Abstract and Executive Summary

Recently, negative interest rates have been observed more and more, making the existent pricing models for financial derivatives flawed or altogether useless.

Consequently, new solutions are needed. Some of them have been developed, and in this master thesis, we study them and compare them, noting that every model has its pros and cons, but also recommending one of them, namely the methodology of Hull and White.

The reason for this is that most proposed solutions understandably try to use existent methods that cannot cope with possible negative rates anymore, making appropriate alterations to them. That way, the adaptations for corporations remain limited. However, they struggle to find a satisfying solution.

The Hull and White Model used to have the possibility of negative rates as a con, but it has now become a pro, and we should simply embrace that fact and apply their model and the recent adaptations they made. That is our recommendation to the financial industry.
Contents

Abstract and Executive Summary ................................................. i

1 Introduction .......................................................................... 1

2 Preliminaries ........................................................................... 3

2.1 Introduction .......................................................................... 3

2.2 The Black-Scholes-Merton Model ........................................... 6

2.2.1 Itô’s Lemma .................................................................... 6

2.2.2 Risk-Neutral Valuation ..................................................... 8

2.2.3 Assumptions ................................................................. 9

2.2.4 Derivation of the Black-Scholes-Merton Model .................... 9

2.3 The SABR Model .................................................................. 10

2.3.1 Introduction .................................................................... 10

2.3.2 Notation .......................................................................... 11

2.3.3 Motivation for the Need of SABR .................................. 12

2.3.4 The Method ................................................................. 12

2.4 The Hull-White (Extended Vasicek) Model ............................... 13

2.4.1 Introduction .................................................................... 13

2.4.2 The Hull-White Model .................................................... 13

2.4.3 Using the Hull-White Model on Swaptions ...................... 14

2.4.4 Binomial Trees ............................................................. 15

2.4.5 Trinomial Trees ............................................................. 17

2.4.6 Conclusion ................................................................. 20

2.5 From Options to Swaptions in General ...................................... 20

2.5.1 Introduction .................................................................... 20

2.5.2 Definitions ................................................................. 20

2.5.3 Interest Rate Swaps ...................................................... 22

2.5.4 Swaptions ................................................................. 24

2.6 The Problem ......................................................................... 25

3 The Shifted SABR Model (2014) ................................................. 27

3.1 Introduction .......................................................................... 27

3.2 Gluing Terminal Densities .................................................... 27

3.3 Shifted Lognormal Low-Strike Extrapolation ............................ 29

3.4 Experiments and Conclusion ................................................ 29
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>The Free Boundary SABR (2015)</td>
<td>31</td>
</tr>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>31</td>
</tr>
<tr>
<td>4.2</td>
<td>The Model</td>
<td>31</td>
</tr>
<tr>
<td>4.3</td>
<td>Experiments and Conclusion</td>
<td>32</td>
</tr>
<tr>
<td>5</td>
<td>The Method of Hull and White (2014)</td>
<td>35</td>
</tr>
<tr>
<td>5.1</td>
<td>Introduction</td>
<td>35</td>
</tr>
<tr>
<td>5.2</td>
<td>A Generalized Tree Building Procedure</td>
<td>35</td>
</tr>
<tr>
<td>5.2.1</td>
<td>Model</td>
<td>35</td>
</tr>
<tr>
<td>5.2.2</td>
<td>New Properties</td>
<td>36</td>
</tr>
<tr>
<td>5.2.3</td>
<td>Description</td>
<td>36</td>
</tr>
<tr>
<td>5.2.4</td>
<td>Examples</td>
<td>37</td>
</tr>
<tr>
<td>5.3</td>
<td>Market Implied Volatility Function</td>
<td>39</td>
</tr>
<tr>
<td>5.4</td>
<td>Conclusion</td>
<td>40</td>
</tr>
<tr>
<td>6</td>
<td>Summary and Conclusion</td>
<td>41</td>
</tr>
<tr>
<td>7</td>
<td>Suggestions</td>
<td>43</td>
</tr>
<tr>
<td>8</td>
<td>Results</td>
<td>47</td>
</tr>
<tr>
<td>Bibliography</td>
<td></td>
<td>62</td>
</tr>
<tr>
<td>Appendix</td>
<td></td>
<td>62</td>
</tr>
<tr>
<td>A</td>
<td>P is the Derivative of the Receiver Swaptions Price with respect to K</td>
<td>63</td>
</tr>
<tr>
<td>B</td>
<td>The Jumping Probabilities Lie Between 0 and 1</td>
<td>65</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td></td>
<td>67</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

Existent pricing models for interest rate derivatives typically assume interest rates to be positive, and usually even reasonably large. This assumption has become less acceptable as low and even negative interest rates pop up in the market in recent years.

As a result, valuation software crashes, and the existing models that do not crash do not valuate properly anymore, also causing arbitrage possibilities. In the past, parameters could be plugged into working formulas involving the natural logarithm of a quotient of a rate and a strike, which were both assumed to be negative. The rate was assumed to be positive because negative rates were not only never observed, but also intuitively seemed unrealistic. The strike was assumed to be positive because the positivity of the rate implied that negative strikes made practical nonsense. After all, if rates are always positive, it makes no sense to trade it for a fixed negative rate - this just means that one of the two parties is paying twice.

Since negative rates appear more and more frequently, we need to consider their possibility, as well as negative strike derivatives.

As an example of what happens if we do not, suppose a corporation called Naiveness Inc. offers swaptions to potential buyers, using traditional pricing models. If they agree to pay some fixed rate and receive a floating rate, then they charge less and less for swaptions as the fixed rate approaches zero, since their pricing model assumes that the floating rate the customer has to pay Naiveness Inc. will remain positive.

However, there is a non-negligible possibility of the floating rate dropping below zero, so that the bank actually has to pay twice: receiving the floating rate has become equivalent to making a payment, since it is negative.

The customer pays almost no money for the swaptions, but can expect to extract a substantial amount of money from Naiveness Inc.

This illustrates the need for new pricing models in which negative interest rates are considered a possibility.

The central model assumption is the distribution of the rate. Traditionally, this distribution is defined on the positive numbers, allowing the use of the logarithm function. Altering the assumptions in such a way that we could still use the logarithm function in some way is a possible goal. Because the SABR method, which uses the logarithm function, is widely used within the software of corporations, it would be practically preferable to adapt that method in a consistent way to deal
with negative rates. To our disappointment, this will turn out to be problematic. We will present three possible solutions, all of which have their pros and cons. In the end, we will recommend the method of Hull and White.

For the convenience of the digital reader, we made sure that when we refer to a chapter, section or formula, clicking on the reference results in jumping to the position that was referred to.

In chapter 2, we present the necessary preliminaries we studied in order to properly understand the material, including the Black-Scholes-Merton Model, the SABR Model, the Hull and White Model and how to move from option valuation to swaption valuation in general.

In chapter 3, we present the Shifted SABR method, a simple method which might be used as a temporary solution, but seems to be too crude for long term use.

In chapter 4, we present the Free Boundary SABR method, another attempt to find a practical solution, which also turns out to have serious drawbacks.

In chapter 5, we continue on the path of Hull and White that we started in the preliminaries, present the results and explain why we believe this is the most promising model to deal with negative interest rates.

Chapter 8 consists of all the results, i.e. tables and graphs, that we refer to during the thesis.

Finally, we summarize and conclude our work and suggest further research.
Chapter 2

Preliminaries

2.1 Introduction

In order to understand the material in this thesis, it is necessary to have knowledge and understanding about the underlying financial products and the mathematical models involving them.

We missed this knowledge and understanding at the start of the internship, and started by studying [12] and [6] and occasionally looking up definitions at [1]. In these preliminaries, we will present the relevant parts. We will attempt to provide sufficient preliminaries for any master’s student mathematics to be able to read this thesis without frequently having to look up definitions.

After the necessary definitions, we will present the Black-Scholes-Merton Model, in combination with Itô’s Lemma and the concept of risk-neutral valuation. Then we move on to the SABR Model, which was invented in addition to the Black-Scholes-Merton Model to solve an inconsistency of the Black-Scholes-Merton Model. Combining these two models has been the main practice in valuating financial derivatives in the financial industry.

We will also show the Hull-White Model in combination with binomial and trinomial trees, which allows negative rates, a property which used to be regarded as a down-side of the model, but makes the model much more interesting in light of the recent negative rates developments. We will build upon this later in the thesis.

Finally, we will treat the transition from valuating options to valuation swaptions, showing the problem this thesis is about in more detail.

These preliminaries are somewhat extensive, which allows us to precisely understand the methods we will compare later. We will not only to study the results, but also grasp how the results came to be, which makes it much easier to make recommendations to the financial industry: we do not restrict ourselves to studying tables and graphs, but also look closely at the underlying mathematical concepts.

Options

Options are a specific type of derivative. A derivative is a product of which the value depends on an underlying asset: it is derived from it, hence the name derivative.

An option is a contract which gives the owner the right, but not the obligation, to buy or sell an underlying stock at or before a predetermined moment for a
predetermined price, the strike price $K$.

A call option provides the right to buy, whereas a put option provides the right to sell.

The most common options are European options and American options. A European option provides the right to buy or sell exactly on the agreed date, whereas an American option can be exercised at any time before or on the agreed date.

For example, suppose a certain European stock currently has value $1. Possessing a European call option to buy the stock in 10 days for $1.10 means hoping the value of the stock will rise above $1 in the next 10 days. If it does not, then the right to buy the stock is not exercised.

European options can be valued in order to derive the valuation of the American equivalents.

Valuation is possible with the celebrated and generally used Black-Scholes-Merton formula, which we will present later.

A digital option, or binary option, is an option where the pay-off is fixed if the stock price exceeds the strike price.

Interest rates

There are different types of interest rates. Relevant for this thesis will be:

- The risk free interest rate (usually denoted by $r$): this is the theoretical return rate of a risk-free portfolio (a portfolio is a set of products). So on an investment that in fact does contain risk, an investor will require to have an expected profit of at least the risk-free rate. After all, if his risk-containing investment had a lower expected return, then he could have avoided the risk and still have a higher return by simply setting up a risk-free portfolio. When valuing financial products which depend on uncertain future data, the risk-free interest rate always has to be taken into account, since this rate determines the value of money in the future (which is greater than its value now, if $r > 0$). Taking this $r$ into account is called discounting.

This rate appears in the definition of a bank account in [6]: $B(t)$ is defined to be the value of a bank account at time $t \geq 0$, behaving according to the following differential equation:

$$dB(t) = r_t B(t) dt, B(0) = 1,$$

where $r_t$ is a positive function of time. The solution of this differential equation is

$$B(t) = e^{\int_0^t r_u du}. \quad (2.2)$$

They refer to $r_t$ as the instantaneous spot interest rate or the short rate. This is the rate at which a bank account grows.

- Treasury rates: rates on Treasury Bills and Treasury bonds. These might be used for estimating the risk-free interest rate.
• EURIBOR: Euro Interbank Offer Rate: the rates offered to prime banks on euro interbank term deposits. The EURIBOR is determined periodically and based on average interest rates established by a panel of around 50 European banks (panel banks) that borrow from and lend to each other. Loan maturities (the maturity of a product is the time it exists) vary from a week to a year and their rates are considered among the most important in the European money market.

• LIBOR: the London Interbank Offered Rate: the unsecured short-term borrowing rate between banks, similar to EURIBOR. Sometimes LIBOR is used to refer to rates like EURIBOR and LIBOR in general. Later on, we will discuss deals where for an agreed upon period of time, two parties trade a fixed rate for a floating (uncertain) rate. This floating rate is usually LIBOR or EURIBOR.

As the maturity of a financial instrument (the time during which the instrument exists) changes, interest rates change as well. The corresponding relation between interest rate and maturity is called the term structure or yield curve. We will use the term term structure.

Interest Rate Swaps

A swap is the exchange of one security (stock, bond, option) for another to change the quality of the maturity of issues (a series of offered stocks or bonds) or because investment objects have changed.

An interest rate swap is a derivative where interest rates are being swapped. Given the notional principle (this is a symbolic amount of money), two parties agree to pay each other interest over that amount of money on predetermined dates. One of the parties pays a fixed rate; the other pays a floating rate, e.g. LIBOR or EURIBOR, which is yet uncertain.

Example: A and B agree that 7 days from now, A pays B 1% over the agreed notional principle of $100, and B pays A a floating rate that is now 0.8%, but which A hopes will have risen above 1% in 1 week.

Notice that anyone who borrows from a bank with a fixed interest rate can transform this interest rate into a floating one if he can agree to an interest rate swap where he pays a floating rate over a notional principle that equals his loan and receives a floating rate. He can then pay the fixed rate he received from the interest rate swap to the bank and pay the floating rate in stead. This principle works in two directions.

Interest Rate Swap Options (Swaptions)

An interest rate swap option, or simply swaption, gives the owner the right but not the obligation to enter into a specific interest rate swap at a predetermined moment in time.

We can valuate swaptions using (an extension of) the Black-Scholes-Merton formula, which is in the next section.

We will mostly focus our attention on swaptions in this thesis, and they are more extensively treated in Section 2.5.
Caps and Floors

A cap is an interest rate option that pays off either 0 or a surplus of interest on predetermined dates. For example, a holder of a cap with rate 0.05 and notional $100 agrees to receive, on the payment date, nothing whenever LIBOR is at most 0.05, and receive the extra interest over $100 if the rate exceeds 0.05. E.g. if the rate is 0.07 on a payment date, he receives $(0.07-0.05)\times100=2$, and if the rate is 0.04, he receives $0$. A cap caps the rate, in the sense that if someone borrows money with floating interest, he can set a maximum interest by buying a cap. If the floating rate becomes too high, he will get back the unexpected money he has to pay on his loan by exercising his cap.

A floor has the same structure, only in this case it is vice versa: a payment is made if the floating rate is at most the floor rate, and no payment is made when the floating rate exceeds the floor rate. A floor sets a floor on the rate, in the sense that if someone lends money and receives the floating rate over it, he can make sure he receives at least the floor rate by buying a floor with exactly that floor rate. If the floating rate becomes too low, he will still get the money. Example: if you have a floor of 0.05 over a notional of $100 and the rate drops to 0.04 on the payment date, you receive $(0.05-0.04)\times100=1$, and if the rate has risen to 0.07 on the payment date, you receive nothing.

2.2 The Black-Scholes-Merton Model

Several methods we will study are based on the Black-Scholes-Merton Model, which can be used to value options and other derivatives - we discuss the model here. The material in this section is a concise adaptation of chapters 13 through 15 of [12].

2.2.1 Itô’s Lemma

The Black-Scholes-Merton formula is sometimes denoted by simply the Black-Scholes formula, since Black and Scholes derived it at the same time as Merton, only with a slightly different method. Since we studied [12], which uses Merton’s method, we will call it the Black-Scholes-Merton Model.

We first need the concept of a Wiener Process, from which Itô’s Lemma will follow, which will be needed to arrive at the Black-Scholes-Merton formula.

The reader is assumed to be familiar with the Markov Property. Behavior of stock prices in practice implies they have the Markov property. Hull explains the cause of that as follows:

“It is competition in the marketplace that tends to ensure that weak-form market efficiency and the Markov Property hold. (...) Suppose that it was discovered that a particular pattern in stock prices always gave a 65% chance of subsequent steep price rises. Investors would attempt to buy a stock as soon as the pattern was observed, and demand for the stock would immediately rise. This would lead to an immediate rise in its price and the observed effect would be eliminated, as would any profitable trading opportunities.”
A Wiener Process \( x \) is a process with the Markov property that can be described as

\[ dx = adt + bdz, \]  

where \( z \) is standard Brownian Motion. It has drift rate \( a \) and variance rate \( b^2 \).

An Itô process is such a process, but with \( a \) and \( b \) depending on \( x \) and \( t \):

\[ dx = a(x,t)dt + b(x,t)dz. \]  

Notice that an Itô Process is a Markov Process.

Itô’s Lemma states that a function \( G \) of \( x \) and \( t \) follows the following process:

\[ dG = \left( \frac{\partial G}{\partial x}a + \frac{\partial G}{\partial t} + \frac{1}{2} \frac{\partial^2 G}{\partial x^2} b^2 \right) dt + \frac{\partial G}{\partial x} b dz. \]  

Combining this with \( dS = \mu S dt + \sigma S dz \), which is the most widely used model for stock price behavior, yields

\[ dG = \left( \frac{\partial G}{\partial S} \mu S + \frac{\partial G}{\partial t} + \frac{1}{2} \frac{\partial^2 G}{\partial S^2} \sigma^2 S^2 \right) dt + \frac{\partial G}{\partial S} \sigma S dz. \]  

This is used to derive the Black-Scholes-Merton formula.

A useful interim observation is the lognormal property. This property describes the process followed by \( \ln S \) when \( S \) satisfies \( dS = \mu S dt + \sigma S dz \). If \( G := \ln S \), then we see

\[ \frac{\partial G}{\partial S} = \frac{1}{S}; \]
\[ \frac{\partial^2 G}{\partial S^2} = \frac{1}{S^2}; \]
\[ \frac{\partial G}{\partial t} = 0, \]  

and we can apply (2.6) to get

\[ dG = \left( \mu - \frac{\sigma^2}{2} \right) dt + \sigma dz, \]  

which implies that \( G \) follows a generalized Wiener Process with drift rate \( \mu - \frac{\sigma^2}{2} \) and variance rate \( \sigma^2 \). Hence

\[ \ln(S_T) - \ln(S_0) \sim N\left( \left[ \mu - \frac{\sigma^2}{2} \right] T, \sigma^2 T \right) \Rightarrow \]
\[ \ln(S_T) \sim N(\ln S_0 + \left[ \mu - \frac{\sigma^2}{2} \right] T, \sigma^2 T). \]  

Another concept we study before we finally arrive at the Black-Scholes-Merton Model is the concept of risk-neutral valuation, since this is the discrete analogue of the Black-Scholes-Merton Model, so it will make understanding the latter easier.
2.2.2 Risk-Neutral Valuation

Consider a simplification of the process that the price of a certain stock goes through, where we assume a deterministic interest rate $r$. After a set period of time, the stock price either moves up a known amount or down another known amount, with probabilities $p_{up}$ and $p_{down}$, respectively. For example (taken from [12]), $S_0 = $20, and after 3 months,

$$S_1 = \begin{cases} 22 & \text{with probability } p_{up} \\ 18 & \text{with probability } p_{down}. \end{cases}$$

(2.10)

Now we can create a portfolio, consisting of:

- $\Delta$ shares on the stock;
- -1 call option, $K = $21.

After 3 months, the portfolio will be worth either $22\Delta - 1$ (if the stock’s value increases) or $18\Delta$ (if the stock’s value decreases). If we pick $\Delta = 0.25$, then after 3 months the portfolio will be worth $4.5$, regardless of what happens, making it risk-free.

Now suppose that e.g. the risk-free rate equals $r = 0.12$, and we define $f$ to be the option price. To avoid arbitrage, the value of the portfolio at time 0 must equal the discounted value of the portfolio after 3 months, $4.5e^{0.12\times\frac{3}{12}}$, so that

$$20 \cdot 0.25 - f = 4.5e^{0.12\times\frac{3}{12}},$$

hence $f \approx 0.633$.

We hope this example adequately illustrates the concept of risk-free valuation. Let us now generalize.

- Stock price: $S_0$
- Option price: $f$
- Termination date: $T$
- $u > 1$, $d < 1$
- $S_T = \begin{cases} S_Tu & \text{with probability } p_{up} \\ S_Td & \text{with probability } p_{down} \end{cases}$
- Corresponding option prices: $f_u$ and $f_d$

We compose a portfolio of $\Delta$ shares and -1 option, so that the portfolio after time $T$ will be worth either $S_0u\Delta - f_u$ or $S_0d\Delta - f_d$. These two expressions are risklessly equal if and only if

$$\Delta = \frac{f_u - f_d}{S_0u - S_0d}.$$ 

(2.11)

The initial discounted worth of the portfolio is $(S_0u\Delta - f_u)e^{-rT}$, and setting up the portfolio costs $S_0\Delta - f$. Again to avoid arbitrage, these must be equal, which results in

$$f = S_0\Delta(1 - ue^{-rT}) + f_u e^{-rT}.$$ 

(2.12)
Substituting $\Delta = \frac{f_u - f_d}{S_0u - S_0d}$ yields

$$f = S_0 \frac{f_u - f_d}{S_0u - S_0d} (1 - u e^{-rT}) + f_u e^{-rT} = e^{-rT} (pf_u + (1 - p)f_d), \quad (2.13)$$

with $p := \frac{e^{rT}-d}{u-d}$.

This value of $p$ makes sense. If $u > e^{rT}$, then $p \in (0, 1)$ and we can interpret $p$ as $P(up)$. $f$ should equal $e^{-rT} (pf_u + (1 - p)f_d)$, and substituting $p := \frac{e^{rT}-d}{u-d}$ into $E(S_T) = pS_0u + (1 - p)S_0d$ gives us

$$pS_0(u - d) + S_0d = \left(\frac{e^{rT}-d}{u-d}\right)S_0u + S_0d = (e^{rT} - d)S_0 + S_0d = e^{rT}S_0,$$

which is exactly what we would expect.

It might seem counter-intuitive that $p_{up}$ and $p_{down}$ are missing, but they are in fact present in disguise, since we are expressing the option price in terms of the price of the underlying stock, which itself depends on $p_{up}$ and $p_{down}$. To make sure that the expected growth of the stock price matches with the risk-free interest rate, $p_{up}$ and $p_{down}$ determine the ratio between $S_0u - S_0$ and $S_0 - S_0d$, i.e. between the rise in the stock price and the decline in the stock price, a ratio which itself affects the option price.

### 2.2.3 Assumptions

The Black-Scholes-Merton Model follows the same approach, but in continuous time. It must work for any derivative dependent on a non-dividend-paying stock, and the return of a risk-free portfolio must be the risk-free interest rate $r$.

Assumptions:

- $S$ follows the process as we have been describing it thus far.
- All securities are infinitely divisible.
- No dividends are paid during the life of the derivative.
- No riskless arbitrage opportunities exist.
- $r$ is constant and independent of the maturity.
- Security trading is continuous.
- Short selling (selling a security you do not own) is allowed.
- There are no transaction costs.

### 2.2.4 Derivation of the Black-Scholes-Merton Model

We can now arrive at the Black-Scholes-Merton formula. At time $t$, the maturity is $T - t$. Recall that $dS = \mu Sdt + \sigma Sdz$. From Itô's lemma, it then follows that

$$df = \left(\frac{\partial f}{\partial S}\mu S + \frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} \sigma^2 S^2\right) dt + \frac{\partial f}{\partial S} \sigma Sdz. \quad (2.14)$$
Now we set up a portfolio with $\frac{\partial f}{\partial S}$ shares and $-1$ derivative, which has value
\[ \Pi = -f + \frac{\partial f}{\partial S} S. \] (2.15)
To avoid arbitrage, we need
\[ d\Pi = r\Pi dt. \] (2.16)
Combining the three previous equations yields
\[ \frac{\partial f}{\partial t} + rS \frac{\partial f}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} = rf. \] (2.17)
The differential equation does not have boundary conditions yet. Since for a European call option, we always have, for $t = T$,
\[ f = \max(S - K, 0), \] (2.18)
and for a European put option
\[ f = \max(K - S, 0), \] (2.19)
we can set these as boundary conditions. The solutions for the price $C$ of a European call and the price $P$ of a European put, which are known as the Black-Scholes-Merton formulas, then become
\[ C = S_0 \Phi(d_1) - Ke^{-rT} \Phi(d_2); \] (2.20)
\[ P = Ke^{-rT} \Phi(-d_2) - S_0 \Phi(-d_1). \] (2.21)
Here,
\[ d_1 = \frac{\ln(S_0/K) + (r + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}, \quad d_2 = \frac{\ln(S_0/K) + (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T}. \] (2.22)
We used $\Phi$ to denote the distribution function of a standard normally distributed random variable.

### 2.3 The SABR Model

#### 2.3.1 Introduction

The unknown variable in the Black-Scholes-Merton formula is the volatility. It has to be estimated, based on historical data.

The model can also be used to valuate other derivatives. If the volatility is unknown but all the other variables are known, and the option prices can be observed, then we have an equation with only the volatility as unknown variable, so we can solve it for $\sigma$ and thus obtain what is referred to as the implied volatility. This implied volatility can then be implemented in other formulas to estimate e.g. the
The value of a swaption, assuming the implied volatility is a good estimation for the actual volatility.

The observed implied volatility is not constant for different strikes. It behaves like what is known in the literature as the \textit{volatility smile}, which means that the implied volatility is lowest for strikes around the current stock price, i.e. \textit{At The Money} (ATM) strikes, and higher for low or high strikes. When stock prices increase, the smile should move to the right (i.e. to higher strikes), and when the stock price decreases, the smile should move to the left. However, in the estimation methods that were used e.g. in [9], the implied volatility moves in the opposite direction. This makes using the Black-Scholes-Merton Model to valuate derivatives in the described manner problematic. The technical details of this problem are presented in Section 2.3.3.

An alternative way to do so was described in [21]. It is called the SABR method (later fine-tuned in [20]). We present [21] in this section.

\subsection*{2.3.2 Notation}

- $A$ is the asset.
- $t_{ex}$ denotes an exercise date, e.g. of a European call option or a European swaption.
- $t_{set}$ denotes a settlement date, e.g. of a European call option or a European swaption.
- $K$ is the strike price of an option.
- $\hat{F}(t)$ is the forward price of the asset for a forward contract that matures on $t_{set}$.
- $f := \hat{F}(0)$.
- $D(t)$ is the discount factor for date $t$ (the value today of $1$ to be delivered on date $t$).
- $V_{\text{call}}$ is the value of a European call option.
- $V_{\text{put}}$ is the value of a European put option.
- $R_{fix}$ is the fixed rate (the strike) of a swaption.
- $\hat{R}_s(t)$ is the forward swap rate.
- $R_0 := \hat{R}_s(0)$ is the forward swap rate today.
- $V_{\text{pay}}$ is the value of a payer swaption.
- $V_{\text{rec}}$ is the value of a receiver swaption.
2.3.3 Motivation for the Need of SABR

By martingale pricing theory, $\hat{F}(t)$ is a martingale, so the Martingale Representation Theorem implies that

$$d\hat{F} = C(t, \ast) dW; \hat{F}(0) = f$$  \hspace{1cm} (2.23)

for some coefficient $C(t, \ast)$, where $dW$ is Brownian Motion.

The same holds for European swaptions:

$$d\hat{R}_s = C(t, \ast) dW, \hat{R}_s(0) = R_0.$$  \hspace{1cm} (2.24)

Black’s specification of $C(t, \ast)$ in [5] was

$$d\hat{F} = \sigma_B \hat{F}(t) dW, \hat{F}(0) = f$$  \hspace{1cm} (2.25)

(the subscript $B$ is just notation so we know it comes from Black), which resulted, using the notation of [21], in:

$$V_{\text{call}} = D(t_\text{set})(f \Phi(d_1) - K \Phi(d_2));$$  \hspace{1cm} (2.26)

$$V_{\text{put}} = V_{\text{call}} + D(t_\text{set})(K - f),$$  \hspace{1cm} (2.27)

where

$$d_{1,2} = \frac{\log \frac{K}{f} \pm \frac{1}{2} \sigma^2 t_{\text{ex}}}{\sigma_B \sqrt{t_{\text{ex}}}}.$$  \hspace{1cm} (2.28)

The problem with this assumption for $C(t, \ast)$ is, as mentioned in the introduction, lies in the volatility smiles. The valuation of down-and-out knock-out barrier options, hedging and the evolution of the implied volatility curve fail when making this assumption.

In [8], [9] and [10], a solution is proposed by assuming that $C(t, \ast)$ is Markovian, creating the local volatility model. Unfortunately, this model led to the wrong dynamics: in typical market behavior, smiles and skews move in the same direction as the underlying, but in this model, they move in the opposite direction: if the forward price $f$ increases, the implied volatility curve moves to the right, and if $f$ decreases, the implied volatility curve moves to the left.

The SABR method provides a celebrated and widely used alternative.

2.3.4 The Method

We cannot use a Markovian model based on a single Brownian Motion, and we do not want the model to be non-Markovian or to be based on non-BM, so instead we choose a two-factor model, with two, possibly correlated, Brownian Motions:

$$C(t, \ast) = \hat{\alpha} \hat{F}^\beta.$$  \hspace{1cm} (2.29)

For the volatility $\hat{\alpha}$ the simplest stochastic process is chosen:
\[
d\hat{F} = \hat{\alpha}\hat{F}\beta dW_1, \hat{F}(0) = f; \\
d\hat{\alpha} = \nu\hat{\alpha}dW_2, \hat{\alpha}(0) = \alpha.
\]

(2.30)

Here, the Brownian Motions \(W_1\) and \(W_2\) are correlated by:

\[
dW_1dW_2 = \rho dt.
\]

(2.31)

SABR is the simplest model which is homogeneous in \(\hat{F}\) and \(\hat{\alpha}\). It provides an explicit expression for \(\sigma_B(K, f)\), which can be implemented in Black’s formula. The formula for the volatility becomes:

\[
\sigma_B(K, f) = \alpha \frac{1}{(fK)^{1-\beta}} \left\{ \frac{1}{2} \left[ 1 + \left(1 - \beta\right)^2 \log^2 \frac{f}{K} + \left(1 - \beta\right) \log \frac{f}{K} + \ldots \right] \cdot \frac{z}{x(z)} \right\} \\
\cdot \left\{ 1 + \frac{(1 - \beta)^2}{24} \frac{\alpha^2}{(fK)^{1-\beta}} + \frac{1}{4} \frac{\rho \beta \nu \alpha}{(fK)^{1-\beta}} + \frac{2 - 3 \rho^2}{24} \nu^2 \right\} \text{ex} + \ldots
\]

(2.32)

where

\[
z = \frac{\nu}{\alpha} (fK)^{1-\beta} \log \frac{f}{K}
\]

(2.33)

and

\[
x(z) = \log \left\{ \frac{\sqrt{1 - 2\rho z + z^2} + z - \rho}{1 - \rho} \right\}.
\]

(2.34)

In the next section, we present yet another alternative to the SABR model, which will become very relevant in view of the recent developments that created our thesis subject.

### 2.4 The Hull-White (Extended Vasicek) Model

#### 2.4.1 Introduction

An alternative to the SABR model is the model of Hull and White. One of the classical drawbacks of this model is the possibility of negative rates. However, since rates are actually becoming negative, this model could very well be the basis of solving the problem those negative rates give rise to. Before we can study their recent model developments, we first need to understand their original model in [13], which can also be used in a tree building procedure: a useful numerical method. In this chapter, we present both the model and the tree building procedure.

#### 2.4.2 The Hull-White Model

Hull and White present an extension to the Vasicek Model in [24]. First, they criticize other available methods, noting that several of their parameters are unobservable. Furthermore, they explain the distinction between equilibrium models and
no-arbitrage models. In equilibrium models, assumptions are made about economic variables, and those assumptions are then used to derive a process for the short rate \( r \). These models generally do not exactly fit the initial term structure of interest rates.

**No-arbitrage** models do not have this problem, since the current term structure of interest rates is an input. To build an adequate no-arbitrage model, Hull and White first recap the Vasicek Model for interest rates,

\[
\begin{align*}
   dr &= a(b - r)dt + \sigma dz, \\
   \text{(2.35)}
\end{align*}
\]

where \( a, b \) and \( \sigma \) are positive constants and \( dz \) is a Wiener Process. Vasicek derived an analytic solution for the price of a discount bond.

“One drawback (...) is that the short-term interest rate, \( r \), can become negative.” An example of this happening (where the Vasicek Model is discretized) can be found in [11] - our Figure 8.1. As mentioned in the introduction, this drawback of (extensions of) Vasicek-models seems to have become an advantage.

To reach consistency with the term structure and volatilities of interest rates, Hull and White add a time-dependent drift \( \theta(t) \) and drop the assumption that \( a \) and \( \sigma \) are constant: they now depend on time. The results were summarized in [12]. The model becomes:

\[
\begin{align*}
   dr &= [\theta(t) - ar]dt + \sigma dz. \\
   \text{(2.36)}
\end{align*}
\]

Furthermore,

\[
\begin{align*}
   \theta(t) &= F_t(0,t) + aF(0,t) + \frac{\sigma^2}{2a}(1 - e^{-2at}). \\
   \text{(2.37)}
\end{align*}
\]

Then the assumption is made that the market price of interest-rate risk is a function of time that is bounded in any finite interval. This allows the derivation of bond prices:

\[
\begin{align*}
   P(t,T) &= A(t,T)e^{-B(t,T)r}, \\
   \text{with} \\
   B(t,T) &= \frac{1 - e^{-a(T-t)}}{a} \\
   \text{(2.38)}
\end{align*}
\]

\[
\ln A(t,T) = \frac{P(0,T)}{P(0,t)} + B(t,T)F(0,t) - \frac{1}{4a^2}\sigma^2(e^{-aT} - e^{-at})^2(e^{2at} - 1).
\]

### 2.4.3 Using the Hull-White Model on Swaptions

In short, the Hull-White Model can be used to valuate European swaptions by viewing a swaption as an option on a coupon-bearing bond, which can be expressed as the sum of zero-coupon bonds, which can be valuated with the Hull-White Model.

Using the Hull-White Model, Hull gives the price of an option that matures on time \( T \) on a zero-coupon bond that matures on time \( s \), with strike \( K \) and principal
\( L: \)

Call: \( LP(0, s)N(h) - KP(0, T)N(h - \sigma_p); \)

Put: \( KP(0, T)N(\sigma_p - h) - LP(0, s)N(-h), \)

where \( h = \frac{1}{\sigma_p} \ln \frac{LP(0, s)}{P(0, T)K} + \frac{\sigma_p}{2} \)

and \( \sigma_p = \frac{\sigma}{a} [1 - e^{-a(s-T)}] \sqrt{\frac{1}{2} - e^{-2aT}}. \)

Because the price of a zero-coupon bond depends on \( r \), we can express a European option with strike \( K \) and maturity \( T \) on a coupon-bearing bond with price \( B(r) \) as the sum of European options on zero-coupon bonds:

1. Calculate the value \( r^* \) for which \( B(r^*) = K. \)

2. Define zero-coupon bonds that together comprise the coupon-bearing bond: each zero-coupon bond represents an interest payment of the coupon-bearing bond.

3. Calculate the prices of options with maturity \( T \) on the zero-coupon bonds described in the previous point.

4. Take the sum over the prices of the options on the zero-coupon bonds. This sum is the price of the option on the coupon-bearing bond.

We can also represent this model in the form of a trinomial tree, so let us first discuss binomial trees.

### 2.4.4 Binomial Trees

We return to Chapters 13 and 21 of [12] to describe the Binomial Tree procedure, which was first developed in [17].

Recall the concept of risk-neutral valuation. We treated a one-step binomial tree, but this is of course a simplification of reality. If analytic solutions are not available, e.g. when analytic solutions are only available when all rates are positive, then we might extend the tree to create an approximation of reality (see Figure 2.1). At each time step, the stock price increases with a factor \( u > 1 \) with probability \( p \) and with a factor \( d < 1 \) (implying a decrease) with probability \( 1 - p \).

#### Determination of \( p, u \) and \( d \)

An asset with yield \( q \) and price \( S \) at \( t = 0 \) must have an expected price (after \( \Delta t \)) of \( Se^{(r-q)\Delta t} \), so it is required that \( Se^{(r-q)\Delta t} = pSu + (1 - p)Sd \), which, dividing by \( S \), is equivalent to

\[
e^{(r-q)\Delta t} = pu + (1 - p)d. \tag{2.40}
\]
If the percentage change of the asset price in $\Delta t$ has variance $\sigma^2 \Delta t$, then this implies

$$pu^2 + (1 - p)d^2 - e^{2(r-q)\Delta t} = \sigma^2 \Delta t,$$

and combining this with (2.40) yields

$$e^{(r-q)\Delta t}(u + d) - ud - e^{2(r-q)\Delta t} = \sigma \Delta t.$$  \hfill (2.41)

To make recombination of the tree possible, we add a third condition (proposed in [17]):

$$u = \frac{1}{d}. \hfill (2.42)$$

Ignoring terms of order higher than $\Delta t$, these three conditions give solutions:

$$p = \frac{e^{(r-q)\Delta t} - d}{u - d};$$
$$u = e^{\sigma \sqrt{\Delta t}};$$
$$d = e^{-\sigma \sqrt{\Delta t}}.$$  \hfill (2.43)

Now suppose there are $N$ intervals of length $\Delta t$, and the jth node (counting from bottom to top) at $i\Delta t$ is node $(i,j)$, then the asset price is $S_0 u^i d^{N-j}$, so, using risk-neutral valuation, we have at the final nodes:

Call option: $f_{N,j} = \max\{S_0 u^j d^{N-j} - K, 0\}, j \in \{0, 1, \ldots, N\};$

Put option: $f_{N,j} = \max\{K - S_0 u^j d^{N-j}, 0\}, j \in \{0, 1, \ldots, N\},$ \hfill (2.44)
and for $i \in \{0, 1, ..., N-1\}$:

- **European call or put option:**
  \[
  f_{i,j} = e^{-r\Delta t}(pf_{i+1,j+1} + (1 - p)f_{i+1,j});
  \]
  American call option:
  \[
  f_{i,j} = \max\{S_0u^jd^{j-i} - K, e^{-r\Delta t}(pf_{i+1,j+1} + (1 - p)f_{i+1,j})\}
  \]
  American put option:
  \[
  f_{i,j} = \max\{K - S_0u^jd^{j-i}, e^{-r\Delta t}(pf_{i+1,j+1} + (1 - p)f_{i+1,j})\}.
  \]

### 2.4.5 Trinomial Trees

Now that the idea is clear, let us move on to trinomial trees. This proves to provide a more complicated model but also a model that is equivalent to finite difference methods. In [12], Hull summarizes [14] and [15] to explain trinomial trees, and we follow his example.

Hull and White build the tree in two stages. Recall the Hull-White Model:

\[
\begin{align*}
\frac{dr}{\Delta t} &= [\theta(t) - ar] dt + \sigma dz, \\
\frac{dR}{\Delta t} &= [\theta(t) - aR] dt + \sigma dz.
\end{align*}
\]

The time steps on the tree are assumed to be constantly $\Delta t$.

Another assumption is that the $\Delta t$ rate, $R$, follows the same process:

\[
\frac{dR}{\Delta t} = [\theta(t) - aR] dt + \sigma dz.
\]

#### First Stage

The first step is to build a trinomial tree for almost the same process, but with $\theta(t)$ left out:

\[
\frac{dR^*}{\Delta t} = -aR^* dt + \sigma dz.
\]

Ignoring terms of higher order than $\Delta t$, we have $E(R^*(t + \Delta t) - R^*(t)) = -aR^*(t)\Delta t$, and $\text{Var}(R^*(t + \Delta t) - R^*(t)) = \sigma^2\Delta t$.

By an error minimization argument, the rates on the tree have distance $\Delta R = \sigma\sqrt{3\Delta t}$.

At each node, $R$ can move up $\Delta R$, or move down $\Delta R$, or remain unchanged, except at the top and bottom node of the final time step, in which case the top node can move down $\Delta R$, move down $2\Delta R$ or remain unchanged, and the bottom node can move up $\Delta R$, move up $2\Delta R$ or remain unchanged. Figure 2.2 explains this visually. Hull and White show that if $j$ moves at most $0.184\Delta t$ up and down, probabilities remain positive. This also means that negative rates are a possibility in this model: if the tree includes negative rates, the positiveness of the probabilities of moving downwards guarantees this possibility.

The mean and variance we mentioned, plus the fact that probabilities need to add up to 1, give three equations with three unknowns ($p_u$, $p_d$ and $p_m$, which are the upper, lower and middle probabilities of moving, respectively).
Figure 2.2: A Trinomial Tree
For all the nodes but the top right and bottom right ones, we get

\[ \begin{align*}
    p_u \Delta R - p_d \Delta R &= -aj \Delta R \Delta t; \\
    p_u \Delta R^2 + p_d \Delta R^2 &= \sigma^2 \Delta t + a^2 j^2 \Delta R^2 \Delta t^2; \\
    p_u + p_m + p_d &= 1, 
\end{align*} \]

with solutions

\[ \begin{align*}
    p_u &= \frac{1}{6} + \frac{1}{2} (a^2 j^2 \Delta t^2 - aj \Delta t); \\
    p_m &= \frac{2}{3} - a^2 j^2 \Delta t^2; \\
    p_d &= \frac{1}{6} + \frac{1}{2} (a^2 j^2 \Delta t^2 + aj \Delta t). 
\end{align*} \]

Similarly, for the bottom right node, we get

\[ \begin{align*}
    p_u &= \frac{1}{6} + \frac{1}{2} (a^2 j^2 \Delta t^2 + aj \Delta t); \\
    p_m &= -\frac{1}{3} - a^2 j^2 \Delta t - 2aj \Delta t; \\
    p_d &= \frac{7}{6} + \frac{1}{2} (a^2 j^2 \Delta t^2 + 3aj \Delta t), 
\end{align*} \]

and for the top right node:

\[ \begin{align*}
    p_u &= \frac{7}{6} + \frac{1}{2} (a^2 j^2 \Delta t^2 - 3aj \Delta t); \\
    p_m &= -\frac{1}{3} - a^2 j^2 \Delta t^2 + 2aj \Delta t; \\
    p_d &= \frac{1}{6} + \frac{1}{2} (a^2 j^2 \Delta t - aj \Delta t). 
\end{align*} \]

**Second Stage**

The second stage consists of converting the tree for \( R^* \) into the tree for \( R \), choosing \( \theta \) in such a way that the initial term structure is matched exactly. This is done by defining \( \alpha(t) = R(t) - R^*(t) \), \( \alpha_i = \alpha(i \Delta t) \) and \( Q_{i,j} \) as the Arrow Debreu price of node \((i,j)\). The Arrow Debreu price is the present value of a security that pays $1 precisely when node \((i,j)\) is reached and $0 otherwise, where node \((i,j)\) stands for the node with \(i\) time steps and \(j\) rate steps, where \(j\) might be negative. Then, match the initial term structure in a forward induction manner, calculating all the \( \alpha_i \)s.

For a better understanding of this stage, Hull included an instructive section “Illustration of Second Stage,” which is quite informative. Then, he gives general formulas for the \( \alpha_i \)s and \( Q_{i,j} \)s: if the \( Q_{i,j} \)s have been determined up to \( i = m \), \( P_{m+1} \)
is the price of a zero-coupon bond maturing at \((m + 1)\Delta t\) and \(n_m\) is the number of nodes at each side of the central node at \(m\Delta t\), then

\[
\alpha_m = \ln \frac{\sum_{j=-n_m}^{j=n_m} Q_{m,j} e^{-j\Delta R\Delta t - \ln P_{m+1}}}{\Delta t},
\]

after which the \(Q_{i,j}\)'s that follow up can be calculated by summing over all the values of \(k\) for which the probability \(q(k, j)\) of moving from node \((m, k)\) to node \((m + 1, j)\) is positive:

\[
Q_{m+1,j} = \sum_k Q_{m,k} q(k, j) e^{-(\alpha_m + k\Delta R)\Delta t}.
\]

### 2.4.6 Conclusion

For reasons described in this chapter, the Hull-White method was already arguably preferable to other methods. Hull states in addition that their method has the advantage of being analytically tractable, and the disadvantage of allowing negative rates, while lognormal models have the advantage of not allowing negative rates but the disadvantage of not being analytically tractable. The analytic tractability of the Hull-White Model remains an advantage, but the allowance of negative rates, in perspective of the recent developments, actually becomes an advantage. This is the main reason why we will recommend the Hull-White Model, of which we present the recent improvements in Chapter 5.

### 2.5 From Options to Swaptions in General

#### 2.5.1 Introduction

The mathematical models we described so far seem to focus on options rather than swaptions, but as soon as we can value options, we can extend the methodology to swaptions. We do so in this chapter, which is based on chapter 1 of [6]. We first build up the necessary mathematical framework to be able to deal with swaptions in addition to options, and then extend the available option theory to swaptions, as announced in Section 2.1. This will result in a more detailed explanation of the problems that arise when negative interest rates become a possibility.

#### 2.5.2 Definitions

**Stochastic Discount Factor**

Recall the definition of a bank account \(B(t)\) in (2.1). So far, we assumed that the risk-free interest rate was deterministic. In that case, at time \(t\) one unit of currency available at time \(T\) is worth \(\frac{B(t)}{B(T)}\). Generalizing, since the rate is clearly important in interest rate derivatives, we define the **stochastic discount factor** by

\[
D(t, T) = \frac{B(t)}{B(T)} = e^{-\int_t^T r_s ds}.
\]
Zero-coupon Bonds

A $T$-maturity zero-coupon bond or pure discount bond is a contract that guarantees its holder the payment of one unit of currency at time $T$, with no intermediate payments. The contract value at time $t < T$ is denoted by $P(t,T)$. Notice that $P(T,T)=1$.

It is worth noting that for deterministic $r$, $D(t,T) = P(t,T)$ for any $t$ and $T$. Later in [6], it is explained that under a certain probability measure, $P(t,T)$ can be seen as the expectation of $D(t,T)$.

Various, Differently Compounded Spot Interest Rates

Before we can move on to interest rates, we need to define and distinguish between spot interest rates that are compounded in various ways. In these definitions, $\tau(t,T)$ is the chosen time measure between the dates $t$ and $T$, usually referred to as year fraction between the dates $t$ and $T$.

The continuously-compounded spot interest rate $R(t,T)$, prevailing at time $t$ for the maturity $T$, is the constant rate at which an investment of $P(t,T)$ units of currency at time $t$ accrues continuously to yield a unit amount of currency at maturity $T$. Expressed in zero-coupon bonds:

$$R(t,T) := -\frac{\ln P(t,T)}{\tau(t,T)}.$$  \hfill (2.56)

The converse expression is

$$P(t,T) = e^{-R(t,T)\tau(t,T)}.$$  \hfill (2.57)

When accruing occurs proportionally to the investment time, the above definition becomes the definition of the simply-compounded spot interest rate:

$$L(t,T) := \frac{1 - P(t,T)}{\tau(t,T)P(t,T)},$$  \hfill (2.58)

with converse expression

$$P(t,T) = \frac{1}{1 + L(t,T)\tau(t,T)}.$$  \hfill (2.59)

If we adjust the definition to reinvesting the obtained amounts once a year, we get the annually-compounded spot interest rate:

$$Y(t,T) := \frac{1}{P(t,T)^{\frac{1}{\tau(t,T)}}} - 1,$$  \hfill (2.60)

so that

$$P(t,T) = \frac{1}{(1 + Y(t,T))^{\tau(t,T)}}.$$  \hfill (2.61)
Finally, if the reinvestment is made $k$ times a year, we get the $k$-times-per-year compounded spot interest rate:

$$Y^k(t, T) := \frac{k}{P(t, T)^{\frac{1}{k\tau(t, T)}}} - k,$$

(2.62)

hence

$$P(t, T) = \frac{1}{\left(1 + \frac{Y^k(t, T)}{k}\right)^{k\tau(t, T)}}.$$  
(2.63)

All these rates go to $r$ as $T \to t^+$.

**Forward Rates**

Recall the definition of interest rate swaps in Section (2.1). Before we give a formal definition, we will define forward rates. To define these, we need the concept of a forward-rate agreement (FRA), which is a particular interest rate swap. It is an agreement on time $t$ to exchange on time $S$ a fixed rate $K$ against the simply-compounded spot interest rate $L(T, S)$ over an agreed nominal value (notional principal) $N$, where $t < T < S$. The value of such a contract is clearly $N\tau(T, S)(K - L(T, S))$, and the value at time $t$ is, less clearly but explained in [6],

$$\text{FRA}(t, T, S, \tau(T, S), N, K) = N(P(t, S)\tau(T, S)K - P(t, T) + P(t, S)).$$

(2.64)

The unique $K$ for which the contract is fair and FRA is thus 0 is the simply-compounded forward interest rate $F(t; T, S)$. To stay consistent and express it in terms of zero-coupon bonds:

$$F(t; T, S) := \frac{1}{\tau(T, S)} \left(\frac{P(t, T)}{P(t, S)} - 1\right).$$

(2.65)

We can use this definition to rewrite (2.64) as:

$$\text{FRA}(t, T, S, \tau(T, S), N, K) = N(P(t, S)\tau(T, S)(K - F(t; T, S))).$$

(2.66)

This confirms that $F(t; T, S)$ is the rate which makes the FRA fair.

We would of course like $F(t; T, S)$ to be the expectation of $L(T, S)$. This turns out to be the case: again, the details are in [6].

**2.5.3 Interest Rate Swaps**

Generalizing the FRAs brings us to interest rate swaps (IRSs) as already preliminary defined in Section 2.1. To be more precise, the holder of a Payer (Forward-Start) Interest Rate Swap (PFS) pays at dates $T_{a+1}, ..., T_{\beta}$ a fixed $N\tau_{i}K$ with $\tau_{i}$ the year fraction between $T_{i-1}$ and $T_{i}$, and receives $N\tau_{i}L(T_{i-1}, T_{i})$, and for the holder of a Receiver Interest Rate Swap (RFS) it is the other way around.
To determine the value of an interest rate swap, first notice that the discounted payoff at a time \( t < T_\alpha \) of a PFS is

\[
\sum_{i=\alpha+1}^{\beta} D(t, T_i) N \tau_i (L(T_{i-1}, T_i) - K),
\]

and the discounted payoff of an RFS is

\[
\sum_{i=\alpha+1}^{\beta} D(t, T_i) N \tau_i (K - L(T_{i-1}, T_i)).
\]

Now set \( \mathcal{T} := \{T_0, ..., T_\beta\} \) and \( \tau := \{\tau_{\alpha+1}, ..., \tau_\beta\} \).

We can view (2.68) as a portfolio of FRAs, for which we already have formulas, so we can add them up:

\[
\text{RFS}(t, \mathcal{T}, \tau, N, K) = \\
\sum_{i=\alpha+1}^{\beta} \text{FRA}(t, T_{i-1}, T_i, N, K) = \\
N \sum_{i=\alpha+1}^{\beta} \tau_i P(t, T_i)(K - F(t; T_{i-1}, T_i)) = \\
-NP(t, T_\alpha) + NP(t, T_\beta) + N \sum_{i=\alpha+1}^{\beta} \tau_i KP(t, T_i).
\]

The final equality holds by a telescoping argument and remembering that \( \tau_i \) is the year fraction between \( T_{i-1} \) and \( T_i \).

We now define two prototypical contracts which represent the two legs of an IRS: the IRS can be viewed as an exchange of these two contracts.

Firstly, a **prototypical coupon-bearing bond** is a contract that ensures the payment at future times \( T_{\alpha+1}, ..., T_\beta \) of the deterministic amounts of currency (cash flows) \( c := \{c_{\alpha+1}, ..., c_\beta\} \) which is typically defined as \( N \tau_i K \) for \( i < \beta \) and \( c_\beta = N \tau_\beta K + N \), where \( K \) is a fixed interest rate and \( N \) the bond nominal value, which is reimbursed in the last cash-flow. Taking discounting for each cash flow into account, notice that the value of a prototypical coupon-bearing bond is

\[
\text{CB}(t, \mathcal{T}, c) = \sum_{i=\alpha+1}^{\beta} c_i P(t, T_i).
\]

Secondly, a **prototypical floating-rate note** is a contract that ensures the payment at future times \( T_{\alpha+1}, ..., T_\beta \) of floating rates that reset at the previous instants \( T_\alpha, ..., T_{\beta-1} \) and the payment of the notional value \( N \) at \( T_\beta \). The value of holding a prototypical floating-rate note is the same as being short an RFS, except we add the present value of \( N \) paid at time \( T_\beta \), which is \( NP(t, T_\beta) \):

\[
\text{FR}(t, \mathcal{T}, \tau, N) = -\text{RFS}(t, \mathcal{T}, \tau, N, 0) + NP(t, T_\beta),
\]
which equals $NP(t, T_\alpha)$ by (2.69), so a prototypical floating-rate note has the same worth as $N$ units of currency at its first reset date.

Finally, the **forward swap rate** is the rate for which the IRS is fair. Expressed in terms of zero-coupon bonds, this comes down to

$$S_{\alpha,\beta}(t) = \frac{P(t, T_\alpha) - P(t, T_\beta)}{\sum_{\alpha+1}^\beta \tau_i P(t, T_i)}.$$ (2.72)

And in terms of forward rates:

$$S_{\alpha,\beta}(t) = \frac{1 - \prod_{\alpha+1}^\beta \frac{1}{1 + \tau_i F_i(t)}}{\sum_{\alpha+1}^\beta \tau_i \prod_{\alpha+1}^\beta \frac{1}{1 + \tau_i F_i(t)}}.$$ (2.73)

### 2.5.4 Swaptions

Recall the definition of swaptions in Section 2.1. To adapt the notation in [6]: the **tenor** is the underlying IRS-length $T_\beta - T_\alpha$. $T_\alpha$ is assumed to be the swaption maturity. Then at maturity, the value of the underlying payer IRS, as we already saw in (2.69), is

$$N \sum_{\alpha+1}^\beta \tau_i P(T_\alpha, T_i)(F(T_\alpha; T_{i-1}, T_i) - K).$$ (2.74)

Since the swap option will only be exercised if this value is positive, the discounted payer-swaption payoff is

$$ND(t, T_\alpha) \left( \sum_{\alpha+1}^\beta \tau_i P(T_\alpha, T_i)(F(T_\alpha; T_{i-1}, T_i) - K) \right)^+. $$ (2.75)

In valuing a swaption, terminal correlation between the rates has to be taken into account. We use the term “terminal” because the rates are the terminal rates of the time periods, so we are not considering instantaneous correlations.

It is market practice to price a payer swaption as follows,

$$PS^{Black}(0, T, \tau, N, K, \sigma_{\alpha,\beta}) = N Bl(K, S_{\alpha,\beta}(0), \sigma_{\alpha,\beta} \sqrt{T_\alpha}, 1) \sum_{i=\alpha+1}^\beta \tau_i P(0, T_i),$$ (2.76)

and a receiver swaption as follows,

$$RS^{Black}(0, T, \tau, N, K, \sigma_{\alpha,\beta}) = N Bl(K, S_{\alpha,\beta}(0), \sigma_{\alpha,\beta} \sqrt{T_\alpha}, -1) \sum_{i=\alpha+1}^\beta \tau_i P(0, T_i),$$ (2.77)

where $\Phi$ is the standard Gaussian cumulative distribution function, and

$$Bl(K, S_{\alpha,\beta}(0), \sigma_{\alpha,\beta} \sqrt{T_\alpha}, \pm 1) = S_{\alpha,\beta}(0)\Phi(\pm d_1(K, S_{\alpha,\beta}(0), \sigma_{\alpha,\beta} \sqrt{T_\alpha})) \pm K\Phi(\pm d_2(K, S_{\alpha,\beta}(0), \sigma_{\alpha,\beta} \sqrt{T_\alpha})).$$ (2.78)
where

\[
\begin{align*}
    d_1(K, S_{α,β}(0), σ_{α,β}\sqrt{T_α}) &= \frac{\ln S_{α,β}(0) + σ_{α,β}^2 T_α}{σ_{α,β}\sqrt{T_α}}, \\
    d_2(K, S_{α,β}(0), σ_{α,β}\sqrt{T_α}) &= \frac{\ln S_{α,β}(0) - σ_{α,β}^2 T_α}{σ_{α,β}\sqrt{T_α}}.
\end{align*}
\]  

(2.79)

A derivation of these formulas can e.g. be found in [7], and the most commonly used method to estimate the volatilities is the previously discussed SABR method.

### 2.6 The Problem

The above described market practice used to work. The idea is that negative strikes do not make sense. A negative strike of K would simply mean that one of the two parties in a swaption agreement would pay the fixed K, i.e. receive $-K > 0$, and also receive the floating rate. This is of course assuming that the floating rate remains positive at all times. Rates approaching zero made the same practical nonsense. This is also observable in the formulas: the input of the logarithms must be positive, and the denominator of the fractions shouldn’t approach zero, since that makes the fractions explode. Since the financial crisis in 2008, however, low or even negative rates have been observed in e.g. LIBOR, which calls for other solutions. \(S_{α,β}\) and \(K\) might be negative. If they are negative simultaneously and \(K\) is sufficiently large, then this formula at least still produces an output, but clearly a questionable one, and otherwise everything explodes (for \(K ↓ 0\)) or is not even defined (for \(S_{α,β}K < 0\)). The remaining of this thesis studies the possibility to create such solutions. The underlying idea is that in traditional models, where the model in [5] was used most commonly, the terminal short rate was assumed to be lognormally distributed, and hence always positive. This assumption has to be modified in some way. We could apply some appropriate modification of SABR, or leave it altogether. A practical advantage of modifications of SABR as opposed to creating a whole new model is that corporations that already use SABR can keep on doing so, with the modification, instead of having to implement something completely new. Therefore, we will first focus on methods that use SABR in any way.
Chapter 3
The Shifted SABR Model (2014)

3.1 Introduction

Now that we have arrived at the core of the problem, it is time to compare the different possible solutions that have been posed by different teams of researchers. Naturally, we started by examining the Shifted SABR Model. We say “naturally,” since this is the method that comes closest to the existing SABR method. It would be practical for corporations to have a method available with which a small alteration to the method they are already using (the SABR method) suffices to give good results. That is the idea behind Shifted SABR: shift the rates enough so that they become positive again, and apply SABR with that shifted rate. With an unexpected problem like rates becoming negative, it makes sense to start by trying to find a solution as simple as possible, and see if the result satisfy.

The Shifted SABR Model appears in [22], where Schlenkrich et al. first give practical examples of rates dropping near or below zero, and an overview of existent methods for swaption pricing. Those methods do not cope with low or negative rates. Then they present the Shifted SABR. The method is described from a point that is as general as possible, after which the choices the authors made are filled in. This contributes to the value of the article, because we might alter the choices of the authors and still adapt their method without too much difficulty. The method is based on keeping the standard SABR formula unchanged around ATM, and changing it appropriately for lower strikes. To do this, the terminal density, which is the probability density of the terminal rate implied by the method, is composed out of two parts that are being glued together.

3.2 Gluing Terminal Densities

A concise way to denote the price of any European Swaption is:

\[
Swpt(t; K, \phi) = An_{\alpha,\beta}(t) \cdot \mathbb{E}^A[(\phi[S_{\alpha,\beta}(T) - K])^+],
\]

where \(S_{\alpha,\beta} = \mathbb{E}^A[S_{\alpha,\beta}(T)]\) is the forward swap rate for a swap running from \(T_{\alpha}\) to \(T_{\beta}\) \((t \leq T \leq T_{\alpha} \leq T_{\beta})\), \(K\) the strike rate, \(\phi \in \{-1, 1\}\) denotes whether it is a receiver or payer swaption and \(An_{\alpha,\beta}(t)\) is the underlying swap’s fixed leg annuity. If we
capture the pricing model in the density function \( p(S) \) of the terminal swap rate, then this must equal

\[
\mathbb{E}^A \left[ (\phi[S_{\alpha,\beta}(T) - K])^+ \right] = \int_{-\infty}^{\infty} (\phi[s - K])^+ p(s) ds. \tag{3.2}
\]

This \( p(s) \) can of course be obtained by differentiating the cumulative distribution function \( P(S) = \int_{-\infty}^{S} p(s) ds \), but the latter in fact equals the derivative of the receiver swaptions price with respect to \( K \): see Appendix A for a proof of this statement, which results in

\[
P(K) = \frac{\partial}{\partial K} \int_{-\infty}^{\infty} (K - s)^+ p(s) ds = \frac{\partial}{\partial K} \left[ \text{Swpt}(t; K, -1) \right], \tag{3.3}
\]

and hence

\[
p(K) = \frac{\partial P(K)}{\partial K} = \frac{\partial^2}{\partial K^2} \left[ \text{Swpt}(t; K, -1) \right]. \tag{3.4}
\]

The SABR method applies Black, but with an improved way to calculate implied volatilities. This implies a terminal density of the rate. The Shifted SABR method aims to repair the terminal rate such that negative rates are taken into account.

A cut-off \( S^* \) will be chosen: if we denote the density that is implied by the regular SABR by \( p_h(S) \) and an alternative density by \( p_l(S) \), then we can define a combined terminal distribution:

\[
p(S) = \begin{cases} p_h(S) & \text{for } S < S^* \\ p_l(S) & \text{for } S \geq S^* \end{cases} \tag{3.5}
\]

If, in addition, we impose this \( p \) to be non-negative and sum up to 1 (so \( P_l(S^*) := \int_{-\infty}^{S^*} p_l(s) ds = 1 - \int_{S^*}^{\infty} p_h(s) ds := P_h(S^*) \)), then

\[
\text{Swpt}(t; K, \phi) = A_{\alpha,\beta}(t) \cdot \int_{-\infty}^{\infty} (\phi[s - K])^+ p(s) ds \tag{3.6}
\]

is an arbitrage-free formula for pricing swaptions.

Similarly,

\[
\text{Swpt}_{l/h}(t; K, \phi) = A_{\alpha,\beta}(t) \cdot \int_{-\infty}^{\infty} (\phi[s - K])^+ p_{l/h}(s) ds. \tag{3.7}
\]

Finally, in order to be able to work reasonably with swaption formulas, we need continuity:

\[
\int_{-\infty}^{S^*} (S^* - s)^+ p_l(s) ds = \int_{S^*}^{\infty} (s - S^*)^+ p_h(s) ds - [S_{\alpha,\beta}(t) - S^*]. \tag{3.8}
\]

We refer to [22] for a detailed proof that these conditions suffice to imply the desired properties.
3.3 Shifted Lognormal Low-Strike Extrapolation

The previous section gives the general framework that was announced in the introduction of this chapter. We still have to define \( p_l \) and \( S^* \).

As announced before, we will shift the forward and strike by a shift parameter \( \lambda \) to deal with the assumption we make, namely that the distribution is shifted lognormal:

\[
p_l(s) = \frac{1}{\nu_{\lambda}(s + \lambda)} \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{\left[ \ln\left( \frac{s_{\alpha,\beta}(t) + \lambda}{s + \lambda} \right) - \frac{\nu^2}{2} \right]^2}{2\nu^2_{\lambda}} \right\}.
\]  

The problem is the choice of \( S^* \). In [22], \( S^* \) is chosen to be the lowest strike for which liquid swaption quotes are available, but this makes the method very temporary. What if we define and use the method today, and tomorrow lower strikes are observed? A method which can change on a day-to-day basis but has to price swaptions that run for sometimes 10 years is disputable. This problem is acknowledged by the authors.

Receiver swaption prices were expressed in basis points and then used to express normalized digital receiver swaption prices:

\[
Swpt_{bps}(t; K, -1) = \frac{Swpt(t; K, -1)}{An_{\alpha,\beta}(t)} = \int_{-\infty}^{K} [K - s]^+ p(s)ds;
\]

\[
DigSwpt_{bps}(t; K, -1) = \frac{\partial Swpt_{bps}(t, K, -1)}{\partial K} = \int_{-\infty}^{K} p(s)ds.
\]

3.4 Experiments and Conclusion

The market data, from September 19, 2014, calibrated to a SABR model, are in Figure 8.2 (\( \beta = 50\% \) everywhere). For both the 2Yx2Y and the 10Yx10Y swaptions, first the prices that result from the shifted SABR are compared with those of the regular SABR and the implied densities and volatilities are compared with those of SABR (Figures 8.3 and 8.5). To give some attention to the subjectively chosen \( S^* \), the choice of \( S^* \) is also varied, and the results are shown in Figures 8.4 and 8.6.

In short, the results are reasonably acceptable, but do present us with several drawbacks.

The implied probability rate for the forward swap rate has been extended to the negative area, as we can see in Figure 8.3. However, the most important drawback is that the cut-off rate is too arbitrary. The research team of Schlenkrich et al. failed to invent any parametric form to let \( S^* \) depend on the maturity. Although they do show different results in prices, volatilities and standard deviations for different choices of \( S^* \) (Figure 8.4 and Figure 8.6), the comparison with the actual prices, volatilities and standard deviations are not made. So it seems to be impossible to determine the right cut-off or cut-off function that can be sufficiently used to price financial derivatives. Another problem is that the focus of this method lies on avoiding zero strike swaptions to have price $0. This objective has obviously been reached, as becomes apparent from Figure 8.5, but the way in which this has...
been done is too arbitrarily chosen to be completely satisfied with the method. We do note that there are other elements of the results that do indeed satisfy. Most importantly, the method successfully solved the problem of negative densities, as can be seen in Figure 8.5. With Shifted SABR, they have become positive, which was one of the objectives.

We conclude that without the presence of a more satisfying solution, we might temporarily stick to the Shifted SABR. We decided to search for a better solution, bearing this one in mind in case we do not find an improvement.
Chapter 4

The Free Boundary SABR (2015)

4.1 Introduction

Another possible solution is the Free Boundary SABR, which was presented in [4], as a follow-up to [2] and [3], adapting it to negative rates without needing a problematic cut-off point as in the Shifted SABR. Unfortunately, after presenting the method and results in [4] and analyzing them, we must conclude that this method is unsatisfactory.

4.2 The Model

The name of the model refers to the boundary conditions of the SABR Model, and the absence of these conditions in the Free Boundary SABR. Instead of making a shift, the idea is to take the absolute value of the rate to deal with negative ones and still be able to use SABR.

When the classical SABR Model is denoted as

\[ dF_t = F_t^\beta \nu_t dW_1; \]
\[ d\nu_t = \gamma \nu_t dW_2, \]

(4.1)

where \( F \) is the forward rate, \( \nu \) the stochastic volatility and \( W_1 \) and \( W_2 \) are Wiener Processes, the following change is made:

\[ dF_t = |F_t|^\beta \nu_t dW_1; \]
\[ d\nu_t = \gamma \nu_t dW_2, \]

(4.2)

with \( \beta \in [0, \frac{1}{2}) \).

Half of the paper is dedicated to sections that the authors included to “get intuition.” When they arrive at the Free Boundary SABR itself, first an option
pricing formula is given for the case where \( W_1 \) and \( W_2 \) are uncorrelated:

\[
O_{\text{F}}^{SABR}(T, K) = \frac{1}{\pi} \sqrt{KF_0} \left\{ 1_{K \geq 0} A_1 + \sin(\nu|\phi|) A_2 \right\},
\]

with

\[
A_1 = \int_0^\pi \frac{d\phi}{\sin(\nu|\phi|)} \frac{\sin(\nu|\phi|) G(T, \gamma^2, s(\phi))}{\cosh(\phi)},
\]

and

\[
A_2 = \int_0^\infty d\phi \sinh(\nu|\phi|) \left( 1_{K \geq 0} \cosh(\nu|\phi|) + 1_{K < 0} \sinh(\nu|\phi|) \right) \frac{G(T, \gamma^2, s(\phi))}{\cosh(\phi)}.
\]

(4.3)

Here, \( T \) is the maturity, and \( \gamma \) the volatility of the volatility.

Then, they move on to the case where \( W_1 \) and \( W_2 \) are correlated, where they use a trick which is often referred to as the map: the general correlation option price is approximated using the zero-correlation price - the parameters are different, but the initial forward price is not. This trick was already introduced in [2]. We refer to the article for the details, and focus on the results of the experiments in this thesis.

4.3 Experiments and Conclusion

To test the method, Antonov et al. chose two different inputs and calculated the implied volatility, both with Monte Carlo simulations and with their own method. The inputs are in Figure 8.7 and the results are in Figure 8.8.

The authors are remarkably satisfied with these results, although their comments on them are brief: “We observe and excellent approximation quality for 3Y and positive strikes \( K > \frac{1}{2} F_0 \) for 10Y and a light degeneration for other strikes. We see that the normal implied volatility possesses significant smiles with the bottom between zero and the ATM strike. In general, increasing the volatility-of-volatility and maturity moves the vertex of the smile to the ATM strike.”

After these comments, the section ends. We do agree that, for a 3 year maturity, the results are decent, but for the 10 year maturity, more significant differences with Monte Carlo simulations already arise, which are referred to as “a light degeneration”. The differences are small for positive strikes and get larger for negative strikes. That the differences are small for positive strikes is not surprising, since the model is designed such that as few changes as possible for regular strikes are made to the SABR method, which has already been successfully tested and used for years.

However, the strikes which might or might not pose problems are the strikes we are interested in. The authors made an excellent choice by studying implied volatilities: if the results satisfy, we can use this method for a broad range of financial products, since the implied volatility is often the only non-trivial input. But the differences get larger for the rates we would actually have liked to be able to deal
with. If this starts happening for 10 year maturities already, we cannot rely on
the method to produce satisfying results for financial products in general. The
differences rise above 2% of the implied volatilities that result from using Monte
Carlo simulations. This is not a difference between an actual valuation of a financial
derivative, but a difference between the volatilities that will be used to valuate
derivatives, which makes matters even worse. This makes us conclude that we still
need another solution.

In the next chapter, we will discuss a more promising method.
Chapter 5

The Method of Hull and White (2014)

5.1 Introduction

So far, we have tried and failed to modify the SABR method in a satisfactory way, but there is a different path. Hull and White made a generalization of the Hull-White Model that seems to be able to deal with the problem reasonably well. We present their recent work here.

First, recall the concept of trinomial trees in Section 2.4.5. The corresponding model choice (the Hull and White Model) is justified there, but in [16], a more general procedure is proposed, in which not only the Hull and White Model can be used, but also other one-factor models.

The material in Section 2.4.5 is summarized in Section 2 of [16], which uses slightly different notation, considering models of the form:

\[ dx = [\theta(t) - ax]dt + \sigma dz, \]

with \( x = f(r) \). The tree is first built for \( x^* \), and then for \( x \), in the same way as before. This known procedure works only for an \( x \) with linear drift. The new, generalized procedure works for general one-factor models of the short rate. We will discuss this in the next section.

5.2 A Generalized Tree Building Procedure

5.2.1 Model

As a general assumption about the short rate \( r \), we take

\[ dr = [\theta(r) + F(r)]dt + G(r)dz. \]

As before, \( \theta \) is chosen to fit the initial term structure. \( F \) determines the drift of \( r \), and \( G \) determines the volatility of \( r \) and should be continuously differentiable. We define \( x = f(r) \), where \( f \) is the primitive of \( \frac{1}{G} \), so that \( \frac{dx}{dr} = \frac{1}{G(r)} \). \( x \) follows the
process
\[dx = H(x, t)dt + dz,\]  
(5.3)

with drift
\[H(x, t) = \frac{\theta(t) + F(r)}{G(r)} - \frac{1}{2} G'(r).\]  
(5.4)

5.2.2 New Properties

Notice that, in contrast to the model as we know it, we do not first build a tree which we shift. Instead, we create a rectangular grid, with horizontal time steps \(\Delta t\) and vertical rate steps \(\Delta x\). In each time step, three adjacent nodes can be reached, but now there is no restriction to which nodes. Finally, the branching process and the branching probabilities depend on \(\theta(t)\).

5.2.3 Description

Let us give a more detailed description of the methodology. We start with \(x_0 = f(r_0)\), where \(r_0\) is the initial interest rate. The rate values are \(x_0 + k\Delta x\), where each \(k\) is a possibly negative integer, and \(\Delta x = \sqrt{3\Delta t}\). Denote by \((i, j)\) the node at time \(i\Delta t\) and rate \(x = x_0 + j\Delta x\).

To branch from \(i\Delta t\) to \((i+1)\Delta t\), Hull and White first assume \(\theta\) to be known by assigning it a trial value. Then they branch and calculate the implied price of a zero-coupon bond. Then, \(\theta\) is chosen such that the price of such a bond is consistent. This is equivalent to making \(\theta\) variable, letting the prices depend on it, and then choosing the one that is consistent with the bond price.

The branching from node \((i, j)\) is done by moment matching. This is not a new concept, and can be found e.g. in [23]. The first and second moment of \(x\) at \((i+1)\Delta t\), conditional on the eventuality that \(x\) is at node \((i, j)\) at time \(i\Delta t\), are defined as \(m_1\) and \(m_2\), respectively, and are correct for \(\Delta t \rightarrow 0\) if:

\[m_1 = f(r_j + (\theta + F(r_j) - 0.5G(r_j)G'(r_j))\Delta t);\]
\[m_2 = \Delta t + m_1^2,\]  
(5.5)

where \(r_j\) is the \(\Delta t\) rate at node \((i, j)\).

Now we choose the three nodes that can be reached from \((i, j)\). We pick \((i + 1, j^* - 1), (i + 1, j^*)\) and \((i + 1, j^* + 1)\), where \(j^*\) is chosen in such a way that \(x_0 + j^*\Delta x\) is as close as possible to \(m_1\):

\[j^* = \text{int}(\frac{m_1 - x_0}{\Delta x} + 0.5).\]  
(5.6)

The function int rounds off any input to the the largest integer which is at most the input. We could have chosen \(j^* = \text{ROUND}(\frac{m_1 - x_0}{\Delta x})\), but we prefer to respect the choice of the authors.
Next, we choose the branching probabilities \( p_d, p_m \) and \( p_u \) such that the moments are matched:

\[
p_u = \frac{\Delta t + (m_1 - x_0 - j^* \Delta x)^2}{2 \Delta x^2} + \frac{m_1 - x_0 - j^* \Delta x}{2 \Delta x};
\]

\[
p_d = \frac{\Delta t + (m_1 - x_0 - j^* \Delta x)^2}{2 \Delta x^2} - \frac{m_1 - x_0 - j^* \Delta x}{2 \Delta x};
\]

\[
p_m = 1 - p_u - p_d.
\]

We should have \( p_u, p_d, p_m \in [0, 1] \), of course, and we do: see Appendix B. Not only do we prove the statement that the three probabilities are between 0 and 1; we get the additional information that \( p_m \geq \frac{1}{3} \).

Recall the Arrow Debreu price of a node \((i, j)\) in Section 2.4.5, which is defined as the price \( Q_{i,j} \) of the derivative that pays of $1 if the node is reached, and $0 otherwise. As before, we start with \( Q_{0,0} = 1 \) and work through the nodes at \( \Delta t \), \( 2\Delta t \) etc. In this way, all the \( Q_{i,j} \) are determined. If we denote the probability of moving from node \((i-1,k)\) to \((i,j)\) by \( q(k,j) \) - notice that \( q \) is only positive for three values of \( j \) - and \( j_d(i) \) and \( j_u(i) \) are the lowest and highest values of \( j \) that can be reached at \( i \Delta t \), then

\[
Q_{i,j} = \sum_{k=j_d(i-1)}^{j_u(i-1)} Q_{i-1,k} q(k,j) e^{-r_k \Delta t},
\]

where \( r_k = f^{-1}(x_0 + k \Delta x) \). Then the price of a zero-coupon bond with maturity \((i + 1) \Delta t\) is

\[
P_{i+1} = \sum_{j=j_d(i)}^{j_u(i)} Q_{i,j} e^{-r_j \Delta t}.
\]

Both \( Q \) and \( P \) depend on \( m_1 \) and thus on \( \theta \) - we choose the value of \( \theta \) for which \( P_{i+1} \) is consistent with the initial term structure, and then move on to the next step. In this way, the complete tree is constructed.

### 5.2.4 Examples

To illustrate the procedure we just described, we consider two examples from [16] - a simple example to get a feeling for the material, and a more realistic example.

**A Simple Example**

Let us choose:

- \( F(r) = -ar \)
- \( G(r) = \sigma r \),
with \( a, \sigma > 0 \), so that the short rate follows the process

\[
dr = \left[ \theta(t) - ar \right] dt + \sigma dz. \tag{5.10}
\]

- \( x = \frac{\ln r}{\sigma} \)
- \( r = e^{\sigma x} \)
- \( \Delta t = 0.5 \) years
- \( a = 0.2 \)
- \( \sigma = 0.15 \)

The results are in Figure 8.9. It is a very basic example to introduce the method, so we obviously cannot draw any realistic conclusions from it, but we included this nonetheless, because it adequately illustrates the different aspects of the method - the values of \( \theta \), the transition probabilities and the Arrow Debreu Prices are listed, and the resulting interest rates tree is shown. Let us continue with a more elaborate and more realistic example.

**A More Realistic Example**

More realistically, let

\[
G(r) = \begin{cases}
    s \left[ \frac{2r}{r_1} - \left( \frac{r}{r_1} \right)^2 \right] & \text{when } r \leq r_1; \\
    s + K(r - r_1)^2 & \text{when } r_1 < r \leq r_2; \\
    \alpha + \beta r & \text{when } r > r_2.
\end{cases} \tag{5.11}
\]

Here, \( K \) is some positive constant. Continuity and differentiability is ensured at \( r_1 \) automatically and at \( r_2 \) by setting \( K = \frac{\beta}{2(r_2 - r_1)} \) and \( \alpha = s + K(r_2 - r_1)^2 - \beta r_2 \). The choice of the function is based on [19], with drift \( a = 0.05 \) and:

- \( r_1 = 0.02 \)
- \( r_2 = 0.1 \)
- \( s = 0.02 \)
- \( \beta = 0.2 \)

Then \( K = 1.25 \) and \( \alpha = 0.008 \). From (5.11) it follows that

\[
x(r) = \begin{cases}
    \frac{r}{2s} \ln \frac{r}{r_3 - r} & \text{when } r \leq r_3; \\
    \frac{1}{\sqrt{sK}} \tan^{-1} \left( (r - r_3) \sqrt{\frac{K}{s}} \right) & \text{when } r_3 < r \leq r_4; \\
    \frac{1}{\beta} \ln(\alpha + \beta r) + C & \text{when } r > r_4,
\end{cases} \tag{5.12}
\]

with, to ensure continuity,

\[
C = -\frac{1}{\beta} \ln(\alpha + \beta r_4) + \frac{1}{\sqrt{sK}} \tan^{-1} \left( (r_4 - r_3) \sqrt{\frac{K}{s}} \right), \tag{5.13}
\]

38
and inverse function

\[
 r(x) = \begin{cases} 
 2r_3 & \text{when } x \leq 0; \\
 r_3 + \frac{1+e^{-sx}}{s^2K} \tan(x\sqrt{sK}) & \text{when } 0 < x \leq x_4; \\
 \frac{r_3}{\frac{r_4}{r_3} - \frac{r_4}{r_3} - \alpha}{\beta} & \text{when } x > x_4,
\end{cases}
\]  

with

\[
 x_4 = \frac{1}{\sqrt{sK}} \tan^{-1} \left( (r_4 - r_3) \sqrt{\frac{K}{s}} \right). 
\]

Hull and White valued caps with this example, as shown, as well as the interest rates, in Figure 8.10. These results are not yet compared with actual prices, since this example simply builds on the method, assuming a yet more realistic structure than Example 1. We do already see the values taking more realistic values, but to produce genuinely convincing results, we need an even more sophisticated \( G \), which is what Hull and White did in the remainder of [16], which we will present now.

### 5.3 Market Implied Volatility Function

Pointing out that something like \( G(r) = \sigma \sqrt{r} \) is possible, Hull and White let \( G \) instead be an almost piece-wise linear function, where the pieces lie between points \( r_i \) (\( i \in \{1, \ldots, n\} \)) with \( G(r_i) =: s_i \), and a rounding procedure at these points is applied to preserve continuous differentiability. Calibration in this case means setting the values of \( s_i \) such that market prices are best matched. As an example, they applied this to the model with \( F(r) = ar \),

\[
dr = (\theta(t) - ar)dt + G(r)dz, 
\]

chose \( a = 0.05 \), and observed implied prices of 10-year caps, observed in 2013. Let us give a more detailed description of \( G \). If the function is divided up in points \( r_i \) with \( G(r_i) = s_i \), and the rounding region around every point has radius \( \epsilon \), then they derive:

\[
 G(r) = \begin{cases} 
 a_0 + b_0r & 0 \leq r \leq r_1 - \epsilon \\
 x_1 + y_1r + z_1r^2 & r_1 - \epsilon \leq r \leq r_1 + \epsilon \\
 a_1 + b_1r & r_1 + \epsilon \leq r \leq r_2 - \epsilon \\
 x_2 + y_2r + z_2r^2 & r_2 - \epsilon \leq r \leq r_2 + \epsilon \\
 \cdots & \cdots 
\end{cases} 
\]  

(5.17)

where

- \( a_i = \frac{s_i r_{i+1} - s_{i+1} r_i}{r_{i+1} - r_i} \);
- \( b_i = \frac{s_{i+1} - s_i}{r_{i+1} - r_i} \);
- \( z_{i+1} = \frac{b_{i+1} - b_i}{4\epsilon} \);
- \( y_{i+1} = b_i - 2z_{i+1} (r_{i+1} - \epsilon) \);
They also give $x(r)$ and $r(x)$, as well as the term structure, corner points and cap prices, which we also present here in Figure 8.11, where the caps, originating from December 2013, have annual payments, cap rate 0.04 and a principal of $100. Also, $\{r_1, ..., r_7\} = \{0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.10\}$ (so $N = 7$), and the goodness-of-fit function was $\sum_{i=1}^{N} \frac{(U_i - V_i)^2}{U_i}$, where $U_i$ and $V_i$ are the market and model price of cap $i$. Subsequently, the market prices are compared with the real market prices and the market prices that would result when applying the Hull-White (normal) Model, to which the model that they generalized was restricted, and the Black-Karasinski (lognormal) Model. See 8.12.

The results are quite satisfying. The Black-Karasinski (lognormal) Model is the model which has been used successfully throughout the years, so the choice of comparing the results with this method is a convincing one. And as we can see, the quality of the results comes remarkably close to the quality of the Black-Karasinski Model, which performs only slightly superior to the new method of Hull and White, especially for low strikes. This also justifies the choice for the lognormal model in the case where negative rates are impossible. But since we lack this choice now, we the new Hull and White Model seems to be our best alternative choice. One problem did rise: the procedure was time-consuming. To tackle this problem, the procedure was repeated with $N = 4$ and $\{r_1, ..., r_4\} = \{0.01, 0.05, 0.10\}$, and this did not deteriorate the results: see Figure 8.13. This is a convincing success, since not only does the method produce satisfying results, it is also practically applicable.

### 5.4 Conclusion

Of the possible solutions to the problem of rates becoming negative, the recent adaptation of the method of Hull and White seems to be the most promising. In the past, the method had the drawback of allowing negative rates, which has now become an advantage. The results in this chapter show that the new method works very well. We hope that Hull and White will extend their extensive numerical work to even more recent data than they already did, to make sure that this method will indeed show no problems for negative rates. If we would have another 6 months to work on this subject, we might have the time to work out even that. Now we leave that as a suggestion for further research - perhaps for another graduate intern at EY.
Chapter 6

Summary and Conclusion

We searched for an adequate solution to the problem of negative interest rates. This is a problem, because existent valuing software do not work or simply crash when the input contains negative rates. First, we stated the necessary preliminaries, including the Black-Scholes-Merton formula, the SABR method and the Hull-White Model. We did so, because SABR is the most commonly used model in the financial industry, so we would like to keep using the model, or an adaptation of the model. Then, we considered the Shifted SABR, a simple version of SABR, and the Free Boundary SABR, a more recent version of SABR, with disappointing results, although in the absence of a satisfactory model, the Shifted SABR is acceptable. Finally, we arrived at the Hull and White methodology, which was recently majorly improved by Hull and White and seems to be the most promising.

Their results were based on fairly recent data, but still data where negative rates did not yet occur. Hence, further research is needed to be completely conclusive about the strong suspicion that their method is the best candidate for dealing with negative rates.
Chapter 7

Suggestions

One thing that we have not yet mentioned is that it might be possible that there is in fact an adaptation to SABR that can deal with negative rates in a satisfying way. We decided to compare existent solutions, since it usually takes several professionals combining their strength to develop such a method, making it unlikely that a single student can develop one in 6 months. This does not mean that it is certainly impossible to use some version of SABR. We can only say that it has become less likely, since both the first, simple try and the second try failed to satisfy.

As mentioned before, the work in Chapter 5 should be extended to be tested to even more recent data, to make sure that negative rates are no problem in practice. Until that is done, the Shifted SABR model is a reasonable alternative for the time being.

Additionally, what we would like to bring to the reader’s attention is the follow-up paper of [16], which is [18]. In that article, market observations are used to deal with “the local price of risk.” In the past, the focus lay mostly on the theoretical Q-measure, but recently, attention has shifted more towards the real-world P-measure. In their paper, Hull and White present an interest rate model that represents the risk evolution under both \( P \) and \( Q \), remarking that approximating the real-world drift of the rate based on risk-neutral drift is generally questionable. “Market prices of contingent claims are independent of investor risk preferences.” Hence observable market prices alone do not determine real-world risk preferences. But they can be estimated with historical data in addition to market prices. When the risk-free model is

\[
dr = \mu(t,r)dt + \sigma(t,r)dz,
\]

where \( \mu \) and \( \sigma \) are deterministic and \( z \) a Wiener Process, then the corresponding real-world model is

\[
dr = (\mu(t,r) + \lambda(t,r)\sigma(t,r))dt + \sigma(t,r)dz,
\]

where \( \lambda \) is the market price of interest rate risk. In the real world, risk aversion exists, implying a higher bond price drift, which is equivalent to a lower interest rates drift, hence \( \lambda(t) < 0 \). At the least, \( \lambda \) should be assumed to depend on \( t \). This overcomes shortcomings of using a one-factor model. Generally, \( \lambda \) depends on \( r \) as well and may also depend on other variables, which explains the notation.
"\(\lambda(t, r, \ldots)\)." Finally, market implied data can be incorporated into the estimation of drift and volatility, even when we need real-world computations.

Earlier estimations of interest rate risk market prices are claimed to be inconsistent when looking 30 years ahead. For models where \(\lim_{r \to 0} \sigma(t, r) = 0\), we get \(\lambda = -1.0\), and "the expected future real-world rate in 30 years is unrealistically close to zero." We have to simply believe this in order to continue, because exact numbers are not given. This part is also questionable because rates are currently possibly negative, so future negative rates do not seem as impossible as they used to seem in the past. That is exactly what this thesis has been about. Strangely, they do comment that these low and sometimes negative rates are "also apparent in empirical analysis of interest rate data," but make no further comments about it. This is the main reason for not including this paper in the regular body of our thesis, but it might be of interest, hence the appearance in this section.

In any case, if \(\sigma\) is constant and the maturity is \(T\), it is shown that

\[
\lambda(T) = -\frac{F(T) - r_0}{\sigma T}.
\]

Using [13], the historical price of risk for 1982-2014 with mean reversion rates 0\%, 5\% and 10\% are estimated, as well as those for the two separate periods 1982-1997 and 1998-2014. See Figure 8.14. We can observe the phenomenon of risk per time unit decreasing in time by observing the decreasing price of risk.

The model that is used is the one described in Chapter 5, and is a mean-reverting version of the model in [19].

Assumption:

\[
dr = [\theta(t) - ar]dt + \sigma(r)dz,
\]

where \(a = 0.05\) (the mean reversion parameter), and

\[
\sigma(r) = \begin{cases} 
\frac{sr}{0.015} & r \leq 0.015 \text{ (lognormal volatility)}; \\
\frac{s}{0.015} & 0.015 < r < 0.06 \text{ (normal volatility)}; \\
\frac{sr}{0.06} & r \geq 0.06 \text{ (lognormal volatility)},
\end{cases}
\]

which leaves only the parameter \(s\) free, which isn’t defined, but appears to be the initial volatility.

The goal is to estimate the market price of risk that is consistent with the assumed model from historical data.

1. First estimate the historical average term structure.

2. Then estimate the average value of \(s\), in the article since 1982 - it turns out to be 1.05\%.

3. Finally, calculate the market price in each time step: if we have a market price of risk \(\lambda_i\) and at node \(j\) a risk-neutral short rate drift of \(\mu_{ij}\), then the real-world drift is \(\mu_{ij} + \lambda_i \sigma_{ij}\). The historical average short rate was 4.4\%, so just advance through the tree, searching for the \(\lambda_i\) that reduces the expected short rate at step \(i\), calculated using the tree, to 4.4\%.
Now we want to calibrate the term structure model parameters to current market data. To do so, the method in Chapter 5 is used, again with data from December 2013, resulting in $s \approx 2.1\%$, which is also consistent with current market data. The assumption is made that historical local prices of risk also apply in the future.

We could plug in other short rates. For the long-run rate, rates observed in time periods of at least 15 years suffice.

The results are in figures 8.15, 8.16 and 8.17. They show predictions for the short rate, as well as expectations and standard deviations, depending on the maturity of interest rates. This is the first time this is done, so combining the risk-neutral drift with the local price of risk may replace the approximations that are currently widely used. Notice that low rates in the distant future are predicted by this model, but not yet negative rates. So there is still work to do on this new subject Hull and White have started to investigate. If we choose their path, this material might be worth delving into more.
Chapter 8

Results
Figure 8.1: Generated Paths of the Short rate (discretized Vasicek Model)
### 6M Euribor Forward Swap Rate

<table>
<thead>
<tr>
<th>Expiry \ Swap Term</th>
<th>1Y</th>
<th>2Y</th>
<th>5Y</th>
<th>10Y</th>
<th>15Y</th>
<th>20Y</th>
<th>30Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>3M</td>
<td>7.21%</td>
<td>3.67%</td>
<td>3.98%</td>
<td>4.22%</td>
<td>3.80%</td>
<td>3.83%</td>
<td>3.77%</td>
</tr>
<tr>
<td>6M</td>
<td>6.14%</td>
<td>3.36%</td>
<td>4.03%</td>
<td>4.16%</td>
<td>3.83%</td>
<td>3.83%</td>
<td>3.77%</td>
</tr>
<tr>
<td>1Y</td>
<td>5.19%</td>
<td>3.78%</td>
<td>4.29%</td>
<td>4.30%</td>
<td>3.94%</td>
<td>3.83%</td>
<td>3.76%</td>
</tr>
<tr>
<td>2Y</td>
<td>5.04%</td>
<td>4.49%</td>
<td>4.64%</td>
<td>4.48%</td>
<td>4.08%</td>
<td>3.93%</td>
<td>3.72%</td>
</tr>
<tr>
<td>5Y</td>
<td>5.35%</td>
<td>5.01%</td>
<td>4.72%</td>
<td>4.35%</td>
<td>4.00%</td>
<td>3.87%</td>
<td>3.57%</td>
</tr>
<tr>
<td>10Y</td>
<td>4.55%</td>
<td>4.19%</td>
<td>4.08%</td>
<td>3.96%</td>
<td>3.72%</td>
<td>3.56%</td>
<td>3.21%</td>
</tr>
<tr>
<td>15Y</td>
<td>3.98%</td>
<td>3.92%</td>
<td>3.81%</td>
<td>3.69%</td>
<td>3.42%</td>
<td>3.21%</td>
<td>2.89%</td>
</tr>
</tbody>
</table>

### SABR Initial Volatility $\alpha$

<table>
<thead>
<tr>
<th>Expiry \ Swap Term</th>
<th>1Y</th>
<th>2Y</th>
<th>5Y</th>
<th>10Y</th>
<th>15Y</th>
<th>20Y</th>
<th>30Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>3M</td>
<td>25.9%</td>
<td>48.4%</td>
<td>46.1%</td>
<td>19.2%</td>
<td>2.4%</td>
<td>14.2%</td>
<td>13.8%</td>
</tr>
<tr>
<td>6M</td>
<td>12.7%</td>
<td>47.1%</td>
<td>47.0%</td>
<td>16.9%</td>
<td>1.3%</td>
<td>14.2%</td>
<td>16.3%</td>
</tr>
<tr>
<td>1Y</td>
<td>67.4%</td>
<td>49.8%</td>
<td>39.9%</td>
<td>9.7%</td>
<td>1.9%</td>
<td>13.3%</td>
<td>17.2%</td>
</tr>
<tr>
<td>2Y</td>
<td>106.0%</td>
<td>47.5%</td>
<td>16.6%</td>
<td>1.5%</td>
<td>4.8%</td>
<td>10.5%</td>
<td>10.5%</td>
</tr>
<tr>
<td>5Y</td>
<td>10.5%</td>
<td>3.4%</td>
<td>0.8%</td>
<td>3.4%</td>
<td>4.0%</td>
<td>4.5%</td>
<td>5.6%</td>
</tr>
<tr>
<td>10Y</td>
<td>12.5%</td>
<td>5.9%</td>
<td>-5.3%</td>
<td>-10.6%</td>
<td>-8.8%</td>
<td>-7.9%</td>
<td>-6.7%</td>
</tr>
<tr>
<td>15Y</td>
<td>21.3%</td>
<td>13.7%</td>
<td>-11.8%</td>
<td>-13.8%</td>
<td>-14.3%</td>
<td>-13.1%</td>
<td>-15.1%</td>
</tr>
</tbody>
</table>

### SABR Correlation $\rho$

<table>
<thead>
<tr>
<th>Expiry \ Swap Term</th>
<th>1Y</th>
<th>2Y</th>
<th>5Y</th>
<th>10Y</th>
<th>15Y</th>
<th>20Y</th>
<th>30Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>3M</td>
<td>133.8%</td>
<td>76.2%</td>
<td>86.4%</td>
<td>98.5%</td>
<td>109.2%</td>
<td>119.4%</td>
<td>119.2%</td>
</tr>
<tr>
<td>6M</td>
<td>102.0%</td>
<td>76.2%</td>
<td>52.9%</td>
<td>64.8%</td>
<td>64.9%</td>
<td>65.1%</td>
<td>70.6%</td>
</tr>
<tr>
<td>1Y</td>
<td>68.8%</td>
<td>65.9%</td>
<td>45.6%</td>
<td>45.5%</td>
<td>44.0%</td>
<td>42.4%</td>
<td>47.3%</td>
</tr>
<tr>
<td>2Y</td>
<td>35.2%</td>
<td>40.5%</td>
<td>33.7%</td>
<td>34.0%</td>
<td>35.0%</td>
<td>35.8%</td>
<td>45.2%</td>
</tr>
<tr>
<td>5Y</td>
<td>44.5%</td>
<td>33.9%</td>
<td>32.5%</td>
<td>32.2%</td>
<td>30.7%</td>
<td>20.3%</td>
<td>24.4%</td>
</tr>
<tr>
<td>10Y</td>
<td>35.2%</td>
<td>33.8%</td>
<td>32.2%</td>
<td>33.4%</td>
<td>29.7%</td>
<td>26.2%</td>
<td>24.4%</td>
</tr>
<tr>
<td>15Y</td>
<td>37.8%</td>
<td>33.3%</td>
<td>28.7%</td>
<td>32.0%</td>
<td>30.1%</td>
<td>26.2%</td>
<td>24.6%</td>
</tr>
</tbody>
</table>

### SABR Volatility of Volatility $\nu$

Figure 8.2: Market Data
Figure 8.3: 2Yx2Y swaptions: Shifted Lognormal compared to SABR

Figure 8.4: 2Yx2Y swaptions: different cut-offs
Figure 8.5: 10Yx10Y swaptions: Shifted Lognormal Wing compared to SABR

Figure 8.6: 2Yx2Y swaptions: different cut-offs
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Value for Input I</th>
<th>Value for Input II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rate Initial Value</td>
<td>$F_0$</td>
<td>50 bps</td>
<td>1%</td>
</tr>
<tr>
<td>SV Initial Value</td>
<td>$\nu_0$</td>
<td>$0.6 F_0^{1-\beta}$</td>
<td>$0.3 F_0^{1-\beta}$</td>
</tr>
<tr>
<td>Vol-of-Vol</td>
<td>$\gamma$</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>Correlations</td>
<td>$\rho$</td>
<td>-0.3</td>
<td>-0.3</td>
</tr>
<tr>
<td>Skews</td>
<td>$\beta$</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>Maturities</td>
<td>$T$</td>
<td>3Y</td>
<td>10Y</td>
</tr>
</tbody>
</table>

Figure 8.7: Parameter Values for Free Boundary SABR

<table>
<thead>
<tr>
<th>K</th>
<th>Analyt</th>
<th>Exact</th>
<th>Diff</th>
<th>Analyt</th>
<th>Exact</th>
<th>Diff</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.95</td>
<td>30.87</td>
<td>30.93</td>
<td>-0.06</td>
<td>40.05</td>
<td>40.86</td>
<td>-0.81</td>
</tr>
<tr>
<td>-0.8</td>
<td>29.83</td>
<td>29.95</td>
<td>-0.12</td>
<td>38.43</td>
<td>39.24</td>
<td>-0.81</td>
</tr>
<tr>
<td>-0.65</td>
<td>28.80</td>
<td>28.97</td>
<td>-0.17</td>
<td>36.80</td>
<td>37.60</td>
<td>-0.80</td>
</tr>
<tr>
<td>-0.5</td>
<td>27.79</td>
<td>27.99</td>
<td>-0.20</td>
<td>35.18</td>
<td>35.97</td>
<td>-0.78</td>
</tr>
<tr>
<td>-0.35</td>
<td>26.83</td>
<td>27.04</td>
<td>-0.21</td>
<td>33.59</td>
<td>34.33</td>
<td>-0.74</td>
</tr>
<tr>
<td>-0.2</td>
<td>25.95</td>
<td>26.15</td>
<td>-0.20</td>
<td>32.05</td>
<td>32.73</td>
<td>-0.68</td>
</tr>
<tr>
<td>-0.05</td>
<td>25.30</td>
<td>25.46</td>
<td>-0.16</td>
<td>30.67</td>
<td>31.25</td>
<td>-0.58</td>
</tr>
<tr>
<td>0.1</td>
<td>25.77</td>
<td>25.85</td>
<td>-0.08</td>
<td>30.20</td>
<td>30.63</td>
<td>-0.43</td>
</tr>
<tr>
<td>0.25</td>
<td>26.63</td>
<td>26.69</td>
<td>-0.06</td>
<td>30.19</td>
<td>30.51</td>
<td>-0.31</td>
</tr>
<tr>
<td>0.4</td>
<td>27.33</td>
<td>27.39</td>
<td>-0.06</td>
<td>30.14</td>
<td>30.41</td>
<td>-0.27</td>
</tr>
<tr>
<td>0.55</td>
<td>27.90</td>
<td>27.97</td>
<td>-0.06</td>
<td>30.06</td>
<td>30.31</td>
<td>-0.25</td>
</tr>
<tr>
<td>0.7</td>
<td>28.38</td>
<td>28.45</td>
<td>-0.07</td>
<td>30.00</td>
<td>30.22</td>
<td>-0.23</td>
</tr>
<tr>
<td>0.85</td>
<td>28.80</td>
<td>28.87</td>
<td>-0.07</td>
<td>29.98</td>
<td>30.18</td>
<td>-0.20</td>
</tr>
<tr>
<td>1</td>
<td>29.18</td>
<td>29.25</td>
<td>-0.07</td>
<td>30.05</td>
<td>30.22</td>
<td>-0.17</td>
</tr>
<tr>
<td>1.15</td>
<td>29.53</td>
<td>29.60</td>
<td>-0.07</td>
<td>30.24</td>
<td>30.36</td>
<td>-0.12</td>
</tr>
<tr>
<td>1.3</td>
<td>29.87</td>
<td>29.94</td>
<td>-0.07</td>
<td>30.56</td>
<td>30.63</td>
<td>-0.07</td>
</tr>
<tr>
<td>1.45</td>
<td>30.22</td>
<td>30.29</td>
<td>-0.06</td>
<td>31.03</td>
<td>31.04</td>
<td>-0.01</td>
</tr>
<tr>
<td>1.6</td>
<td>30.58</td>
<td>30.63</td>
<td>-0.06</td>
<td>31.63</td>
<td>31.58</td>
<td>0.04</td>
</tr>
<tr>
<td>1.75</td>
<td>30.95</td>
<td>30.99</td>
<td>-0.05</td>
<td>32.35</td>
<td>32.26</td>
<td>0.09</td>
</tr>
<tr>
<td>1.9</td>
<td>31.33</td>
<td>31.37</td>
<td>-0.04</td>
<td>33.17</td>
<td>33.04</td>
<td>0.13</td>
</tr>
</tbody>
</table>

Figure 8.8: Implied volatilities Free Boundary SABR in bps - ATM bold. “Exact” means “With Monte Carlo simulation.”
Figure 8.9: Simple Example
Zero coupon interest rates for Example 2 and cap calibration

<table>
<thead>
<tr>
<th>Time (yrs)</th>
<th>Zero Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>0.20%</td>
</tr>
<tr>
<td>0.5</td>
<td>0.22%</td>
</tr>
<tr>
<td>1</td>
<td>0.27%</td>
</tr>
<tr>
<td>2</td>
<td>0.38%</td>
</tr>
<tr>
<td>3</td>
<td>0.67%</td>
</tr>
<tr>
<td>4</td>
<td>1.08%</td>
</tr>
<tr>
<td>5</td>
<td>1.52%</td>
</tr>
<tr>
<td>6</td>
<td>1.93%</td>
</tr>
<tr>
<td>7</td>
<td>2.27%</td>
</tr>
<tr>
<td>8</td>
<td>2.55%</td>
</tr>
<tr>
<td>9</td>
<td>2.78%</td>
</tr>
<tr>
<td>10</td>
<td>2.98%</td>
</tr>
<tr>
<td>15</td>
<td>3.60%</td>
</tr>
<tr>
<td>20</td>
<td>3.85%</td>
</tr>
<tr>
<td>25</td>
<td>3.96%</td>
</tr>
<tr>
<td>30</td>
<td>4.02%</td>
</tr>
</tbody>
</table>

Values of caps with annual payments on December 2, 2013 when volatility function is as shown in Figure 4, cap rate is 4%, and the principal is 100.

<table>
<thead>
<tr>
<th>Steps per year</th>
<th>Cap Life (yrs)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10</td>
</tr>
<tr>
<td>1</td>
<td>8.56</td>
</tr>
<tr>
<td>2</td>
<td>7.84</td>
</tr>
<tr>
<td>5</td>
<td>7.65</td>
</tr>
<tr>
<td>10</td>
<td>7.67</td>
</tr>
<tr>
<td>15</td>
<td>7.63</td>
</tr>
<tr>
<td>20</td>
<td>7.64</td>
</tr>
</tbody>
</table>

Figure 8.10: More Realistic Example
<table>
<thead>
<tr>
<th>Time (yrs)</th>
<th>Zero Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>0.20%</td>
</tr>
<tr>
<td>0.5</td>
<td>0.22%</td>
</tr>
<tr>
<td>1</td>
<td>0.27%</td>
</tr>
<tr>
<td>2</td>
<td>0.38%</td>
</tr>
<tr>
<td>3</td>
<td>0.67%</td>
</tr>
<tr>
<td>4</td>
<td>1.08%</td>
</tr>
<tr>
<td>5</td>
<td>1.52%</td>
</tr>
<tr>
<td>6</td>
<td>1.93%</td>
</tr>
<tr>
<td>7</td>
<td>2.27%</td>
</tr>
<tr>
<td>8</td>
<td>2.55%</td>
</tr>
<tr>
<td>9</td>
<td>2.78%</td>
</tr>
<tr>
<td>10</td>
<td>2.98%</td>
</tr>
<tr>
<td>15</td>
<td>3.60%</td>
</tr>
<tr>
<td>20</td>
<td>3.85%</td>
</tr>
<tr>
<td>25</td>
<td>3.96%</td>
</tr>
<tr>
<td>30</td>
<td>4.02%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Steps per year</th>
<th>Cap Life (yrs)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10</td>
</tr>
<tr>
<td>1</td>
<td>8.56</td>
</tr>
<tr>
<td>2</td>
<td>7.84</td>
</tr>
<tr>
<td>5</td>
<td>7.65</td>
</tr>
<tr>
<td>10</td>
<td>7.67</td>
</tr>
<tr>
<td>15</td>
<td>7.63</td>
</tr>
<tr>
<td>20</td>
<td>7.64</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$r_t$</th>
<th>0.0%</th>
<th>1.0%</th>
<th>2.0%</th>
<th>3.0%</th>
<th>4.0%</th>
<th>5.0%</th>
<th>6.0%</th>
<th>10.0%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_t$</td>
<td>0.00%</td>
<td>1.48%</td>
<td>1.68%</td>
<td>1.68%</td>
<td>1.80%</td>
<td>1.97%</td>
<td>2.33%</td>
<td>3.43%</td>
</tr>
</tbody>
</table>

Figure 8.11: Term Structure, Corner Points and Cap Prices
### Figure 8.12: Cap Prices Compared

<table>
<thead>
<tr>
<th>Cap Strike</th>
<th>Market Volatility</th>
<th>Market Price</th>
<th>Model Price</th>
<th>Price Difference</th>
<th>HW Price Difference</th>
<th>BK Price Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00%</td>
<td>50.75%</td>
<td>19.18%</td>
<td>18.87%</td>
<td>−0.31%</td>
<td>2.19%</td>
<td>−1.03%</td>
</tr>
<tr>
<td>2.00%</td>
<td>38.73%</td>
<td>14.06%</td>
<td>13.86%</td>
<td>−0.20%</td>
<td>1.89%</td>
<td>−1.27%</td>
</tr>
<tr>
<td>3.00%</td>
<td>32.30%</td>
<td>9.96%</td>
<td>9.99%</td>
<td>0.03%</td>
<td>1.60%</td>
<td>−0.89%</td>
</tr>
<tr>
<td>4.00%</td>
<td>30.15%</td>
<td>7.21%</td>
<td>7.12%</td>
<td>−0.09%</td>
<td>0.89%</td>
<td>−0.67%</td>
</tr>
<tr>
<td>5.00%</td>
<td>28.50%</td>
<td>5.19%</td>
<td>5.10%</td>
<td>−0.09%</td>
<td>0.28%</td>
<td>−0.39%</td>
</tr>
<tr>
<td>6.00%</td>
<td>26.50%</td>
<td>3.56%</td>
<td>3.68%</td>
<td>0.12%</td>
<td>−0.01%</td>
<td>0.05%</td>
</tr>
<tr>
<td>7.00%</td>
<td>25.72%</td>
<td>2.56%</td>
<td>2.63%</td>
<td>0.07%</td>
<td>−0.35%</td>
<td>0.17%</td>
</tr>
<tr>
<td>8.00%</td>
<td>25.50%</td>
<td>1.92%</td>
<td>2.01%</td>
<td>0.09%</td>
<td>−0.61%</td>
<td>0.19%</td>
</tr>
<tr>
<td>9.00%</td>
<td>25.55%</td>
<td>1.50%</td>
<td>1.46%</td>
<td>−0.04%</td>
<td>−0.76%</td>
<td>0.13%</td>
</tr>
<tr>
<td>10.00%</td>
<td>25.70%</td>
<td>1.20%</td>
<td>1.12%</td>
<td>−0.08%</td>
<td>−0.80%</td>
<td>0.10%</td>
</tr>
</tbody>
</table>

### Figure 8.13: Cap Prices for $N=4$

<table>
<thead>
<tr>
<th>Cap Strike</th>
<th>Market Volatility</th>
<th>Market Price</th>
<th>Model Price</th>
<th>Price Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00%</td>
<td>50.75%</td>
<td>19.18%</td>
<td>19.01%</td>
<td>−0.17%</td>
</tr>
<tr>
<td>2.00%</td>
<td>38.73%</td>
<td>14.06%</td>
<td>14.07%</td>
<td>0.01%</td>
</tr>
<tr>
<td>3.00%</td>
<td>32.30%</td>
<td>9.96%</td>
<td>10.18%</td>
<td>0.22%</td>
</tr>
<tr>
<td>4.00%</td>
<td>30.15%</td>
<td>7.21%</td>
<td>7.23%</td>
<td>0.01%</td>
</tr>
<tr>
<td>5.00%</td>
<td>28.50%</td>
<td>5.19%</td>
<td>5.08%</td>
<td>−0.11%</td>
</tr>
<tr>
<td>6.00%</td>
<td>26.50%</td>
<td>3.56%</td>
<td>3.59%</td>
<td>0.04%</td>
</tr>
<tr>
<td>7.00%</td>
<td>25.72%</td>
<td>2.56%</td>
<td>2.58%</td>
<td>0.02%</td>
</tr>
<tr>
<td>8.00%</td>
<td>25.50%</td>
<td>1.92%</td>
<td>1.95%</td>
<td>0.03%</td>
</tr>
<tr>
<td>9.00%</td>
<td>25.55%</td>
<td>1.50%</td>
<td>1.41%</td>
<td>−0.09%</td>
</tr>
<tr>
<td>10.00%</td>
<td>25.70%</td>
<td>1.20%</td>
<td>1.09%</td>
<td>−0.11%</td>
</tr>
</tbody>
</table>
Dependence of the Estimated Market Price of Risk on the Maturity of the Bonds Used in the Estimation for Different Time Periods. Reversion rate = 5%.

Figure 8.14: Historical Price of Risk.
Distribution of the Short-term Interest Rate in 1 Year
Estimated from Data on December 2, 2013

Distribution of the Short-term Interest Rate in 5 Years
Estimated from Data on December 2, 2013

Figure 8.15: Short Rate in 1 and 5 Years
Figure 8.16: Short Rate in 30 Years
Figure 8.17: Expectation and Standard Deviation of the Short Rate
Bibliography


Appendix A

P is the Derivative of the Receiver Swaptions Price with respect to K

\[
\frac{\partial}{\partial K} \int_{-\infty}^{\infty} (K - s)^+ p(s) ds = \frac{\partial}{\partial K} \int_{-\infty}^{K} (K - s)^+ p(s) ds \\
= \frac{\partial}{\partial K} \left( K \int_{-\infty}^{K} p(s) ds - \int_{-\infty}^{K} sp(s) ds \right) \\
\text{Product Rule} \quad (A.1)
\]

\[
\int_{-\infty}^{K} p(s) ds + kp(k) - kp(k) = \int_{-\infty}^{K} p(s) ds = P(K).
\]

\[\square\]
Appendix B

The Jumping Probabilities Lie Between 0 and 1

To show this, let us first make sure that $p_u$ and $p_d$ are non-negative.

If $m_1 - x_0 - j^* \Delta x \leq 0$, then automatically $p_d \geq 0$. Also,

$$p_u = \frac{\Delta t + (m_1 - x_0 - j^* \Delta x)^2}{2\Delta x^2} + \frac{m_1 - x_0 - j^* \Delta x}{2\Delta x} = \frac{1}{6} + \left(\frac{1}{2} - \frac{1}{2}\sqrt{2}\right) \frac{m_1 - x_0 - j^* \Delta x}{\Delta x},$$

since $\Delta x = \sqrt{3j}$ and $(\frac{m_1 - x_0 - j^* \Delta x}{2\Delta x})^2 = -\frac{1}{2}\sqrt{2} \frac{m_1 - x_0 - j^* \Delta x}{\Delta x}$ by the assumption of non-positivity of the numerator, and

$$\frac{1}{6} + \left(\frac{1}{2} - \frac{1}{2}\sqrt{2}\right) \frac{m_1 - x_0 - j^* \Delta x}{\Delta x} \geq \frac{1}{6},$$

since both $\frac{1}{2} - \frac{1}{2}\sqrt{2}$ and $m_1 - x_0 - j^* \Delta x$ are at most 0.

If $m_1 - x_0 - j^* \Delta x \geq 0$, then automatically $p_u \geq 0$. Also,

$$p_d = \frac{\Delta t + (m_1 - x_0 - j^* \Delta x)^2}{2\Delta x^2} - \frac{m_1 - x_0 - j^* \Delta x}{2\Delta x} = \frac{1}{6} + \left(\frac{1}{2} \sqrt{2} - \frac{1}{2}\right) \frac{m_1 - x_0 - j^* \Delta x}{\Delta x} \geq \frac{1}{6} + \left(\frac{1}{2} \sqrt{2} - \frac{1}{2}\right) \frac{m_1 - x_0 - \text{int}(\frac{m_1 - x_0}{\Delta x} + 0.5) \Delta x}{\Delta x} \geq \frac{1}{6} + \frac{1}{4} - \frac{1}{4} \sqrt{2} > 0.06 \geq 0.$$  

The non-negativity of these two probabilities ensures that $p_m \leq 1$. Finally, by showing that $p_u + p_d \leq 1$, we automatically get $p_u \leq 1$, $p_d \leq 1$ and $p_m \geq 0$, since
\[ p_m = 1 - p_u - p_d: \]

\[
p_u + p_d = \frac{\Delta t + (m_1 - x_0 - j^* \Delta x)^2}{\Delta x^2} = \frac{\frac{1}{6} + \left| \frac{m_1 - x_0 - j^* \Delta x}{\Delta x} \right|}{\Delta x}
\]

(B.4)

If the numerator is non-negative, this equals

\[
\frac{1}{6} + \left| \frac{m_1 - x_0 - \text{int}(\frac{m_1 - x_0}{\Delta x} + 0.5)\Delta x}{\Delta x} \right| \leq \frac{1}{6} \leq 1.
\]

(B.5)

And if the numerator is negative, it equals

\[
\frac{1}{6} - \frac{m_1 - x_0 - \text{int}(\frac{m_1 - x_0}{\Delta x} + 0.5)\Delta x}{\Delta x} \leq \frac{1}{6} + \frac{1}{2} = \frac{2}{3} \leq 1.
\]

(B.6)
Acknowledgments

I would like to thank Dr. Diederik Fokkema, my supervisor at EY. He has been an energetic supervisor, with both a great mind and effort to help me. At one point he asked me if he could do anything besides the things he was already doing, which was not only making time for one-on-one meetings but also being available for questions whenever he was present. He wanted to do more, although he was already being a great help. More than once did I ponder over something I did not understand for an hour, only to ask Dr. Fokkema and have an answer within 30 seconds. I developed great respect for him.

Furthermore, I would like to thank Dr. Svetlana Borovkova, my Primary Supervisor at VU University. Many professors from the Finance department were reluctant to take a student from the Exact Sciences department, but Dr. Borovkova consented and also had the mathematical background I was hoping for. Even from outside the Netherlands, she answered my electronic questions, with one of which she saved me an enormous amount of time. Not unwelcome, in view of the time I had already spent on the preliminaries. I was planning to study and summarize a series of articles, of which Dr. Borovkova luckily knew that they had already been summarized in Hull’s book. This is just one example, but worth mentioning.

I would also like to thank Dr. Andre Ran, my Second Reader at VU University. Although it was his right as Second Reader to wait for my semi-final thesis, read it and comment on it and leave it at that, he wanted to be part of the process from the beginning. This was especially helpful whenever Dr. Borovkova and Dr. Fokkema were abroad and I had need of face-to-face support. One occasion worth mentioning is when I wanted to ask a small question about something in a paper which did not require knowledge about the rest of the paper. He sat down with me and started from the beginning, simply because he was interested in what I was reading. After helping me with the part I had asked as well, he requested me to send him the article, so he could finish it himself as well. Having such a helpful and motivated second reader was very inspiring.

My thanks go to my colleagues at EY as well. They made me feel welcome, not only inviting me to lunch on day 1 but even showing up with over twenty of them at one of my performances as a comedian, and two months later coming to support me at the semi-finals of the Amsterdam Studenten Cabaret Festival, which I won partly because of them. Dr. Fokkema and his wife even showed up at the finals, even though they already knew 70% of my jokes. I also felt free in asking any of my colleagues questions at any time, and they were always able and willing to help me with any issues. If I end up working at EY, I will be lucky to have such a wonderful group of colleagues.
Finally, I would like to thank my best friend and former peer student Frank Blom MSc for providing me with the .TeX-file of his thesis, which I could use as a template for writing mine, and for always being ready to help me with IT-related issues. For example, the possibility for digital readers to click on citations and references to chapters, sections and formulas and immediately jumping to the relevant part was his idea.

Thanks, all. It has been a blast.