the regime-switching model and the Black-Scholes model are equal if the transition intensities are all zero.