

Pricing and hedging options in a negative interest rate environment

Delft University of Technology
Faculty of Electrical Engineering, Mathematics and Computer Science
Delft Institute of Applied Mathematics
Mekelweg 4, 2628 CD Delft

ABN Amro Bank N.V.
CRM, Regulatory Risk, Model Validation, Product Analysis
Gustav Mahlerlaan 10, 1082 PP Amsterdam

A thesis submitted to the
Delft Institute of Applied Mathematics
in partial fulfillment of the requirements

for the degree

Master of Science
in
Applied Mathematics

by

Luuk Hendrik Frankena

February 29, 2016

MSc THESIS APPLIED MATHEMATICS

”Pricing and hedging options in a negative interest rate environment”

Luuk Hendrik Frankena

Delft University of Technology

Daily supervisors

S. Wandl MSc

S. van der Stelt PhD

Responsible professor

Prof. dr. ir. C.W. Oosterlee

Members of the thesis committee

Prof. dr. ir. C.W. Oosterlee

S. Wandl MSc

Dr. R.J. Fokkink

Dr. ir. M.B. van Gijzen

February 29, 2016

Delft

Acknowledgements

This thesis has been submitted for the degree Master of Science in Applied Mathematics at the Delft University of Technology. The academic supervisor of this thesis was Kees Oosterlee, professor at the Numerical Analysis group of the Delft Institute of Applied Mathematics. The main part of the research was conducted during an internship at the Model Validation team of ABN Amro Bank N.V. under the supervision of Sabrina Wandl and Sjors van der Stelt.

I would like to thank my supervisors Kees Oosterlee, Sabrina Wandl and Sjors van der Stelt for their advice and support throughout this thesis. I also would like to thank my colleagues of the Model Validation team and Anton van der Stoep for his advice about the time dependent SABR model.

Abstract

This thesis is about pricing interest rate options in a negative interest rate environment and about pricing foreign exchange barrier options. Conventional interest rate option pricing models are unable to price interest rate options in the current negative interest rate environment. Displaced versions and free boundary versions of the conventional models are proposed as a solution. Also normal models are proposed as a solution. Moreover, it is important to use risk metrics consistent with the model.

Foreign exchange barrier options are priced with local volatility, stochastic volatility and stochastic local volatility models. The valuation of a proprietary trading model is compared with industry standards such as the local volatility model and the constant parameter SABR model. Furthermore, it is compared with an extension of the SABR model with time dependent parameters. This time dependent SABR model can be calibrated to volatilities of multiple expiries, in contrast to the constant parameter SABR model. Finally, a local volatility component is added to guarantee a perfect calibration.

Keywords: negative interest rates, interest rate option pricing, Black's model, local volatility model, SABR model, CEV process, free boundary SABR, Bachelier's model, displaced models, time dependent SABR, effective parameters, stochastic local volatility, foreign exchange barrier option.

Contents

1	Introduction	2
1.1	Interest rates and interest rate derivatives	5
1.2	Mathematical framework and preliminaries	10
2	Positive rate modelling	12
2.1	Black’s model	12
2.2	Local volatility model	15
2.3	Constant Elasticity of Variance model	16
2.4	Stochastic Alpha Beta Rho model	18
2.5	Summary	24
3	Negative rate modelling	25
3.1	Bachelier’s model and the normal SABR model	25
3.2	Displaced models	27
3.3	Free boundary models	28
3.4	Summary	33
4	Pricing path dependent FX options	34
4.1	FX products and mathematical framework	35
4.2	Time dependent FX-SABR model	39
4.3	Mappings of time dependent to effective parameters	41
4.4	Calibration procedure	44
4.5	Improving calibration accuracy by adding a local volatility component	46
4.6	Normal version of the time dependent SABR model	46
5	Experiments	50
6	Conclusion	54
6.1	Further research	55
	References	56
	Appendix A Solution to Black’s model	58
	Appendix B Risk metrics of Black’s model	59
	Appendix C Risk metrics under Bachelier’s model	60
	Appendix D Transition probability density of the CEV process	61
	Appendix E Integrals in the volvol mapping for piecewise constant parameters	64
	Appendix F Recovery procedure of the characteristic function	66
	Appendix G Local and stochastic local volatility	70
0.1	Dupire’s formula	71
0.2	Dupire’s formula expressed in implied volatilities	72

1 Introduction

Economic Background

Negative interest rates is a very present-day topic in the current economy. The European Central Bank (ECB) and the central banks of Switzerland, Denmark, Sweden, and Japan have set negative interest rates on reserves¹, see also figure 1a. Switzerland is the first government in history to sell 10-year debt at a negative interest rate². A quarter of worldwide central bank reserves have a negative yield in February 2015 according to the Commonwealth Bank of Australia. In February 2016 the total size of government bonds with negative yields raises to a record high of almost six trillion US Dollar³, see figure 1b. In June 2014, the ECB was the first in decreasing their rate below zero⁴. The ECB aims to influence inflation by setting interest rates⁵.

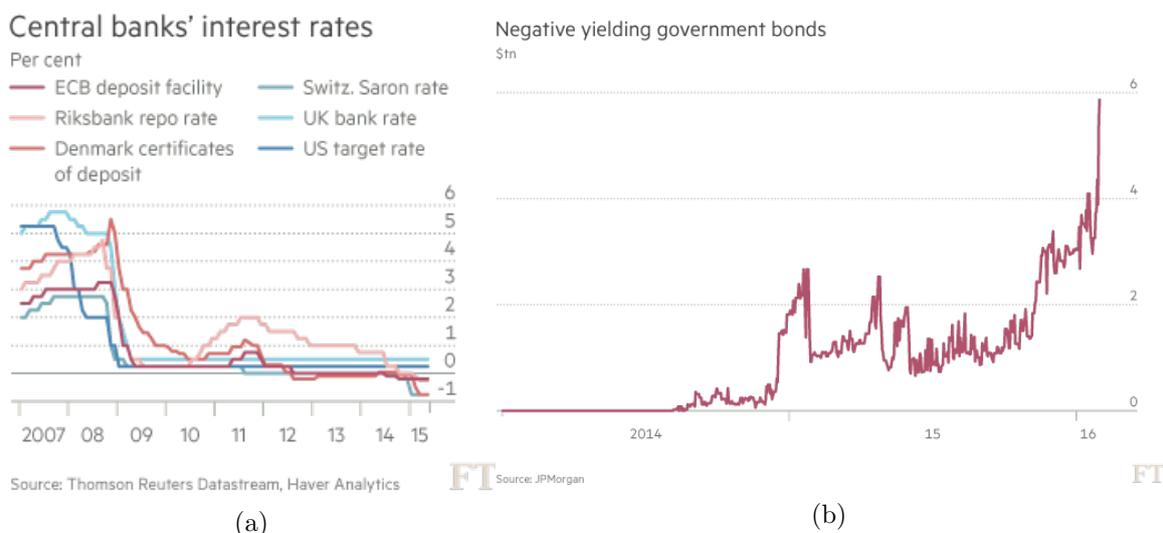


Figure 1: Figure (a) shows central bank's interest rates from 2007 until mid 2015. Figure (b) show the market value of outstanding government bonds with negative yield worldwide in February 2016, in trillion US Dollars. Source of both figures: Financial Times

Deposit (depo) rates are set low (or negative) by central banks to stimulate economic growth rates. There are several ways in which low interest rates improve economic growth. The monetary stimulus measure aims to enlarge credit to the real economy, increases asset prices, forces investors towards riskier instead of safe assets and lowers the exchange rate. Furthermore, the policy helps to lift inflation to the central bank's target, which is around 2% in the case of the ECB.

Now follows a short explanation of each of these arguments.

The main objective of a low interest rate policy is to deter saving and to encourage borrowing due to the lower costs of financing. This is the credit argument. Sub-zero depo rates are presented by policy makers as a tax imposed by the central bank on commercial banks to encourage them to increase lending to companies and consumers. A commercial bank can choose to pass on the tax to their customer by reducing their lending rates and by charging negative rates for deposits. This punishes depositors but increases bank lending. So the policy measure is also a tax for consumers with deposits in this case. When commercial banks choose to not pass the

¹ft.com/intl/cms/s/2/dba246d8-faf4-11e4-84f3-00144feab7de.html#axzz3ibRsjUjR

²ft.com/intl/cms/s/0/35ddc68e-dde7-11e4-8d14-00144feab7de.html

³next.ft.com/content/90ca12c0-d0b0-11e5-831d-09f7778e7377

⁴bloombergview.com/quicktake/negative-interest-rates

⁵ecb.europa.eu/home/html/faqinterestrates.en.html

tax to their costumers this does not result in an incentive for these banks to lend more to the real economy. Swiss wealth manager Julius Baer is one of the few examples that passes the cost of negative interest rates to its (institutional) clients ⁶.

Lower interest rates result in reduced discount factors on cash flows from assets. By definition, this increases the net present value of an asset. Furthermore, the monetary policy may raise the expectations of improved economic conditions and therewith higher future cash flows from assets.

Investors are encouraged to move from safe assets like government bonds to riskier assets. When an investor wants to maintain his target yield while the government bond yield is decreasing this implies a shift in balance of the portfolio to riskier assets. The increased risk taking has led to a convergence of sovereign spreads in the euro zone. Governments in the euro zone with a low credit quality can borrow cheaper due to this effect. This may lead to greater economic stability in the euro zone.

The exchange rate is depreciated indirectly since investors will change currency to invest in government bonds of countries that have a higher yield. A depreciated exchange rate boosts net export by making products and services cheaper for foreign companies and consumers. This causes growth, employment and an increased inflation due to higher import prices.

Economic consequences of this monetary policy are not in the scope of this thesis, more information can be found in (BIS, 2015)⁷.

Until recently it was assumed that interest could not go below the 'zero bound', since depositors could withdraw cash when rates go negative. However, cash needs to be stored and insured, which costs money. Furthermore, a bank account is more convenient in use. So apparently there is willingness to pay for having a bank account, which is equivalent to being charged negative interest rates. The question is how low the interest rates can go before cash becomes more attractive. Interchange fees, which credit card companies charge to costumers, can be seen as a proxy of how low rates can go⁸. Interchange fees are on average 2 to 3 per cent, so negative rates in this area should be feasible as well. If people would act completely rational, the rates cannot become more negative than the cost of carrying cash. Arbitrage is possible when the negative rate causes the coupon payments to be worth more than the cost of carrying the cash equivalent of the notional of the loan. The strategy is to borrow money at the negative rate, convert the money to cash, store it till maturity and then pay the face value (also called par value) in cash or convert it back first.

The major risks for financial institutions in a situation of negative rates are IT, operational and model risk. Their IT systems might not be able to deal with data containing negative rates. The assumption of positive rates can be hardwired in spreadsheets. Outcomes of entering negative rates in these spreadsheets are unknown. The provided data from Reuters or Bloomberg might be incorrect itself when those systems are not able to deal with negative rates. It can be compared with the millennium bug in the year 2000 when IT systems could not deal with dates past the twentieth century. Operational risk results from an incorrect implementation of solutions to negative rates. If traders plug ordinary implied volatilities in an adjusted model called the displaced model, this leads to incorrect prices. Furthermore, risk management may face issues when risk calculations such as Value at Risk (VaR) or the risk metrics of financial derivatives depend on a displaced volatility. This research will focus on model risk as a consequence of the negative rates.

⁶bloomberg.com/news/articles/2015-02-26/julius-baer-charges-institutional-clients-for-snb-negative-rate

⁷Hannoun, Ultra-low or negative interest rates: what they mean for financial stability and growth, BIS, <http://www.bis.org/speeches/sp150424.pdf>

⁸ft.com/intl/cms/s/2/dba246d8-faf4-11e4-84f3-00144feab7de.html#axzz3ibRsjUjR

Problem statement

Many mathematical models in finance (implicitly) assume that interest rates are positive, including the conventional models used in interest rate derivatives pricing. Models with this assumption fail to price (and hedge) financial products in the current economic situation of low or even negative interest rates. They break down or produce incorrect prices. According to Black's model [7] for instance, the price of a zero strike floor should equal zero, while current market values of these products are positive⁹. Note that interest rates do not have to be negative to cause the mathematical models to fail. It is sufficient that the market assigns a positive probability to the event of a negative interest rate in the future.

Interest rate derivatives in particular are sensitive to the changed interest rate environment because interest rates are their underlying asset. So they are especially exposed to interest rate risk. Examples of interest rate derivatives are interest caps/floors and swaptions. These products, together with swaps, are the most traded (simple, and in general) interest rate derivatives at ABN Amro. Swaps prices can be extracted from market quotes. It is more interesting to investigate interest rate derivatives with optionality, since their conventional valuation models assume that interest rates are positive, which leads to problems in the current economic situation. In the valuation of swaps this assumption is not made and their valuation in the current economic situation is not a problem.

One of the main (unstated) reasons for a central bank to decrease its interest rate is to depreciate its currency. A depreciated currency stimulates the export and therewith the domestic economy. Every policy rate change causes capital inflows and outflows. The current active interest rate policy of central banks throughout the world, leads to increased volatility in the foreign exchange market¹⁰. This in turn increases the demand in the financial industry for accurate pricing models for (complex) foreign exchange options.

This master thesis project compares model risk for pricing (and hedging) interest rate options in a negative interest rate environment in sections 2 and 3, and model risk for pricing (and hedging) foreign exchange barrier options in 4 and 5. Barrier options are a special type of options that will be explained later.

A valuation model is said to perform well if

- valuation is *arbitrage-free*: the model matches market quotes and interpolation is smooth.
- it produces risk metrics that are *stable* and *in line with market dynamics*.

Structure of the thesis

Section 1.1 introduces interest rates in general, together with interest rate derivatives. Interest rate derivatives are financial products whose value depend on the level of the underlying interest rate. Section 1.2 introduces the mathematical framework and notation. Chapters 2 and 3 provide an overview of models used in the pricing and hedging of simple interest rate options. The 2nd chapter provides an overview of conventional models used for interest rate derivatives pricing in an economy with positive rates. ABN Amro uses the same or similar models to price caps/floors [33]. The same models can be used for the pricing and hedging of simple foreign exchange options. The first model discussed is the Black's model, also known as diffusion or lognormal model. Secondly, the local volatility model, an extension of the Black's model, is discussed. Thirdly, the SABR model is explained in detail together with its popular approximation of implied volatility by Hagan's formula. Furthermore, the Constant Elasticity of Variance (CEV) model is analysed as an introduction to the SABR model. Chapter 3 contains modified versions of these models

⁹risk.net/risk-magazine/feature/2218691/negative-rates-dealers-struggle-to-price-0-floors

¹⁰Saxena. Capital flows, exchange rate regime and monetary policy. BIS. bis.org/publ/bppdf/bispap35c.pdf

that are used to price and hedge interest rate options in a negative rate environment. The models discussed are Bachelier's model, the normal SABR model, and displaced versions and free boundary versions of models from chapter 2. Chapter 4 introduces a generalization of the SABR model with time dependent parameters, and it introduces the foreign exchange market. This model can be used to price and hedge more complex options, so called path dependent options, such as a barrier option. The time dependent SABR model is introduced in a foreign exchange context. In chapter 5 the time dependent SABR model, the constant parameter SABR model, the local volatility model and ABN Amro's proprietary trading model are calibrated to foreign exchange data in order to price foreign exchange barrier options. An analysis of the particular foreign exchange markets and a comparison of the valuations produced by these models is contained in the same section. Section 6 contains the conclusion of the results of this research project.

1.1 Interest rates and interest rate derivatives

The Bank for International Settlements (BIS) collects data about the amounts of outstanding of over-the-counter (OTC) derivatives. In June 2014 interest rate contracts represented 77% gross market value in these OTC derivatives markets. In the same period the largest portion of these OTC single-currency interest rate derivatives - around 88% in gross market value - were swaps [19]. The market for swaps is around ten times larger than the market for options on interest rates, which is around ten times larger than the market for forward rate contracts. These numbers show that interest rate derivatives play an important role in the world economy.

Interest is the amount of money the borrower (debtor) promises to pay the lender (creditor) for the use of its money. An interest rate is the rate at which the interest is paid. For example; if the principal is 150 euro and the interest for borrowing the money one year is 7,50 euro, the yearly interest rate is $7,50/150=5\%$.

Interest can be periodically compounded or continuously compounded. Assume that there are m periods (a year) of payments with (yearly) rate r , this implies that putting the amount A in the bank account results in an amount of $A(1 + \frac{r}{m})^{mT}$ after T years. Continuously compounding means compounding over an infinitesimal small period. This corresponds to taking the limit $m \rightarrow \infty$ in the last expression, yielding Ae^{rT} after T years.

The interest rate depends on the credit risk; the risk of a default by the borrower of the funds [24]. When such a default occurs, the interest and principal are not paid to the lender as is agreed in the contract. The higher the credit risk, the higher the interest rate that the lender asks for.

Important interest rates for the European financial markets are the LIBOR and EURIBOR.

LIBOR

LIBOR stands for London Interbank Offer Rate and it is the rate of interest that a selection of major banks charge each other for short-term loans. It is an indication of the average rate at which contributor banks can borrow money in the London interbank market for a particular period and currency. Examples of banks contributing to LIBOR are HSBC, UBS, Société Générale, Rabobank, Bank of Tokyo-Mitsubishi and JP Morgan Chase. LIBOR is administered by the InterContinentalExchange (ICE) Benchmark Administration (IBA) and comes in seven different maturities; overnight, one week and one, two, three, six and twelve months¹¹. LIBOR rates are fixed daily for five different currencies: the Euro (EUR), US Dollar (USD), Pound Sterling (GBP), Japanese Yen (JPY) and Swiss Franc (CHF). This means that there is a total of thirtyfive different LIBOR rates fixed by ICE each business day.

¹¹theice.com/iba/libor

EURIBOR

The European Money Markets Institute (EMMI) publishes the EURIBOR, which is similar to LIBOR. It stands for Euro Interbank Offered Rate and it is the rate at which Euro interbank term deposits are being offered by one highly creditworthy bank to another within the European Monetary Union¹². EURIBOR is different from the EURO LIBOR and is more commonly used.

The EUR LIBOR rates that ICE¹³ published on 13 April 2015 can be found in table 1.

Zero rates

An n -year zero coupon interest rate is defined as the rate earned on an investment that starts now and where all the interest and principal is realized at the end of n years. This means that zero coupon instruments do not have intermediate payments. From a borrower's perspective for example, the 10-year zero rate with continuous compounding, market quote 5% per annum and principal 100 has a cashflow of -100 at time 0 and a cashflow of $100 \cdot e^{0.05 \cdot 10} = +164.87$ after ten years and zero cashflows in between.

Zero rates are important in the construction of the index and discount curve, which will we discuss next and which are important in the valuation of interest rate derivatives and foreign exchange derivatives.

Index and discount curves

Forward interest rates are rates for periods of time in the future implied by current zero rates [24]. The following example illustrates this idea. Suppose that the per annum rate for one year is 4.0% and per annum rate for two years is 5.0%, both with continuous compounding. The forward rate for the second year is the rate that is implied by the zero rates for the period in time between the end of the first year and the end of the second year. Suppose that the notional is 100, since $100 \cdot e^{2 \cdot 0.05} = 100 \cdot e^{0.04} \cdot e^f$, this means that $f = 6\%$ is the forward rate from the first to the second year. It is the rate for year two that, when starting with the 4.0% rate of year one, produces the overall rate of 5.0% for two years.

The graph of forward rates is called the index curve. It can be constructed for small maturities from deposit rates [13]. For longer maturities it can be obtained from (LIBOR) swap rates. The discount curve is built from quoted overnight indexed swap (OIS) rates, these rates will be explained later. The index curve as well as the discount curve are assumed to be given.

Forward Rate Agreement

A *forward contract* is an agreement to buy or sell a certain underlying at a certain time in the future at a certain price [24]. Forward contracts are usually traded in the OTC market. In contrast to an exchange market, contracts in an OTC market are not standardized and trades between two parties can be executed without others knowing. In general, an exchange market is more transparent and more liquid. A *forward rate agreement* (FRA) is a forward contract on an interest rate. In a typical FRA, two parties exchange a fixed rate for a floating rate with each other at a certain time in the future. The agreement is settled at the termination date, usually in cash. Only the difference of the payments needs to be transferred. Imagine a two-year FRA with a principal of 100.000 USD on the one-year LIBOR and a one-year fixed rate of 3.0%. The LIBOR turns out to be 3.5% after two years. The result is a payment of $(3.5\% - 3.0\%) \cdot 100.000 = 500$ USD from the floating rate payer to the fixed rate payer in two years from now.

¹²emmi-benchmarks.eu/euribor-org/about-euribor.html

¹³theice.com/iba/historical-data

Swaps and swap rates

Interest rate swaps (IRS) allow financial managers to effectively hedge their interest rate exposure. An IRS can be used to transform a floating-rate loan into a fixed-rate loan, or the other way around. It is a financial product in which two parties exchange interest rate cash flows during a fixed period of time. One of the cash flows has a fixed rate and the other has a floating rate indexed to a reference rate such as the LIBOR or the EURIBOR. The cash flow with the fixed rate is called the *fixed leg* and the cash flow based on the floating rate is called the *floating leg*. Identification of the payer and the receiver of the swap is based on the fixed leg. The one that pays the fixed rate and receives the floating rate is called the *payer swap*. The *receiver swap* pays the floating rate and receives the fixed rate. An IRS is usually structured so that one side transfers the difference between the two payments to the other side.

The *tenor* is the period of payment of a cash flow (or the inverse of its payment frequency). For a swap, the tenor of the floating leg does not have to match the tenor on the fixed leg. The standard IRS in the US for instance, is a swap with quarterly LIBOR payments and semiannual fixed payments [24].

The notional principal of both legs is equal. Note that in most swaps the principal is used only for the calculation of interest payments; the principal itself is not exchanged. That is why it is called the *notional principal*, or simply *notional*. Exchanging the principal at the end of the life of the contract would add no financial value to either the swap payer or the swap receiver, since the value of the principal is equal for both sides at that time. This is different for *cross-currency swaps*, swaps with a different currency for each leg. Despite having equal valued notionals at the beginning of the contract, such swaps can have different principal values at the end of the contract due to changes in the exchange rate. The reason for this is that the notional of each leg is expressed in its own currency.

A swap can be characterised as a portfolio of forward rate agreements, since an FRA is equivalent to a single-period interest rate swap. Another characterisation of a swap is a portfolio of two bonds, one bond paying fixed rates and one bond paying floating rates.

Swap rates

Swaps are usually not arranged as agreements between two (nonfinancial) companies directly, since it is unlikely that they need to hedge their interest exposure at exactly the same time and for exactly the same notional principal. This means that nonfinancial companies make use of financial intermediaries, so called market makers. A market maker in swaps quotes two rates: its bid rate and its offer rate. The *bid rate* is the fixed rate that it is prepared to pay in exchange for receiving LIBOR and its *offer rate* is the fixed rate that is prepared to receive in return for paying LIBOR. There is usually a small difference between these rates, called *spread*, which is the main source of profit for the market maker. The *swap rate* is defined to be the average of the bid rate and the offer rate.

LIBOR swap rates can be obtained from ICE, just like LIBOR. The ICE Swap Rate is the most important benchmark for swap rates and spreads for interest rate swaps globally¹⁴. The rate is published for tenors ranging from one year to thirty years and for the currencies EUR, GBP and USD. As an example, the EUR swap rates that ICE¹⁵ published on 13 April 2015 can be found in table 2.

Direct observation of LIBOR rates is possible only upto twelve months. Traders use LIBOR swap rates to extend a LIBOR zero curve.

¹⁴theice.com/iba/ice-swap-rate

¹⁵theice.com/iba/historical-data

Table 2: EUR ICE swap rates 1100 13-APR-2015

Tenor	Swap Rate
1 Year	0.001
2 Years	0.072
3 Years	0.104
4 Years	0.157
5 Years	0.217
6 Years	0.284
7 Years	0.355
8 Years	0.422
9 Years	0.482
10 Years	0.535
12 Years	0.625
15 Years	0.720
20 Years	0.799
25 Years	0.825
30 Years	0.844

Table 1: EUR ICE LIBOR 13-APR-2015

Tenor	LIBOR Rate
Overnight	-0.16643
1 Week	-0.10500
1 Month	-0.03643
2 Months	-0.00786
3 Months	0.00714
6 Months	0.06571
1 Year	0.18857

OIS rates

The *risk-free rate* is the theoretical rate of return of an investment with zero risk. The rate plays an important role in the mathematical theory of asset pricing. In the past interest rate derivatives traders used LIBOR rates as proxies for the risk-free rate when valuing derivatives. They used a single curve constructed from the LIBOR rates for both calculating future rates as well as discounting. Since the credit crunch it is market practice to discount with overnight index swap (OIS) rates, LIBOR turned out to be a poor proxy for the risk free rate under stressed market conditions. An OIS is an interest rate swap in which a fixed rate of interest is exchanged for a floating rate of interest that is the geometric mean of a specific daily overnight rate[25]. The overnight rates for the EUR, USD and GBP market are the Euro Overnight Index Average (EONIA)¹⁶, the effective Federal Funds Rate¹⁷ and the Sterling Overnight Index Average (SONIA)¹⁸ respectively. The OIS market quotes are the fixed rates for an OIS[13].

A *yield spread* is defined to be the difference between the quotes of two different interest rates. The three month LIBOR-OIS spread for instance, is the difference between the three month LIBOR and the three month OIS rate. This spread used to be very small, so it made sense to price derivatives in a single curve framework where discounting and calculation of future rates were based on a single reference rate. The divergence of the LIBOR and OIS rates during the credit crunch caused the market to adopt a multicurve framework where different rates are used for discounting and calculation of future values. The OIS rate is presently the best proxy for the risk-free rate for the valuation of collateralized interest rate derivatives according to Hull and White [25]. Standard agreements stipulate daily collateral calls. Given these daily collateral calls, the choice for an overnight swap rate as the risk-free rate is natural.

Day count convention

Interest is earned over some reference period, and the day count convention is the way of measuring the length of this period. The day count fraction α_i represents the ratio of length of time interval $[T_{i-1}, T_i]$ over the length of a 'year'. This is a simplification, since there are different ways to calculate the number of days, one could only count business days for instance.

¹⁶emmi-benchmarks.eu/euribor-eonia-org/about-eonia.html

¹⁷newyorkfed.org/markets/omo/dmm/fedfundsdata.cfm

¹⁸wmba.org.uk/pages/index.cfm?page_id=31

Calls, puts, caps and floors

The buyer of a European-type *call (put) option* has the right, but not the obligation, to buy (sell) the underlying asset at a certain time, the *expiry date* T , for a certain amount, the *strike* K . The seller is obligated to sell the underlying asset to the buyer, if the buyer *exercises* its option at the expiry date. The payoffs of a call and a put option at the expiry date are:

$$P^{call} = \max(F(T) - K, 0) =: (F(T) - K)^+, \text{ and} \quad (1a)$$

$$P^{put} = \max(K - F(T), 0), \quad (1b)$$

respectively, where $F(T)$ is the level of the underlying asset at maturity. Asset $F(T)$ could represent the foreign exchange forward rate between the Euro and the US Dollar at time T for instance. Mathematical models are used to determine the value of such a put or call option. The classic models used in industry for the pricing of interest rate and foreign exchange options, are introduced in section 2, after introducing the mathematical framework in 1.2. European-type call and put options are more complex in the interest rate market. An interest rate cap is an insurance for a holder of a loan with a floating rate against the floating rate rising above a pre-defined level, the cap-rate K (like a strike). A cap (floor) consists of a series of N subsequent caplets (floorlets). A caplet (floorlet) is a call (put) option on an interest rate for a specific period. It is market practice to price caplets with Black's model - also known as diffusion model [27, 7], which will be introduced in the section 2.1. The payoff function of a caplet (floorlet) in period $[T_{i-1}, T_i]$ is

$$P_i^{caplet} = N_i \cdot \alpha_i \cdot \max(R(T_{i-1}) - K, 0), \quad (1c)$$

$$P_i^{floorlet} = N_i \cdot \alpha_i \cdot \max(K - R(T_{i-1}), 0), \quad (1d)$$

where N_i is the notional and α_i is the day count fraction (or coverage). The rate R is set at T_{i-1} ; the beginning of the period i and the payment is usually made at T_i ; the end of the period i . This means that we should discount from T_i . The payoff of a cap (floor) is simply the sum of the payoffs of its caplets (floorlets):

$$P^{cap} = \sum_{i=1}^N P_i^{caplet}, \quad (1e)$$

$$P^{floor} = \sum_{i=1}^N P_i^{floorlet}. \quad (1f)$$

Barrier options

Barrier options are *path-dependent* options, that means that the payoff of a barrier option depends on historical price levels of the underlying asset [35]. Certain properties of the contract are triggered when the price of the underlying asset becomes too high or too low. Next to a call or a put feature, barrier options have two aspects: the current forward price can be below or above the barrier level and the option can be a so called in or out barrier option. An in option starts worthless and gains value when the barrier is reached. An out option on the contrary, starts with a value and becomes worthless if the barrier level is reached, it is knocked out in other words. In summary, there are four types of barrier options:

1. Up-and-out: the current forward price is below the barrier level and has to move up to reach the barrier level for the option to be knocked out.
2. Down-and-out: the current forward price is above the barrier level and has to move down to reach the barrier level for the option to be knocked out.
3. Up-and-in: the current forward price is below the barrier level and has to move up to reach the barrier level for the option to be knocked in.

4. Down-and-in: the current forward price is above the barrier level and has to move down to reach the barrier level for the option to be knocked in.

Up-and-out and down-and-out barrier options in the foreign exchange market are priced in the experiments chapter 5. Foreign exchange barrier options are explained in more detail in section 4.1.

1.2 Mathematical framework and preliminaries

This subsection will state definitions and theorems from stochastic calculus and financial mathematics. Furthermore, notation will be introduced. Both will be used throughout this thesis. More details can be found in Andersen and Piterbarg [1].

The **normal distribution**, also called Gaussian distribution, is used many times throughout this thesis. The standard normal cumulative distribution function is denoted by $\Phi(\cdot)$ and its derivative, the standard normal probability density function $\Phi'(x) = e^{-x^2/2}/\sqrt{2\pi}$ is denoted by $\phi(\cdot)$.

Wiener process. A stochastic process $\{W_t\}_{t \geq 0}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where \mathbb{P} is the real-world probability measure, is a Wiener process, also known as a Brownian motion, when it satisfies:

1. $W_0 = 0$,
2. Mapping $t \rightarrow W_t$ is continuous everywhere almost surely,
3. Independent increments, where the increments $W_t - W_s$ for $0 \leq s \leq t$ are normally distributed with mean 0 and variance $t - s$.

Filtration. The filtration considered in this thesis is always generated by the relevant Brownian motion $\{W_t\}_{t \geq 0}$ in its context: $\mathcal{F}_t = \sigma\{W(s) \mid 0 \leq s \leq t\}$.

Stochastic differential equation (SDE). An SDE with drift μ and volatility σ is defined as

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t, \quad (2)$$

where $\mu, \sigma : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$.

Itô process. The solution of SDE (2) is the Itô process given by

$$X_t = X_0 + \int_0^t \mu(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s. \quad (3)$$

Itô's lemma. Let $f(t, x)$ denote a continuous function $f : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$, with continuous partial derivatives $\partial f / \partial t \equiv f_t$, $\partial f / \partial x \equiv f_x$, $\partial^2 f / \partial x^2 \equiv f_{xx}$. Let X_t be given by the Itô process (3) and define $Y_t \equiv f(t, X_t)$. Then Y_t is an Itô process with stochastic differential equation (SDE)

$$dY_t = \left(f_t(t, X_t) + f_x(t, X_t)\mu(t, X_t) + \frac{1}{2}f_{xx}(t, X_t)\sigma^2(t, X_t) \right) dt + f_x(t, X_t)\sigma(t, X_t)dW_t. \quad (4)$$

Itô's lemma can be motivated heuristically from a Taylor expansion:

$$f(t + dt, X_{t+dt}) = f(t, X_t) + f_t dt + f_x dX_t + \frac{1}{2}f_{xx}(dX_t)^2 + \dots, \quad (5)$$

where $(dX_t)^2 = \sigma^2(t, \omega)dt$ in the limit.

Fokker-Planck partial differential equation (PDE). The transition probability density function $p(t, x) = p(t_0, x_0, t, x)$ associated to SDE (2) for X_t for $t_0 \leq t \leq T$, satisfies the Fokker-Planck PDE:

$$\frac{\partial}{\partial t} p(t, x) = -\frac{\partial}{\partial x} [\mu(t, x)p(t, x)] + \frac{1}{2} \frac{\partial^2}{\partial x^2} [\sigma^2(t, x)p(t, x)], p(t_0, X_0) = \delta(X_0), \quad (6)$$

where δ is the Dirac delta function and W_t a Wiener process.

Martingale. A stochastic process $\{X_t \mid 0 \leq t \leq T\}$ is a martingale with respect to filtration \mathcal{F} , if X is a \mathcal{F} -adapted process such that X_t is integrable for all $t \in [0, T]$ such that

$$\mathbb{E}[X_t \mid \mathcal{F}_s] = X_s \text{ almost surely, for } 0 \leq s \leq t \leq T. \quad (7)$$

Money market account $B(\cdot)$. The continuously compounded money market account $B(t)$ satisfies

$$B(t) \equiv \exp\left(\int_0^t r(u)du\right). \quad (8)$$

Risk-neutral measure \mathbb{Q} . The risk neutral measure \mathbb{Q} has the money market account as a numeraire. Under \mathbb{Q} and in the absence of arbitrage a contingent claim $V(t)$ is valued as

$$V(t) = B(t)\mathbb{E}_{\mathbb{Q}}[V(T)/B(T) \mid \mathcal{F}_t]. \quad (9)$$

In the context of foreign exchange options in section 4.1 this measure is called the spot measure \mathbb{S} .

Zero coupon bond price $P(\cdot, \cdot)$. In the absence of arbitrage, the time $0 \leq t \leq T$ price $P(t, T)$ of a T -maturity zero coupon bond is given by

$$P(t, T) \equiv \mathbb{E}_{\mathbb{Q}}[B(t)/B(T) \mid \mathcal{F}_t] = \mathbb{E}_{\mathbb{Q}}\left[\exp\left(-\int_t^T r(u)du\right) \mid \mathcal{F}_t\right]. \quad (10)$$

The zero coupon bond $P(t, T)$ is used as a discount factor for certain products.

T -forward measure \mathbb{F}^T . [30] The T -forward measure \mathbb{F}^T uses the T -maturity zero-coupon bond (10) as the numeraire asset. Under \mathbb{F}^T a contingent claim $V(T)$ is valued as

$$V(t) = P(t, T)\mathbb{E}_{\mathbb{F}^T}[V(T) \mid \mathcal{F}_t]. \quad (11)$$

Arbitrage. When pricing financial derivatives, it is crucial to use a method that does not introduce arbitrage. Arbitrage is defined to be a costless trading strategy which at some future time provides a positive profit with a positive probability, but has no possibility of a loss.

Risk neutral probability density function and arbitrage testing. According to Breeden and Litzenberger [8] the risk neutral probability density function can be approximated by the second order derivative of the call value with respect to the strike:

$$P(0, T_j)f_{\mathbb{F}^T_j}(K_i) = \frac{\partial^2 C(K, T_j)}{\partial K^2} \Big|_{K=K_i}. \quad (12)$$

In order to be arbitrage-free, the probability density function should be nonnegative and integrate to one.

2 Positive rate modelling

This chapter introduces conventional models used in the pricing and hedging of interest rate options and foreign exchange options. The modeled forward rate could be an interest rate forward as well as an exchange rate forward. In section 2.1 Black's model is introduced, a variation of the famous Black-Scholes model. Black's model is extended with local volatility in section 2.2, to be able to calibrate to market data and to price path dependent options. The CEV model is introduced in 2.3, as an introduction to the SABR model in 2.4. At the end of the chapter, the models are summarized in section 2.5.

2.1 Black's model

In 1976 Fischer Black introduced a model [7] that is a special case of the original Black-Scholes model:

$$dF_t = \sigma F_t dW_t, \quad (13)$$

where σ is the constant volatility and W_t a Brownian motion. The difference is that it uses the forward rate F_t instead of the spot rate S_t . This makes Black's model useful for pricing interest caps / floors, swaptions, and foreign exchange options. The relationship between the spot rate S_t in Black Scholes' model and the forward rate F_t in Black's model is $S_t = P(0, T)F_t$, where $P(0, T)$ is the discount factor at time to maturity T . The solution of SDE (13) is

$$F_t = F_0 \exp(\sigma W_t - \sigma^2 t/2), \quad (14)$$

so F_t is lognormally distributed with $F_t = \exp(Y_t)$ where $Y_t \equiv Y_0 + \sigma W_t - \sigma^2 t/2$ and therewith $Y_t \sim \mathcal{N}(Y_0 - \sigma^2 t/2, \sigma^2 t)$. See figure 2 for a plot of the probability distribution. The details of the derivation of Black's model are contained in appendix A.

The prices of a European call and put option on a forward rate with strike K and time to maturity T are given by

$$B(K, F_0, \sigma, T) \equiv V^c(0) = P(0, T)[F_0 \Phi(d_1) - K \Phi(d_2)], \text{ and} \quad (15a)$$

$$V^p(0) = P(0, T)[K \Phi(-d_2) - F_0 \Phi(-d_1)], \quad (15b)$$

respectively, where

$$d_1 = \frac{\log(F_0/K) + (\sigma^2/2)T}{\sigma\sqrt{T}} \text{ and } d_2 = d_1 - \sigma\sqrt{T}. \quad (15c)$$

This model assumes the forward rate $F(t)$ process to be log-normal. The *put-call parity*

$$V^c(t) - V^p(t) = P(t, T)(F_t - K), \quad (16)$$

is an important relation between the price of a call and a put option, that holds for any pricing model. It is often used in the derivation of the risk metrics of a pricing model.

Risk metrics for Black's model

For each option pricing model, certain risk metrics can be calculated. The sensitivity of a call option price with respect to the underlying rate is called the call delta Δ_c . The underlying rate can be a spot rate or a forward rate, this depends on the market conventions of the specific product. The delta of a put option Δ_p can be computed from the delta of a call Δ_c by the put-call parity (16). Other important risk metrics are vega Λ and gamma Γ : the sensitivity of the option price to its volatility and the second order sensitivity of the option price to its underlying forward rate respectively. The derivation of vega, gamma and the call delta can be

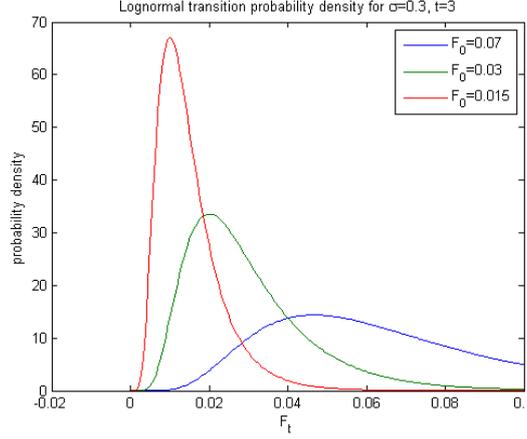


Figure 2: Probability density of Black's model for current forward rates $F_0 \in \{0.015, 0.04, 0.07\}$, volatility $\sigma_B = 0.3$ and time $t = 3$.

found in appendix B. The results are as follows:

$$\Delta_c \equiv \frac{\partial V^c}{\partial F_0} = P(0, T)\Phi(d_1), \quad (17a)$$

$$\Delta_p \equiv \frac{\partial V^p}{\partial F_0} = P(0, T) [\Phi(d_1) - 1], \quad (17b)$$

$$\Lambda \equiv \frac{\partial V^{c,p}}{\partial \sigma} = P(0, T)F_0\sqrt{T}\phi(d_1), \quad (17c)$$

$$\Gamma \equiv \frac{\partial^2 V^{c,p}}{\partial F_0^2} = P(0, T)\phi(d_1)\frac{1}{\sigma\sqrt{T}F_0}. \quad (17d)$$

The metrics will be referred to as Δ_c^B , Δ_p^B and Λ^B , where the B stands for Black's model, since the same metrics will be derived for other models later.

Black's model and zero strike puts

Note that $\log(F_0/K) \rightarrow \infty$ as $K \downarrow 0$. As a result $d_1 \rightarrow \infty$, $\Phi(-d_1) \rightarrow 0$ and thereby $V_p \rightarrow 0$ as $K \downarrow 0$, since the factor $K\Phi(-d_2)$ vanishes. This implies that the Black model assumes that the underlying has a zero probability of becoming negative. In contrast, markets in 2015 and 2016 show us that certain puts of strike zero are traded for nonzero prices. So the market assigns a positive probability to negative rates in the future. Note that this is independent of the sign of the current rate. Also note that this means that the value of a zero strike call option is equal to the value of its underlying, by the put-call parity (16).

The zero strike puts cannot be priced (correctly) with Black's model in this negative interest rate environment. Furthermore, this leads to difficulties in pricing puts with small (but positive) strikes because of the continuity of V^p in K . This is shown with an example next. First, an important concept is introduced that will be used many times throughout this thesis.

The *implied (Black) volatility* is the volatility that when substituted in Black's model (15) results in the market value of the option. So, the implied volatility is the σ that solves the nonlinear equation

$$f(\sigma) = P(0, T)[K\Phi(-d_2) - F\Phi(-d_1)] - V_{market}^p = 0. \quad (18)$$

When Black's model is calibrated to market data of call and put options of the same expiry

and different strikes, it appears that the implied volatility varies per strike. In other words, the assumption of constant volatility does not hold in practice. The implied volatility as a function of the strike is called the *implied volatility smile* or the *implied volatility skew*. The volatility smile is different for each expiry as well; the implied volatility as a function of both strike and expiry is called the *implied volatility surface*. After introducing these concepts, we proceed to the example of puts with a small strike.

If V_p is not 'small' for small K , a large implied Black volatility is necessary to 'compensate' for this. Letting the strike go to zero while fixing a positive value of the put V^p results in blowing up of the volatility, as can be seen in figure 3. In other words, the implied volatility of a put option with a small strike and a (large enough) positive value does not exist.

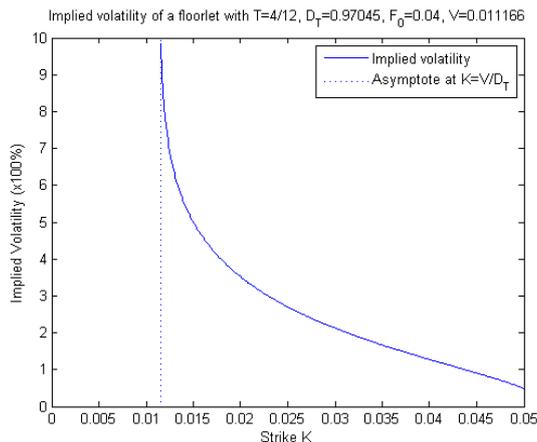


Figure 3: As the strike of the European put option (floorlet) with positive market value approaches zero, the Black implied volatility 'blows up'. Parameters used are $F_0 = 0.04$, $V^p = 0.011166$, $T = 4/12$ and $P(0, T) = 0.9704$.

In order to see that the implied volatility has an asymptote at $K = V^p/P(0, T)$, we rewrite equation (15b) to

$$\frac{V^p}{P(0, T)} = K\Phi(-d_2) - F_0\Phi(-d_1). \quad (19)$$

Let F_0 , V^p , T and $P(0, T)$ be fixed and let $K \downarrow V^p/P(0, T)$. Then it must hold that $\Phi(-d_1) \downarrow 0$ and $\Phi(-d_2) \uparrow 1$ in order to keep the left and right hand side of equation (19) in balance. This holds if and only if $\sigma \rightarrow \infty$.

The implied volatility of a put option with a strike smaller than $V^p/P(0, T)$ cannot be obtained, since formula (19) and $K < V^p/P(0, T)$ imply that

$$\frac{V^p}{P(0, T)} < \frac{V^p}{P(0, T)}\Phi(-d_2) - \underbrace{F_0}_{\geq 0} \underbrace{\Phi(-d_1)}_{0 \leq \dots \leq 1}, \quad (20)$$

and thereby that $\Phi(-d_2) > 1$. This is not correct by the definition of a cumulative distribution function. It means that puts with strikes smaller than $V^p/P(0, T)$ cannot be quoted by their Black implied volatilities. The same analysis can be applied to an option's risk metrics.

Traders and risk managers determine the value of an illiquid option by using information of implied volatilities of liquid options. The previous example illustrates that the price of small strike options is very much dependent on the volatility used. So, Black's model has large model uncertainty for small strikes.

2.2 Local volatility model

In section 2.1 it was described that the implied volatility is not constant in practice. The implied volatility needed to match market quotes usually varies with both the strike K and the expiry T . The usage of a different volatility for each expiry-strike pair actually implies that a different model is used for each combination. This leads to problems with the pricing of exotic options such as barrier options, whose payoff depend on the level of the forward rate at different points in time. To price a barrier option, the implied volatility of both the barrier level B and the strike K should be input of the model at the same time. This is not possible with Black's model, since those volatilities are different.

Dupire, Derman and Kani [17, 16] found a solution for this problem by introducing local volatility. They extended Black's model by replacing the constant volatility σ by the so called *local volatility* function $\sigma_{LV}(t, F_t)$ that is dependent on time t and the underlying forward rate F_t . The stochastic differential equation that describes the dynamics of the forward rate under the local volatility model in the T -forward measure \mathbb{F}^T is given by

$$dF_t = \sigma_{LV}(t, F_t)F_t dW_t. \quad (21)$$

Dupire's [17] local volatility formula, for a forward rate, is given by

$$\sigma_{LV}(T, K) = \sqrt{\frac{2 \frac{\partial C(T, K)}{\partial T}}{K^2 \frac{\partial^2 C(T, K)}{\partial K^2}}}. \quad (22)$$

where $C(T, K)$ is the value of a call option with strike K and expiry T . The derivation of Dupire's formula for a spot rate is described in appendix G0.1. In the spot measure \mathbb{S} , the local volatility is a function of the time and the underlying spot rate. A more practical expression of Dupire's formula in terms of implied volatilities instead of call prices, is described in appendix G0.2. The local volatility model is applied to market data of foreign exchange options in chapter 5.

A major problem with the local volatility model is that it predicts the wrong dynamics of the volatility smile. When the price of the underlying increases, one expects that the smile shifts to higher levels as well. In contrast, the local volatility model predicts that the smile will shift to lower prices after an increase of the underlying. The opposite counterintuitive movement can be seen for a decrease of the underlying. Due to this contradiction, delta and vega risk metrics under the local volatility model may perform worse than the risk metrics of Black's model [20]. To show this problem, ignore the time parameter and consider the special case where

$$dF_t = \sigma_{LV}(F_t)F_t dW_t. \quad (23)$$

In [22, 20] it is shown with perturbation methods that the implied Black volatility of a European call or put option can be expressed in terms of the local volatility by the following approximation:

$$\sigma_B(K, F_0) = \sigma_{LV}([F_0 + K]/2) \left[1 + \frac{1}{24} \frac{\sigma_{LV}''([F_0 + K]/2)}{\sigma_{LV}([F_0 + K]/2)} (F_0 - K)^2 + \dots \right]. \quad (24)$$

Usually $|K - F_0| \ll 1$ holds, since K and F represent rates, like for example 0.04 and 0.12. This implies that the first term of the right-hand-side dominates the behaviour of the approximation. The second term is a small correction to the first term and the total contribution of the other terms is very small. So for calibration purposes the following approximation is useful:

$$\sigma_B(K, F_0) = \sigma_{LV}([F_0 + K]/2). \quad (25)$$

Suppose that the current forward price is F_0 and that the volatility smile observed in the market is $\sigma_B^M(K, F_0)$. According to the above approximation it must hold that

$$\sigma_B^M(2F - F_0, F_0) = \sigma_{LV}\left(\frac{F_0 + (2F - F_0)}{2}\right) = \sigma_{LV}(F). \quad (26)$$

When the initial forward price F_0 shifts to a forward price F , the relation between the initial market smile σ_B^M and the new market smile $\sigma^{M,new}$ is

$$\sigma_B^{M,new}(K, F) \stackrel{(25)}{=} \sigma_{LV} \left(\frac{F+K}{2} \right) \stackrel{(26)}{=} \sigma_B^M \left(2 \left[\frac{F+K}{2} \right] - F_0, F_0 \right) = \sigma_B^M(K + F - F_0, F_0), \quad (27)$$

or equivalently,

$$\sigma_B^{M,new}(K, F_0 + \Delta F) = \sigma_B^M(K + \Delta F, F_0), \quad (28)$$

where $\Delta F := F - F_0$. Suppose σ_M is a convex function, as is usually the case in practice. This means that the volatility smile shifts to the left when $F > F_0$ and to the right when $F < F_0$. Exactly the opposite of what is expected from intuition and from market practice. This is best illustrated with an example. Suppose $\sigma_B^M(K, F_0) = (F_0 - K)^2$, then $\sigma_B^{M,new}(K, F) = \sigma_B^M(K + \Delta F, F_0) = (F_0 - (K + \Delta F))^2 = ((F_0 - \Delta F) - K)^2$. For $\Delta F > 0$ the smile was shifted to the left instead of the right.

This property of the local volatility model leads to incorrect delta hedges. In the local volatility model, European call options are priced with Black's formula (15):

$$V_{LV}^c = B(K, F, \sigma_B(K, F), T), \quad (29)$$

where its Black implied volatility $\sigma_B(K, F)$ is expressed in terms of the local volatility function σ_{LV} by formula (24). The delta of the local volatility model is given by

$$\Delta_c^{LV} \equiv \frac{\partial V_{LV}^c}{\partial F} = \frac{\partial B}{\partial F} + \frac{\partial B}{\partial \sigma_B} \frac{\partial \sigma_B}{\partial F} = \Delta_c^B + \Lambda^B \frac{\partial \sigma_B}{\partial F}. \quad (30)$$

Its first term is equal to the delta Δ_c^B of Black's model. The second term is proportional to $\partial \sigma_B / \partial F$, the change of the implied volatility with respect to the underlying forward rate. Previous analysis showed that exactly these dynamics are predicted wrongly by the local volatility model. In conclusion, the local volatility model is suited for pricing purposes, but not for proper risk management.

2.3 Constant Elasticity of Variance model

The Constant Elasticity of Variance (CEV) model is an important part of the SABR model. The SABR model is an industry standard in the interest rate derivatives market and it will be introduced in section 2.4. An analysis of the CEV model provides insight in the SABR model. In the CEV model, the forward rate F_t follows the SDE

$$dF_t = \sigma F_t^\beta dW_t, \quad (31)$$

where $\sigma > 0$ is the constant volatility, β is the CEV-exponent or the power parameter and $\{W_t\}_{t \geq 0}$ is a Wiener process. In the interest rate market, the CEV-exponent is usually in the range $0 \leq \beta \leq 1$. If $\beta = 0$, the CEV model reduces to Bachelier's model, that will be introduced in section 3.1. When $\beta = 1$, Black's model from section 2.1 is obtained. This section provides an analysis of the CEV model (31) for $0 < \beta < 1$. More specifically, in this section the transition probability density function $P(F_t = f | F_0)$ of the CEV model (31) is analyzed. It will provide insight in the transition probability density function of the SABR model.

The CEV process can be transformed to a time-changed squared Bessel process, for which the transition density is known [27]. The transition density of the CEV process can be derived from the inverse transformation. The derivation is listed in appendix D and the result is stated next.

The transition probability density of the CEV process (31) is of the following form for $0 < \beta < 1$:

1. For $0 < \beta < \frac{1}{2}$ with an absorbing boundary at $F_t = 0$ and for $\frac{1}{2} \leq \beta < 1$ without (the need of) applying a boundary condition:

$$\begin{aligned} p_A(t, f, F_0) &:= P(F_t = f | F_0) \\ &= \frac{1}{\nu(t)} \left(\frac{f}{F_0} \right)^{-\frac{1}{2}} \exp \left(-\frac{f^{2(1-\beta)} + F_0^{2(1-\beta)}}{2(1-\beta)^2 \nu(t)} \right) I_{|\frac{\delta-2}{2}|} \left(\frac{(F_0 f)^{1-\beta}}{\nu(t)(1-\beta)^2} \right) \frac{f^{1-2\beta}}{1-\beta}. \end{aligned} \quad (32a)$$

2. For $0 < \beta < \frac{1}{2}$ with a reflecting boundary at $F_t = 0$:

$$\begin{aligned} p_R(t, f, F_0) &:= P(F_t = f | F_0) \\ &= \frac{1}{\nu(t)} \left(\frac{f}{F_0} \right)^{-\frac{1}{2}} \exp \left(-\frac{f^{2(1-\beta)} + F_0^{2(1-\beta)}}{2(1-\beta)^2 \nu(t)} \right) I_{\frac{\delta-2}{2}} \left(\frac{(F_0 f)^{1-\beta}}{\nu(t)(1-\beta)^2} \right) \frac{f^{1-2\beta}}{1-\beta}, \end{aligned} \quad (33a)$$

where

$$I_a(x) \equiv \sum_{j=0}^{\infty} \frac{(x/2)^{2j+a}}{j! \Gamma(a+j+1)}, \text{ and } \Gamma(x) \equiv \int_0^{\infty} u^{x-1} e^{-u} du, \quad (34)$$

are the modified Bessel function of the first kind and the gamma function respectively.

The CEV process with an absorbing boundary condition is a martingale, while the CEV process with a reflecting boundary is not [23]. Furthermore, the probability that F_t attains the zero boundary, under the CEV process (31) with $0 < \beta < 1$ and an absorbing boundary condition, is [27]:

$$P(F_t = 0 | F_0) = 1 - \gamma \left(\frac{1}{2(1-\beta)}, \frac{F_0^{2(1-\beta)}}{2(1-\beta)^2 \nu(t)} \right) / \Gamma \left(\frac{1}{2(1-\beta)} \right), \quad (35)$$

where $\gamma(x, y) = \int_0^y u^{x-1} e^{-u} du$ is the lower incomplete gamma function.

Paths $\{F_t\}_{t \geq 0}$ that hit zero, stay in zero (forever) with the absorbing boundary condition, but reflect to positive values with the reflecting boundary condition. Figure 4 shows a plot of the transition probability density functions of the CEV process with an absorbing and with a reflecting boundary condition at zero. The parameters used are CEV-exponent $\beta = 0.10$, current forward rate $F_0 = 0.04$, volatility $\sigma = 0.04$ and time $t = 3$. The SABR model uses an absorbing boundary condition, since this allows the underlying process to be a martingale. Therefore, the plot of the CEV process with an absorbing condition gives insight in the transition probability density of the SABR model.

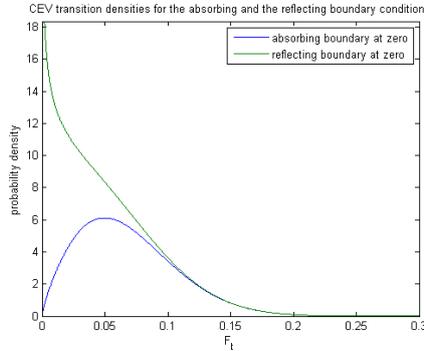


Figure 4: CEV process transition probability density function for absorbing and reflecting boundary conditions, with CEV-exponent $\beta = 0.10$, current forward rate $F_0 = 0.04$, volatility $\sigma = 0.04$ and time $t = 3$.

2.4 Stochastic Alpha Beta Rho model

The dynamics of the market smile predicted by local volatility models is opposite of observed market behaviour [20]. To eliminate this problem, the SABR model was derived by Hagan [20].

Model description

The SABR model [20] is a two factor model with the dynamics given by a system of two stochastic differential equations. The state variables F_t and α_t can be thought of as the forward price of an asset and a volatility parameter respectively. The forward asset can be a forward interest rate or a forward foreign exchange rate. The dynamics of the forward in the SABR model are given by

$$dF_t = \alpha_t F_t^\beta dW_t^{(1)}, \quad F_0 = F > 0, \quad (36a)$$

$$d\alpha_t = \nu \alpha_t dW_t^{(2)}, \quad \alpha_0 = \alpha > 0, \quad (36b)$$

$$dW_t^{(1)} dW_t^{(2)} = \rho dt, \quad (36c)$$

where $0 \leq \beta \leq 1$, $\nu \geq 0$ and $W_t^{(1)}$ and $W_t^{(2)}$ are two ρ -correlated Brownian motions. The parameter ν is the volatility of α_t , so it is the volatility-of-volatility (volvol) of the forward rate.

Hagan's formula

One of the reasons of the popularity of the SABR model is the availability of an explicit expression for the implied Black volatility, called Hagan's formula [20]:

$$\sigma_H(T, K) \approx a(K) b(T, K) \left(\frac{c(K)}{g(c(K))} \right), \quad (37a)$$

where
$$a(K) = \alpha \left[(FK)^{(1-\beta)/2} \left(1 + \frac{(1-\beta)^2}{24} \log^2 \frac{F}{K} + \frac{(1-\beta)^4}{1920} \log^4 \frac{F}{K} \right) \right]^{-1}, \quad (37b)$$

$$b(T, K) = \left[1 + \left(\frac{(1-\beta)^2}{24} \frac{\alpha^2}{(FK)^{1-\beta}} + \frac{\rho\beta\nu\alpha}{4(FK)^{(1-\beta)/2}} + \frac{2-3\rho^2}{24} \nu^2 \right) T \right], \quad (37c)$$

$$c(K) = \frac{\nu}{\alpha} (FK)^{(1-\beta)/2} \log \frac{F}{K} \quad (37d)$$

and
$$g(x) = \log \left(\frac{\sqrt{1-2\rho x + x^2} + x - \rho}{1-\rho} \right). \quad (37e)$$

For at-the-money options, i.e. $K = F$, Hagan's formula reduces to

$$\sigma_H^{ATM} \equiv \sigma_H(T, K = F) \approx \frac{\alpha}{F^{1-\beta}} \left[1 + \left(\frac{(1-\beta)^2}{24} \frac{\alpha^2}{(F)^{2-2\beta}} + \frac{\rho\beta\nu\alpha}{4(F)^{1-\beta}} + \frac{2-3\rho^2}{24} \nu^2 \right) T \right]. \quad (38)$$

Hagan's formula was derived under the condition that $\epsilon = \nu^2 T \ll 1$, where T is the expiry. So Hagan's formula is not suitable for calibration to products with a large expiry, in a market where volatility changes rapidly. Furthermore, Hagan et al. made a small error when deriving (37), which was fixed by Obloj [26]. From here on, the corrected formula of Obloj is referred to as Hagan's formula.

Problems with Hagan's formula

Even though Hagan's formula is easy to implement, it is not suitable for usage in every situation. As one can see in figure 5 the probability density function of the forward rate implied by Hagan's approximation is negative in a part of its domain. This means that Hagan's formula introduces arbitrage [9]. This probability density function is obtained by applying result (12) to call prices corresponding to the implied volatilities generated by Hagan's formula.

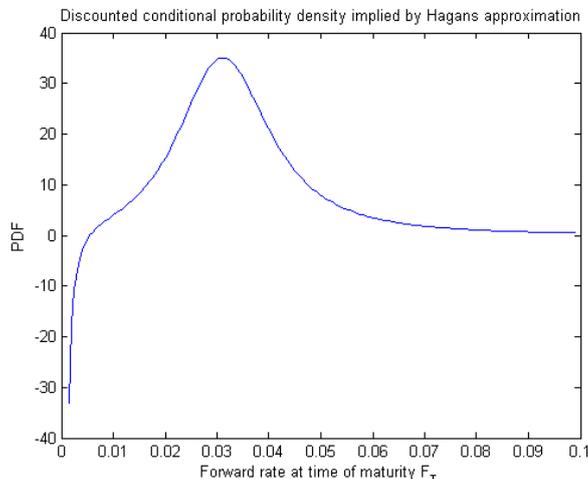


Figure 5: Forward rate probability density function implied by Hagan's approximation. Parameters used are $\rho = 0.5$, $\nu = 0.5$, $\beta = 0.6$, $\alpha = 0.06$, $F_0 = 0.04025$ and $T = 8$.

Exact solution of the zero correlation SABR model

The SABR model gives rise to an exact solution when correlation $\rho = 0$ only. There is no exact solution available for the general correlation case. According to Antonov et al. [2], the time zero value $V^c(0)$ of a call option with strike K , current forward F , discount factor $P(0, t)$, and expiry t , under the SABR model with zero correlation, is given by

$$\begin{aligned} & V^c(0)/P(0, t) - (F - K)^+ \\ &= \frac{2}{\pi} \sqrt{KF} \left\{ \int_{s_-}^{s_+} \frac{\sin(\eta\phi(s))}{\sinh s} G(t\nu^2, s) ds + \sin(\eta\pi) \int_{s_+}^{\infty} \frac{e^{-\eta\psi(s)}}{\sinh s} G(t\nu^2, s) ds \right\}, \end{aligned} \quad (39a)$$

where

$$G(t, s) = 2\sqrt{2} \frac{e^{t/8}}{t\sqrt{2\pi t}} \int_s^{\infty} u e^{\frac{u^2}{2t}} \sqrt{\cosh u - \cosh s} du, \quad (39b)$$

$$\phi(s) = 2 \arctan \sqrt{\frac{\sinh^2 s - \sinh^2 s_-}{\sinh^2 s_+ - \sinh^2 s}}, \quad \psi(s) = 2 \operatorname{arctanh} \sqrt{\frac{\sinh^2 s - \sinh^2 s_+}{\sinh^2 s - \sinh^2 s_-}}, \quad (39c)$$

$$s_- = \operatorname{arcsinh} \left(\frac{\nu|q - q_0|}{\alpha} \right), \quad s_+ = \operatorname{arcsinh} \left(\frac{\nu(q + q_0)}{\alpha} \right), \quad (39d)$$

$$q = \frac{K^{1-\beta}}{1-\beta}, \quad q_0 = \frac{F^{1-\beta}}{1-\beta}, \quad \text{and } \eta = \left| \frac{1}{2(\beta - 1)} \right|. \quad (39e)$$

The inner integral G can be efficiently approximated by a (numerically) simple function:

$$G(t, s) \approx \sqrt{\frac{\sinh s}{s}} e^{-\frac{s^2}{2t} - \frac{t}{s}} \left(R(t, s) + \tilde{R}(t, s) \right), \text{ where} \quad (39f)$$

$$R(t, s) = 1 + \frac{3tg(s)}{8s^2} - \frac{5t^2(-8s^2 + 3g^2(s) + 24g(s))}{128s^4} + \frac{35t^2(-40s^2 + 3g^3(s) + 24g^2(s) + 120g(s))}{1024s^6}, \quad (39g)$$

$$g(s) = s \coth s - 1 \text{ and } \tilde{R}(t, s) = e^{t/8} - \frac{3072 + 384t + 24t^2 + t^3}{3072}. \quad (39h)$$

This last function $\tilde{R}(t, s)$ is a correction to guarantee that $G(t, 0) = 1$.

Since Hagan's formula (37) was derived under the condition that $\epsilon = \nu^2 T \ll 1$, the formula is biased for certain parameter combinations. The zero correlation solution (39) is useful to check whether Hagan's formula is an accurate approximation of the volatility smile produced by the SABR model. Furthermore, Antonov's (39) formula is also useful to check the correctness of a Monte Carlo implementation of the SABR model.

Influence parameters on the shape of the curve

A detailed analysis of the effects of the SABR parameters on the shape of the implied volatility curve is given in section 4.3. A brief summary is listed here, to gain intuition of Hagan's formula and the SABR model. A change in the initial volatility parameter α has the effect of shifting the curve up or down, as can be seen in figure 14b. Figures 14a and 14c show the sensitivity of the curve to the ν and ρ parameters respectively. Volatility of volatility parameter ν controls the curvature of the volatility smile. Correlation parameter ρ mainly effects the skew. Since the parameters all have a different effect on the curve, the fitted parameters tend to be stable [20].

The sensitivity of the curve to β is shown in figure 15. As β is usually fixed before estimating the other parameters, this behaviour is less relevant for a stability analysis of the calibration process. Nevertheless, the figure gives insight in the influence of β on the curve when all parameters would be estimated.

Calibration of SABR parameters to market data

An option pricing model is said to be calibrated to market data when the prices generated by the model coincide with the prices observable in the market. Or equivalently, a pricing model is calibrated when the model implied volatilities match the market implied volatilities. Models are calibrated to market quotes of simple options, such as European-type options, because they contain information about the market. The calibrated model can be used to price similar options, but with a different strike or expiry that is not quoted in the market. In most cases, there are no market quotes of complex options available. Also, these complex options can be priced with a model that is calibrated to (the appropriate) simple options.

In market practice, the parameter β is usually fixed before calibration, since the other parameters can replace the effect of β on the shape of the volatility smile [20]. The risk metrics of the SABR model are sensitive to β though, so for risk management purposes β is important [6].

After β is fixed, the remaining parameters ν, ρ and α need to be estimated. This calibration problem can be seen as a global minimization problem over a loss function between market implied volatilities of options with the same expiry, and implied volatilities of Hagan's formula (37). A common choice for the loss function is the sum of squared errors (SSE):

$$(\hat{\nu}, \hat{\rho}, \hat{\alpha}) = \arg \min_{\nu, \rho, \alpha} \sum_i [\sigma_i^M - \sigma_H(F_i, K_i; \nu, \rho, \alpha)]^2, \quad (40)$$

where σ_H is the Hagan approximation (37) with a fixed expiry T that is equal for all the quotes used. The market volatilities are denoted by σ_i^M and their current forward rates are denoted by F_i . The above parameterization will be referred to as the *first parameterization* of SABR.

The number of parameters in the above optimisation can be reduced from three to two by using the following approximation for formula (38):

$$\log \sigma_H^{ATM} \approx \log \alpha - (1 - \beta) \ln F, \quad (41)$$

in combination with at-the-money volatilities observed in the market [34]. The α -parameter can be expressed as $\alpha = \alpha(\nu, \rho)$, since equation (38) can be rewritten as a polynomial of α :

$$\left[\frac{(1 - \beta)^2 T}{24 F^{2-2\beta}} \right] \alpha^3 + \left[\frac{\rho \beta \nu T}{4 F^{1-\beta}} \right] \alpha^2 + \left[1 + \frac{2 - 3\rho^2}{24} \nu^2 T \right] \alpha - F^{1-\beta} \sigma_H^{ATM} = 0. \quad (42)$$

from which the root can be found numerically. In case there are several real roots, it is optimal to choose the smallest root according to West [34]. This parameterization will be referred to later as the *second parameterization*.

Risk metrics of the SABR model

As mentioned before, a European call option on a forward rate can be priced under the SABR model by using Hagan's approximation (37) of the implied Black volatility together with Black's model (15):

$$V_{SABR}^c = B(K, F, \sigma, T) \text{ with } \sigma = \sigma_H(K, F; \alpha, \beta, \rho, \nu). \quad (43)$$

The first parameterization expresses Black implied volatility σ as a function of $(\alpha, \beta, \rho, \nu)$. In contrast, the second parameterization expresses σ as a function of $(\sigma_H^{ATM}, \beta, \rho, \nu)$, since α is obtained by solving formula (42) to obtain $\alpha(\sigma_H^{ATM}, F)$. According to the chain rule of differentiation their respective delta values are given by

$$\Delta_c^{SABR1} \equiv \frac{\partial V_{SABR}^c}{\partial F} = \frac{\partial B}{\partial F} + \frac{\partial B}{\partial \sigma} \frac{\partial \sigma_H}{\partial F} = \Delta_c^B + \Lambda^B \frac{\partial \sigma_H}{\partial F}, \text{ and} \quad (44a)$$

$$\Delta_c^{SABR2} \equiv \frac{\partial V_{SABR}^c}{\partial F} = \frac{\partial B}{\partial F} + \frac{\partial B}{\partial \sigma} \left[\frac{\partial \sigma_H}{\partial F} + \frac{\partial \sigma_H}{\partial \alpha} \frac{\partial \alpha}{\partial F} \right] = \Delta_c^B + \Lambda^B \left[\frac{\partial \sigma_H}{\partial F} + \frac{\partial \sigma_H}{\partial \alpha} \frac{\partial \alpha}{\partial F} \right]. \quad (44b)$$

In both parameterizations the delta under the SABR model is equal to the delta Δ^B of Black's model (17a) plus a correction term that is proportional to the vega Λ^B of Black's model (17c). The vega under the SABR model does not depend on which of the two parameterizations is used:

$$\Lambda^{SABR} \equiv \frac{\partial V_{SABR}^c}{\partial \alpha} = \frac{\partial B}{\partial \sigma} \frac{\partial \sigma_H}{\partial \alpha} = \Lambda^B \frac{\partial \sigma_H}{\partial \alpha}. \quad (45)$$

These risk metrics tend to be dependent on the choice of β . Bartlett [6] improved the stability of the risk metrics by using information on the dynamics of SABR (36) in the derivation of the risk metrics. In the previous derivation of the delta the value of the second factor α_t was fixed while shifting the first factor F_t :

$$F \rightarrow F + \Delta F, \quad (46a)$$

$$\alpha \rightarrow \alpha. \quad (46b)$$

The derivation of vega was done in exactly the opposite way:

$$F \rightarrow F, \quad (47a)$$

$$\alpha \rightarrow \alpha + \Delta \alpha. \quad (47b)$$

But the two factors in the dynamics of SABR (36) are correlated by correlation ρ , so when one factor changes the other factor is also likely to change. When deriving the delta it may be better to introduce a shift for α that depends on the shift in F :

$$F \rightarrow F + \Delta F, \quad (48a)$$

$$\alpha \rightarrow \alpha + \tilde{\Delta}\alpha. \quad (48b)$$

Since the dynamics are not deterministic but stochastic the shift $\tilde{\Delta}\alpha$ will represent the average change or shift resulting from a shift in F . To obtain the value of $\tilde{\Delta}\alpha$ first express the dynamics in two independent Brownian Motions $dW_t^{(1)}$ and dZ_t by using Cholesky decomposition [31]:

$$dF_t = \alpha_t F_t^\beta dW_t^{(1)}, \quad (49a)$$

$$d\alpha_t = \nu\alpha_t \left(\rho dW_t^{(1)} + \sqrt{1 - \rho^2} dZ_t \right), \quad (49b)$$

$$E \left[dW_t^{(1)} dZ_t \right] = 0, \quad (49c)$$

where $dW_t^{(2)} = \left(\rho dW_t^{(1)} + \sqrt{1 - \rho^2} dZ_t \right)$. These expressions for the factors imply that

$$d\alpha_t = \frac{\nu\rho}{F_t^\beta} dF_t + \nu\alpha_t \sqrt{1 - \rho^2} dZ_t. \quad (50)$$

Differential $d\alpha_t$ represents the infinitesimal version of $\tilde{\Delta}\alpha$. Averaging the $\tilde{\Delta}\alpha$'s corresponds to taking expectations in the above equation. The dZ_t term disappears as increments of a Brownian motion have expectation zero. This results in

$$\tilde{\Delta}\alpha = \frac{\nu\rho}{F^\beta} \Delta F \text{ or } \frac{d\alpha_t}{dF_t} = \frac{\nu\rho}{F_t^\beta} \text{ on average.} \quad (51)$$

As a result the delta of SABR is

$$\Delta_c^{SABR} = \frac{\partial B}{\partial F} + \frac{\partial B}{\partial \sigma} \left[\frac{\partial \sigma_H}{\partial F} + \frac{\partial \sigma_H}{\partial \alpha} \frac{\partial \alpha}{\partial F} \right] = \Delta_c^B + \Lambda^B \left[\frac{\partial \sigma_H}{\partial F} + \frac{\partial \sigma_H}{\partial \alpha} \frac{\nu\rho}{F^\beta} \right]. \quad (52)$$

The last term is in fact equal to $\nu\rho/F^\beta$ times the old vega of SABR in equation (45). In a vega-hedged portfolio this term is by definition zero. So if (old) vega and delta risks are both hedged, the new delta risk is also hedged for this portfolio [6].

The calculation of the improved vega is based on the same idea as is used in the derivation of the improved delta. A shift in α results in a shift in F through correlation:

$$F \rightarrow F + \tilde{\Delta}F, \quad (53a)$$

$$\alpha \rightarrow \alpha + \Delta\alpha. \quad (53b)$$

The SABR dynamics can be expressed in two independent Brownian motions Y_t and $W_t^{(2)}$ by Cholesky decomposition [31]:

$$dF_t = \alpha_t F_t^\beta \left(\rho dW_t^{(2)} + \sqrt{1 - \rho^2} dY_t \right), \quad (54a)$$

$$d\alpha_t = \nu\alpha_t dW_t^{(2)}, \quad (54b)$$

$$E \left[dY_t dW_t^{(2)} \right] = 0. \quad (54c)$$

Combining the above two factors results in

$$dF_t = \frac{F_t^\beta \rho}{\nu} d\alpha_t + \alpha_t F_t^\beta \sqrt{1 - \rho^2} dY_t. \quad (55)$$

So on average

$$\tilde{\Delta}_F \frac{\rho F_t^\beta}{\nu} = \Delta\alpha \text{ or } \frac{dF_t}{d\alpha_t} = \frac{\rho F_t^\beta}{\nu}. \quad (56)$$

The improved SABR vega reads

$$\Lambda^{SABR} = \frac{\partial B}{\partial \sigma} \frac{\partial \sigma_H}{\partial \alpha} + \frac{\partial B}{\partial F} \frac{\partial F}{\partial \alpha} = \Lambda^B \frac{\partial \sigma_H}{\partial \alpha} + \Delta_c^B \cdot \frac{\rho F_t^\beta}{\nu} = \Lambda_{old}^{SABR} + \Delta_c^B \cdot \frac{\rho F_t^\beta}{\nu}, \quad (57)$$

where Λ_{old}^{SABR} is the vega in formula (45). SABR is a single self-consistent model for all strikes: calculated risks at one strike are consistent with risks calculated at other strikes. As a result, the risks of all the options on the same underlying can be added together, only the remaining risk needs to be hedged[20].

If β is chosen to be fixed before calibration it is preferred that the calibrated model is relatively insensitive to changes in β . Figure 6c shows the sensitivity of the call delta to β under a SABR model with the first parameterisation. The (improved) second parameterisation is even less sensitive to changes in β as can be seen in figure 6d. The difference in the two delta's can be expressed in terms of the sensitivity of the implied volatility to the current forward $d\sigma_B/dF$. This quantity for the first parameterisation and the improved second parameterisation are plotted in figures 6a and 6b, respectively. The latter is less sensitive to β .

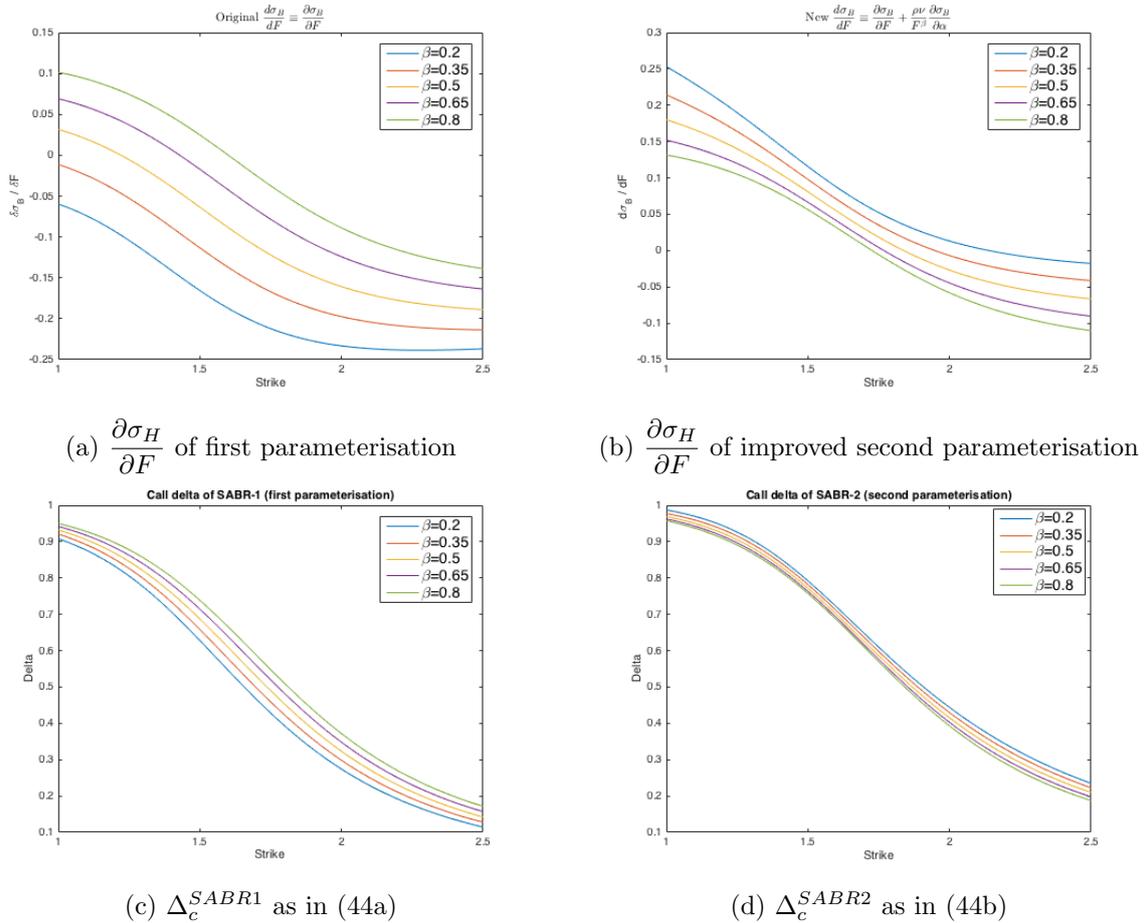


Figure 6: Sensitivity of risk measures under the SABR model. For each fixed β Hagan's formula was calibrated to a certain set of market quotes. The β parameter was varied in both figures. Furthermore, in both figures the other parameters are fixed: time to maturity $T = 1$, discount factor $P(0, T) = 1$, and current forward $F = 1.75$.

2.5 Summary

In the past, industry believed that interest rates were always positive. Both the assumption of lognormal dynamics for the forward rate and the assumption of a boundary condition at zero express this belief. The models discussed in this section were industrial practice until low and negative interest rates appeared in the market. The assumption of positive rates is embedded in their structure; all the models assign a zero probability to the rate going negative. Foreign exchange rates are positive by definition and the models introduced in this section are still appropriate for modeling foreign exchange rates.

Black's model has a constant volatility, while this is not observed in the market. When Black's model is extended with the local volatility model, it can be calibrated to market prices of several expiries. The dynamics of the market implied volatility smile with respect to a change in the underlying rate is different from the implied volatility smile predicted by the local volatility model. The SABR model solves this problem. Hagan's approximation formula of Black's volatility under the SABR model can be used to calibrate an implied volatility smile to market data with dynamics in line with market behaviour.

All the models share the drawback that they cannot be used in an economy with small or negative interest rates (in their current form). Methods to circumvent this problem are explored in chapter 3.

3 Negative rate modelling

The models explained in chapter 2 cannot cope with negative interest rates. The corresponding probability density functions are zero for rates less or equal to zero. Two solutions come to mind: shifting the boundary condition by a positive constant s , such that rates larger than $-s$ are allowed, or removing the boundary condition. In section 3.2 and 3.3 these two methods are applied to the models from chapter 2. But first, the normal model is introduced in section 3.1. This model allows for negative rates in a natural way. In current market practice either the implied shifted lognormal volatility is quoted together with the shift parameter or the implied normal volatility is quoted. At the end of the chapter, the models are compared in the summary in section 3.4.

3.1 Bachelier's model and the normal SABR model

The normal model, introduced in 1900 by Bachelier [4], is the most simple model that comes to mind as a way of modelling negative interest rates. In the normal model, the forward rate F_t follows the SDE

$$dF_t = \sigma_N dW_t, \quad (58)$$

where σ_N represents the normal volatility. The solution of this model can be easily found by Itô integration:

$$F_t = F_0 + \sigma_N W_t, \quad (59)$$

which means that $F_t \sim \mathcal{N}(F_0, \sigma_N^2 t)$. This model allows for negative interest rates in a natural way since its distribution is normal, see figure 7a for an example.

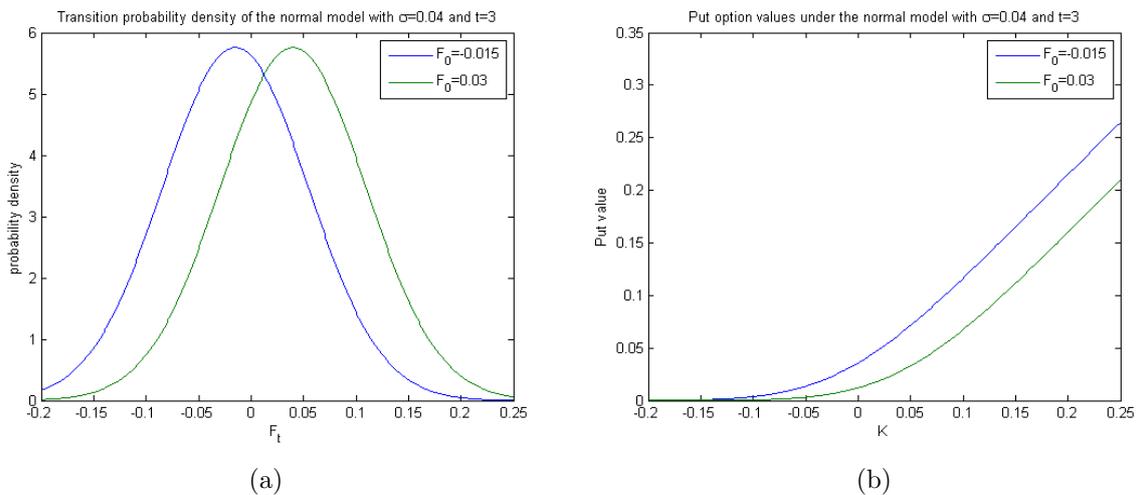


Figure 7: Figure (a) shows the transition density of the normal model for different forwards. Figure (b) shows a plot of the value of a put under the normal model for different strikes.

Pricing calls and puts under the normal model

When the underlying forward rate follows the normal model, the European call and put values are given by:

$$V^c(0) = P(0, T) \left[(F_0 - K)\Phi(d) + \sigma\sqrt{T}\phi(d) \right] \quad \text{and} \quad (60a)$$

$$V^p(0) = P(0, T) \left[(K - F_0)\Phi(-d) + \sigma\sqrt{T}\phi(d) \right], \quad \text{where } d = \frac{F_0 - K}{\sigma\sqrt{T}}. \quad (60b)$$

The derivation of these equations can be found in appendix C. The normal model allows valuation of options with negative strikes and negative current forward rates, in contrast to the lognormal

model. Figure 7b shows the value of a European put for a negative and a positive underlying forward rate. In the lognormal model a put with strike zero has zero value by definition. The value of a put with any strike under the normal model is strictly positive, since any (positive or negative) forward rate has a nonzero probability of being attained.

Risk metrics under the normal model

We can also compute the option's delta, vega, and gamma, as follows

$$\Delta_c \equiv \frac{\partial V^c}{\partial F} = P(0, T)\Phi(d), \Delta_p \equiv \frac{\partial V^p}{\partial F} = -P(0, T)\Phi(-d), \quad (61a)$$

$$\Lambda \equiv \frac{\partial V^c}{\partial \sigma} = P(0, T)\sqrt{T}\phi(d), \quad (61b)$$

$$\Gamma^c \equiv \frac{\partial^2 V^c}{\partial F^2} = P(0, T)\frac{\phi(d)}{\sigma\sqrt{T}} \text{ and } \Gamma^p \equiv \frac{\partial^2 V^p}{\partial F^2} = P(0, T)\frac{\phi(d)}{\sigma\sqrt{T}}. \quad (61c)$$

The derivation of these risk metrics can be found in appendix C. The risk metrics are strictly positive for every combination of strike and forward.

Normal SABR ($\beta = 0$) volatility smile extension for Bachelier's model

Hagan's formula is used to calibrate an implied Black volatility smile. A similar formula is available for Bachelier's model to calibrate an implied Bachelier volatility smile:

$$\sigma_N(K) = \alpha \frac{\zeta}{x(\zeta)} \left(1 + \frac{2 - 3\rho^2}{24} \nu^2 T \right), \quad (62a)$$

$$\text{where } \zeta = \frac{\nu}{\alpha} (F_0 - K), \text{ and} \quad (62b)$$

$$x(\zeta) = \log \left(\frac{\sqrt{1 - 2\rho\zeta + \zeta^2} - \rho + \zeta}{1 - \rho} \right). \quad (62c)$$

This formula is derived in [20]. The domain of the implied Bachelier volatility function (62) is the entire real line, since Bachelier's option pricing formula (60a) allows strikes (and forwards) from the entire real line. The SABR model (36) with $\beta = 0$ is the only version of the SABR model that can model negative forward rates. This version of the SABR model is called the normal SABR model:

$$dF_t = \alpha_t dW_t^{(1)}, \quad F_0 = F, \quad (63a)$$

$$d\alpha_t = \nu\alpha_t dW_t^{(2)}, \quad \alpha_0 = \alpha, \quad (63b)$$

$$E \left[dW_t^{(1)} dW_t^{(2)} \right] = \rho dt, \quad (63c)$$

The normal SABR model (63) is calibrated to an implied Bachelier volatility smile with formula (62).

Risk metrics under the normal SABR model

The method of Bartlett [6] used previously to determine the delta and vega under the SABR dynamics with $\beta > 0$ can be applied to the normal SABR model:

$$\Delta_c \equiv \frac{\partial V^c}{\partial F} = \frac{\partial N}{\partial F} + \frac{\partial N}{\partial \sigma_N} \left(\frac{\partial \sigma_N}{\partial F} + \frac{\partial \sigma_N}{\partial \alpha} \frac{\partial \alpha}{\partial F} \right) = \frac{\partial N}{\partial F} + \frac{\partial N}{\partial \sigma_N} \left(\frac{\partial \sigma_N}{\partial F} + \frac{\partial \sigma_N}{\partial \alpha} \nu \rho \right), \quad (64a)$$

$$\Lambda \equiv \frac{\partial V^c}{\partial \sigma} = \frac{\partial N}{\partial \sigma_N} \frac{\partial \sigma_N}{\partial \alpha} + \frac{\partial N}{\partial F} \frac{\rho}{\nu}, \quad (64b)$$

where $N = N(K, F, \sigma_N, T)$ is the value of a call under Bachelier's model (60a) and $\sigma_N = \sigma_N(K, F, T; \alpha, \rho, \nu)$ is the implied normal volatility (62). The option's put delta Δ_p can be calculated with the put-call parity (16).

Calibration of normal SABR

Calibration of the normal SABR model is similar to calibrating the SABR model for $\beta > 0$;

$$(\hat{\nu}, \hat{\rho}, \hat{\alpha}) = \arg \min_{\nu, \rho, \alpha} \sum_i [\sigma_i^M - \sigma_N(F_i, K_i; \nu, \rho, \alpha)]^2, \quad (65)$$

where σ_i^M are market implied volatilities for a certain maturity T .

3.2 Displaced models

Another class of models is the class of shifted or displaced models. In these models, forward rate F_t is replaced with shifted forward rate $F_t + s$, where s is a constant. In this section the displaced versions of Black's model (15), the CEV model (31), and the SABR model (36) are discussed.

Displaced diffusion

Forward rate F_t is said to be a displaced diffusion process if it is the solution of SDE [11]:

$$dF_t = d(F_t + s) = \sigma(F_t + s) dW_t, \quad (66)$$

where s is a constant displacement parameter. Note that $\tilde{F}_t \equiv (F_t + s)$ follows a lognormal process, or Black process (13). This fact, together with the fact that the payoff of a European call option $\max(F_T - K, 0)$ can be rewritten as:

$$\max(F_T - K, 0) = \max((F_T + s) - (K + s), 0) \equiv \max(\tilde{F}_T - \tilde{K}, 0), \quad (67)$$

leads to the conclusion that European calls and puts can be valued under the displaced diffusion model by plugging in $\tilde{F}_0 \equiv (F_0 + s)$ and $\tilde{K} = (K + s)$ in Black's model (15). Risk metrics can be obtained by plugging in these shifted parameters as well. The same principle can be applied to obtain the solution of SDE (71):

$$\tilde{F}_t = \tilde{F}_0 \exp(\sigma W_t - \sigma^2 t/2), \text{ so } F_t = -s + (F_0 + s) \exp(\sigma W_t - \sigma^2 t/2). \quad (68)$$

This solution shows that rates (and strikes) larger than $-s$ can be modelled with the displaced diffusion model, see figure 8 for some examples. Furthermore, the first moment is conserved:

$$\mathbb{E}[F_t | \mathcal{F}_0] = \mathbb{E}[\tilde{F}_t - s | \mathcal{F}_0] = \tilde{F}_0 - s = (F_0 + s) - s = F_0. \quad (69)$$

Displaced diffusion as an approximation of the CEV process

It is interesting to note that there exists a close relation between the CEV (31) and displaced diffusion (71) processes. The displaced diffusion model may be considered as a first-order approximation to the CEV dynamics [27]. Define $g(x) = x^\beta$, then a first-order Taylor expansion of $g(x)$ around $x = F_0$ is given by,

$$g(x) = g(F_0) + \frac{dg}{dx}(F_0) \cdot (x - F_0) + \mathcal{O}(x^2), \quad (70a)$$

so equivalently, when evaluated at $x = F_t$,

$$F_t^\beta \approx F_0^\beta + \beta F_0^{\beta-1} (F_t - F_0) = c_1 [c_2 + F_t], \text{ where } c_1 \equiv \beta F_0^{\beta-1} \text{ and } c_2 \equiv \left(\frac{1}{\beta} - 1\right) F_0. \quad (70b)$$

So, CEV process (31) can be approximated by a displaced diffusion, as follows

$$dF_t \approx \sigma c_1 [c_2 + F_t] dW_t, \quad (70c)$$

with volatility $c_1 \sigma$ and shift c_2 . Since the SABR model is a CEV process with stochastic volatility, displaced diffusion can be seen as an approximation for SABR as well.

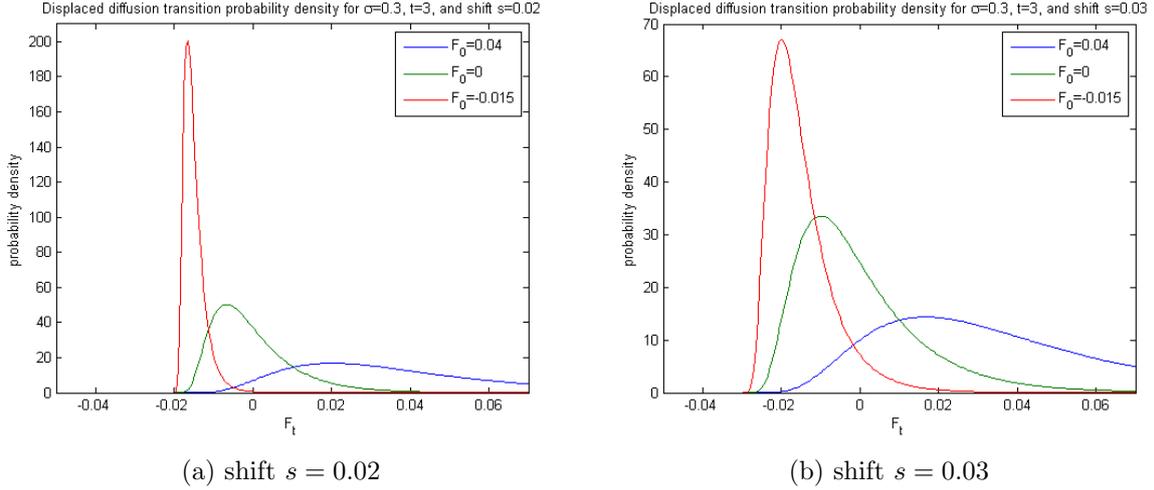


Figure 8: Figure (a) and (b) show the transition probability densities of the displaced diffusion model for several rates F_0 and two different shift parameters s .

Displaced SABR model

Where the classic SABR model only allows rates to be nonnegative, a shifted model with shift $s > 0$ allows rates larger than $-s$ to be modelled. So if the (constant) shift is $s = 0.02$ the displaced model can model rates larger than -2.0% . The displaced SABR model is defined as

$$dF_t = \alpha_t (F_t + s)^\beta dW_t^{(1)}, \quad F_0 = F, \quad (71a)$$

$$d\alpha_t = \nu \alpha_t dW_t^{(2)}, \quad \alpha_0 = \alpha, \quad (71b)$$

$$E \left[dW_t^{(1)} dW_t^{(2)} \right] = \rho dt, \quad (71c)$$

where F_t is a forward rate. A similar analysis as for displaced diffusion shows that $\tilde{F}_t = (F_t + s)$ follows the SABR SDE (36).

Drawback shift parameter

A drawback of the displaced SABR model and the displaced diffusion model is that the shift-parameter needs to be selected a priori, since it is not known how low the interest rates can go. This may lead to a situation when the value of this shift needs to be changed and subsequently the present value and risk of the entire portfolio need to be recomputed[3]. Free Boundary SABR seems a more natural solution to permit negative interest rates.

3.3 Free boundary models

Just as an analysis of the CEV model provides more insight in the SABR model, an analysis of the free boundary CEV model provides more insight in the free boundary SABR model, which will be introduced next. In the free boundary CEV (FB-CEV) model, the forward rate F_t follows the SDE

$$dF_t = \sigma |F_t|^\beta dW_t, \quad (72)$$

where σ is the volatility, β is the CEV-exponent and $\{W_t\}_{t \geq 0}$ is a Wiener process. As later analysis will show, $0 \leq \beta < 1/2$ should hold in order to obtain the constructed solution.

The transition probability density function of the FB-CEV model gives insight in the transition probability density function of the FB-SABR model. The first probability density function is constructed from the absorbing and reflecting solutions of the CEV model as follows. Let $q(t, f, F_0) := P(F_t = f | F_0)$ and $p(t, f, F_0) := P(F_t = f | F_0)$ be the transition probability

density functions of the free boundary CEV process (72) and the ordinary CEV process (31) respectively. The Fokker-Planck equation for the transition density of the free boundary CEV model (72) for $0 \leq t \leq T$ and $F \in \mathbb{R}$ is

$$\frac{\partial}{\partial t} q(t, f, F_0) = \frac{1}{2} \frac{\partial^2}{\partial f^2} \left[|f|^{2\beta} q(t, f, F_0) \right], \text{ with } q(t, F_0, F_0) = \delta(F_0). \quad (73)$$

The Fokker-Planck equation for the (ordinary) CEV process (31) for $0 < T$ and $F \geq 0$ is

$$\frac{\partial}{\partial t} p(t, f, F_0) = \frac{1}{2} \frac{\partial^2}{\partial f^2} \left[f^{2\beta} p(t, f, F_0) \right], \text{ with } p(0, F_0, F_0) = \delta(F_0), \quad (74)$$

which has an absorbing solution (32) called $p_A(t, f, F_0)$, and a reflecting solution (33) called $p_R(t, f, F_0)$. Fokker-Planck equation (74) with $f \geq 0$ is equivalent to Fokker-Planck equation (73). So when looking for a solution of the free boundary model, it is natural to consider a linear(-like) combination of $p_A(t, |f|, |F_0|)$ and $p_R(t, |f|, |F_0|)$.

Antonov et al. [3] propose the solution:

$$q(t, f, F_0) = \frac{1}{2} [p_R(t, |f|, |F_0|) + \text{sign}(f \cdot F_0) p_A(t, |f|, |F_0|)]. \quad (75)$$

which satisfies (73) by linearity of differentiation and the fact that p_A and p_R satisfy (74) for $f \geq 0$. Above solution (75) only holds for $0 \leq \beta < \frac{1}{2}$, since reflecting solution $p_R(t, f, F_0)$ is only available for $0 \leq \beta < \frac{1}{2}$.

This solution can be seen as the average of an absorbing part; $\text{sign}(f \cdot F_0) p_A(t, |f|, |F_0|)$, and a reflecting part: $p_R(t, |f|, |F_0|)$. These parts together with the solution are plotted in figure 9a to give some intuition about the structure of the solution. Figure 9b compares the solution of the FB-CEV model with the absorbing and reflecting solutions of the ordinary CEV model. The FB-CEV model has a spike in its probability density around zero, caused by the reflecting part of the solution. Figure 9c shows the effect of β on the shape of the spike.

As expected, the densities of the FB-CEV model for $\beta = 0$ and the normal model are equal, see figure 9d.

To check if solution (75) of (73) is a proper transition probability density function it needs to be checked if it integrates to one and whether the first moment equals the current forward rate, or not. Furthermore, nonnegativity of q is guaranteed due to the fact that every path that hits zero is reflected by p_R but absorbed by p_A and therewith $p_A(t, f, F_0) \leq p_R(t, f, F_0)$. It appears that the combination of the sign function and the weights $\frac{1}{2}$ in (75) are essential for obtaining a solution that guarantees those first two properties:

$$\mathbb{E}[F_t | F_0], \quad (76a)$$

$$= \int_{-\infty}^0 f \cdot \frac{1}{2} (p_R(t, -f, |F_0|) - p_A(t, -f, |F_0|)) df + \int_0^{\infty} f \cdot \frac{1}{2} (p_R(t, f, |F_0|) + p_A(t, f, |F_0|)) df, \quad (76b)$$

$$= \int_{-\infty}^0 \frac{f}{2} p_R(t, -f, |F_0|) df - \int_{-\infty}^0 \frac{f}{2} p_A(t, -f, |F_0|) df + \int_0^{\infty} \frac{f}{2} p_R(t, f, |F_0|) df + \frac{1}{2} F_0, \quad (76c)$$

$$= - \int_0^{\infty} \frac{f}{2} p_R(t, f, |F_0|) df + \frac{1}{2} F_0 + \int_0^{\infty} \frac{f}{2} p_R(t, f, |F_0|) df + \frac{1}{2} F_0 = F_0. \quad (76d)$$

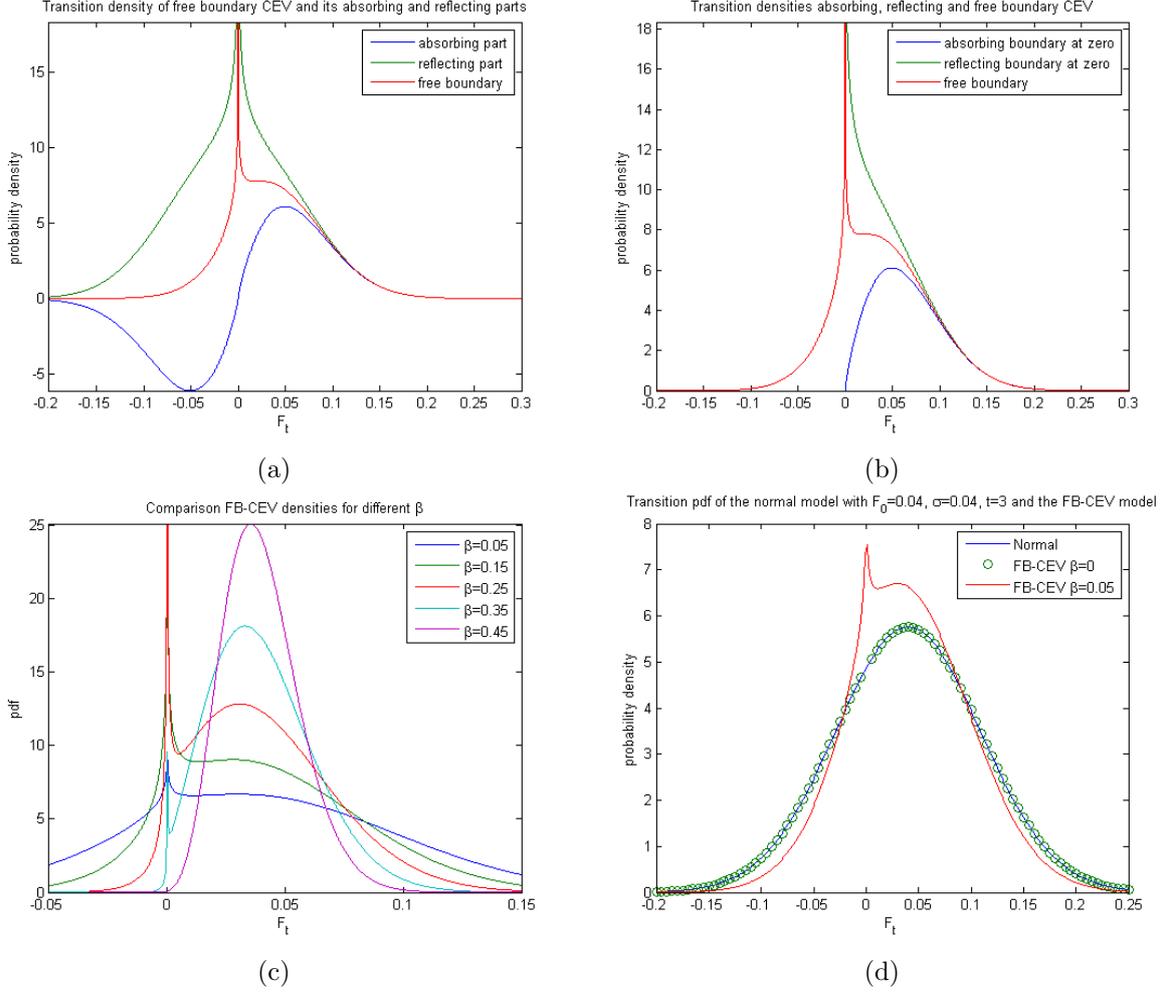


Figure 9: Figure (a) shows a plot of the transition density of the FB-CEV model together with its absorbing and reflecting parts. Figure (b) compares the transition probability density function of the FB-CEV model with the transition probability density functions of the CEV model with an absorbing or reflecting boundary. The CEV model parameters used in the plots are CEV-exponent $\beta = 0.1$, current forward $F_0 = 0.04$, volatility $\sigma = 0.04$ and time $t = 3$. In figure (c) the same parameters were used, except for β , which was varied. Figure (d) shows a comparison between the density of the normal model and the densities of the FB-CEV model for $\beta = 0.05$ and 0.

$$\int_{-\infty}^{\infty} q(t, f, |F_0|) df, \quad (76e)$$

$$= \int_{-\infty}^0 \frac{1}{2} (p_R(t, -f, |F_0|) - p_A(t, -f, |F_0|)) df + \int_0^{\infty} \frac{1}{2} (p_R(t, f, |F_0|) + p_A(t, f, |F_0|)) df, \quad (76f)$$

$$= \int_0^{\infty} \frac{1}{2} (p_R(t, f, |F_0|) - p_A(t, f, |F_0|)) df + \int_0^{\infty} \frac{1}{2} (p_R(t, f, |F_0|) + p_A(t, f, |F_0|)) df, \quad (76g)$$

$$= \int_0^{\infty} p_R(t, f, |F_0|) df = 1. \quad (76h)$$

The sign function cancels out the absorbing factors in the calculation of the integrated probability and it cancels out the reflecting factors in the calculation of the first moment. The weights $\frac{1}{2}$ are necessary to scale the first moment to F_0 and the make sure the probability density function integrates to 1. The free boundary CEV model combines the best of two worlds: it inherits the conservation of first moment from the absorbing CEV model (32) and it inherits the integration-to-one property from the reflecting CEV model (33). Other functions from the solution space

$$\{c_A \text{sign}(f \cdot F_0)^i p_A(t, |f|, |F_0|) + c_R \text{sign}(f \cdot F_0)^j p_R(t, |f|, |F_0|) \mid c_A, c_R \in \mathbb{R}, i, j \in \{1, 2\}\} \quad (77)$$

do not have these properties simultaneously. In summary, the FB-CEV model has a solution (75) based on the absorbing and reflecting solutions of the ordinary CEV model. Moreover, the solution has a spike around zero. The free boundary SABR model is analyzed next.

Free Boundary SABR

The free boundary SABR (FB-SABR) model is an extension of the classic SABR model (36) that allows forward rate values from the full real line, $F_t \in \mathbb{R}$, by putting absolute signs around F_t [3]. This is not the only modification; the classic model uses an absorbing boundary at $F_t = 0$ while the FB-SABR model, as is clear from its name, does not have a boundary condition. This changes the nature of the solutions from the model. The dynamics of the FB-SABR model are given by

$$dF_t = \alpha_t |F_t|^\beta dW_t^{(1)}, \quad F_0 = F, \quad (78a)$$

$$d\alpha_t = \nu \alpha_t dW_t^{(2)}, \quad \alpha_0 = \alpha, \quad (78b)$$

$$E \left[dW_t^{(1)} dW_t^{(2)} \right] = \rho dt, \quad 0 \leq \beta < \frac{1}{2}. \quad (78c)$$

The condition $0 \leq \beta < \frac{1}{2}$ is necessary for stable solutions, as explained in the previous section.

Analytical solutions for the FB-SABR model are available for the cases that $\nu = 0$ or that $\rho = 0$ only. In the first case the model is reduced to the FB-CEV model. In the second case the zero correlation solution can be found in [3]. Figure 10a and 10b show that the Euler Monte Carlo scheme performs well for FB-CEV. When $\nu \neq 0$, there is no analytical solution of the FB-SABR model available in general, but Monte Carlo simulations may give insight. The probability densities are derived by calculating the second order derivatives of the simulated call prices with respect to the strike, according to result (12). Figures 10c and 10d show that FB-SABR is asymmetric in F_t , in contrast to the FB-CEV model in figures 10a and 10b. This asymmetry is caused by the lognormal stochastic volatility process.

Calibrate the FB-SABR model

The original article of Hagan et al. [20] that introduced the SABR model included the derivation of the implied Black and implied Bachelier volatilities for a general coefficient $C(F)$ of system

$$dF_t = \alpha_t C(F_t) dW_t^{(1)}, \quad F_0 = F, \quad (79a)$$

$$d\alpha_t = \nu \alpha_t dW_t^{(2)}, \quad \alpha_0 = \alpha, \quad (79b)$$

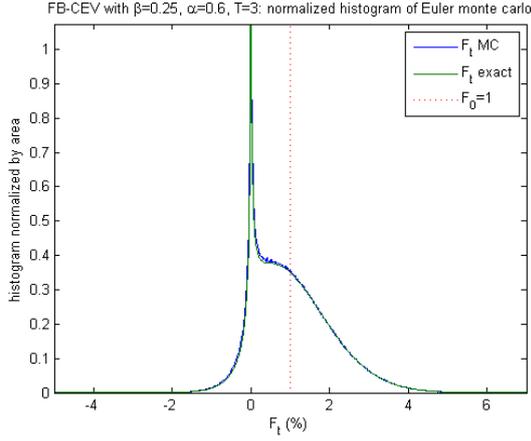
$$E \left[dW_t^{(1)} dW_t^{(2)} \right] = \rho dt. \quad (79c)$$

Hagan's formula is derived by taking $C(F) = F^\beta$. Since FB-SABR can model rates from the whole real line, it is natural to use implied Bachelier volatilities. The implied Bachelier volatility smile for $C(F) = |F|^\beta$ is given by:

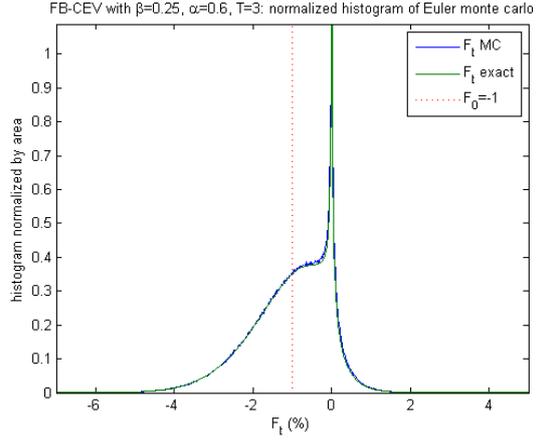
$$\sigma_N(K) = \frac{\alpha(F_0 - K)(1 - \beta)}{F_0/|F_0|^\beta - K/|K|^\beta} \cdot \frac{\zeta}{x} \cdot \left[1 + T \left(\frac{-\beta(2 - \beta)\alpha^2}{24|F_{av}|^{2-2\beta}} + \frac{\rho\nu\alpha\beta \text{sign}(F_{av})}{4|F_{av}|^{1-\beta}} \right) \right], \quad (80a)$$

where,

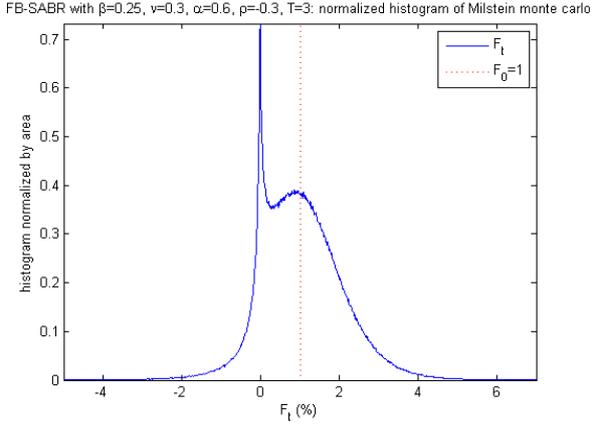
$$\zeta = \frac{\nu(f - K)}{\alpha|F_{av}|^\beta}, \quad x = \log \left(\frac{\sqrt{1 - 2\rho\zeta + \zeta^2} - \rho + \zeta}{1 - \rho} \right), \quad (80b)$$



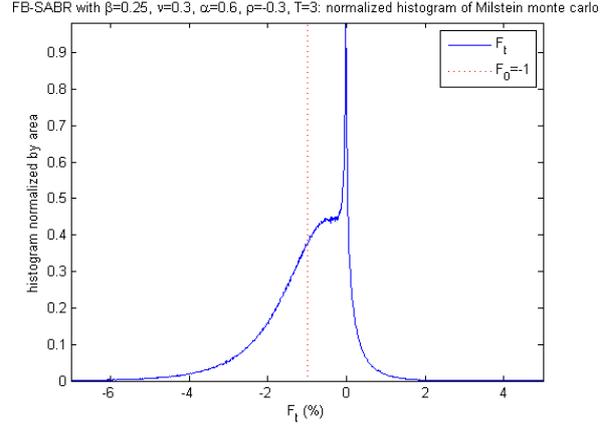
(a) FB-CEV with $F_0 = 1.0$



(b) FB-CEV with $F_0 = -1.0$



(c) FB-SABR with $F_0 = 1.0$, $\rho = -0.3$ and $\nu = 0.3$



(d) FB-SABR with $F_0 = -1.0$, $\rho = -0.3$ and $\nu = 0.3$

Figure 10: Estimated transition probability densities of the FB-CEV and FB-SABR models for parameters $\beta = 0.25$, $\alpha = 0.6$ and $T = 3$. The simulation scheme used $5 \cdot 10^6$ paths and 50 time steps per year.

and F_{av} is chosen as the generalized geometric mean

$$F_{av} = \text{sign}(F_0 K) \sqrt{|F_0 K|}. \quad (80c)$$

Figure 11 shows the probability density function implied from approximation 80 versus the approximation of the exact probability density function of FB-SABR by monte carlo simulations. The method from Hagan et al. seems accurate for the parameters in figure (a), except for a small area around zero. Figure (b) shows that the method is far from accurate for other parameters: its implied probability density is negative for a large area around zero, while this is the most relevant area in a low rate environment. The conclusion is that the method of Hagan et al. is an inappropriate approximation to FB-SABR and an inappropriate model of the volatility smile (in a low rate environment).

In conclusion, the FB-SABR is able to model negative rates, but the uncontrollable spike around zero makes it unreliable and unrealistic. The intuition that makes the original SABR model so popular, disappears in this extension. Furthermore, Hagan's formula for the FB-SABR model is inappropriate to calibrate the FB-SABR model, since its approximation is inaccurate in the low interest rate area.

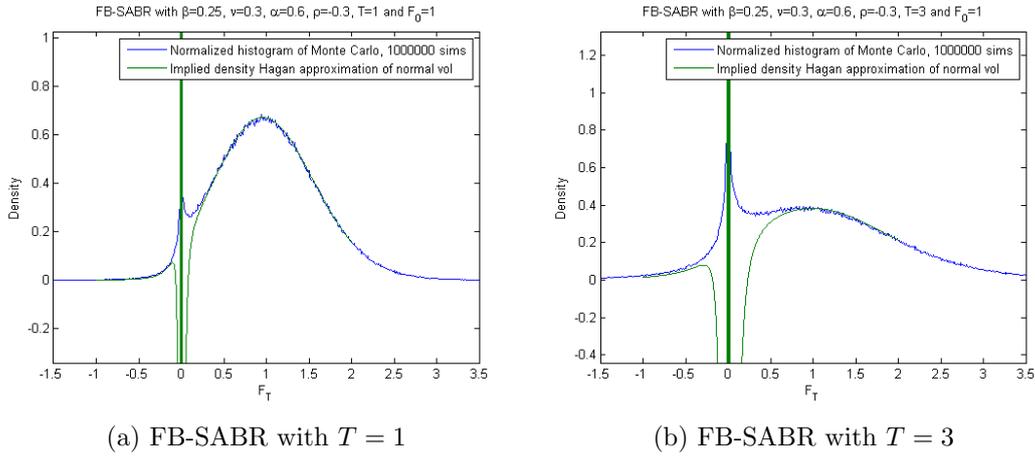


Figure 11: Comparison of the density implied by normal volatility 80 and the approximated density from monte carlo simulations.

3.4 Summary

Black's model and the SABR model both have the assumption of positive rates embedded in their structure. These models can be extended to model rates larger than $-s$ by shifting them with some shift parameter $s > 0$. Another way to extend SABR is to remove the boundary condition at zero, resulting in the free boundary SABR model. This extended model can model rates from the whole real line. Bachelier's model solves the problem of negative rates by using a normal distribution to model the underlying instead of a lognormal distribution as in Black's model. The main (dis)advantages of the models are listed in table 3.

Model class, domain	Pro +	Con -
Normal models $-\infty < F_t < \infty$	Model interpretation Analytical solution No extra parameter	Positive probability on large negative rates
Displaced models $-s < F_t < \infty$	Model interpretation Analytical solution	Additional shift parameter s chosen a priori
Free boundary models $-\infty < F_t < \infty$	No extra parameter	Model interpretation No analytical solution available (in general) Approximation inaccurate

Table 3: Benefits and drawbacks of models in a negative rates environment. Shift parameter s is a positive constant.

4 Pricing path dependent FX options

The initial research goal was to price barrier caps with the time dependent SABR model [15] in a negative interest rate environment. However, during the research it appeared that there is no obvious transformation of the time dependent FX-SABR model to the interest rate market. An important assumption of the time dependent FX-SABR model is deterministic discount curves (due to deterministic interest rates), while a realistic model for pricing interest rate derivatives requires stochastic interest rates. Methods from [10] or [21] are preferred to price path dependent interest rate options under a variation of the SABR model. The research goal was adjusted to the pricing and hedging of foreign exchange barrier options. Time dependent SABR has been proven to work well in this application area and ABN Amro was eager to gain knowledge about how this method performs compared to their own proprietary trading model. In the experiments in chapter 5 the market valuations of both methods will be compared.

The rest of this section explains why a stochastic volatility model, or SABR with time dependent parameters more specifically, is needed for the pricing of path dependent options. Section 4.1 introduces mathematical definitions and knowledge about the foreign exchange market such that the time dependent SABR model can be introduced in section 4.2. Certain mappings that are the core of the calibration procedure of this model are explained in section 4.3. The calibration procedure itself is described in section 4.4. Section 4.5 shows how the calibration performance is improved. Finally, section 4.6 shows how the time dependent SABR model is adjusted for a negative underlying asset.

Pricing path dependent options with local volatility or stochastic volatility models

The pricing and hedging of path dependent financial products requires an accurate calibration to prices of European-type options with different expiries that contain information about market behaviour through time. Accurate calibration to European-type options is not sufficient though, an appropriate model should also reflect market behaviour of implied volatility smile dynamics. The current value of a path dependent product depends on transition probabilities from one future state to another. Even though the local volatility model [17] can be calibrated perfectly to European-type option market data, it produces a flattening of the forward implied volatility smile. Products that are sensitive to the forward implied volatility smile, like path dependent options or forward starting options, could be priced incorrectly due to this property of the local volatility model. In contrast, stochastic volatility models predict that the forward implied volatility smile has a similar shape to the current observed implied volatility smile [5]. This usually leads to more accurate results in the pricing of such products. The dynamics of the implied volatility curve with respect to changes in the underlying, are very important for hedging. One of the main reasons for Hagan et al. [20] to introduce the SABR model is that the local volatility model produces inaccurate volatility smile moves after a change in the underlying. This in turn leads to unstable hedges. The SABR model in contrast, possesses implied volatility curve dynamics corresponding to market behaviour observed in the foreign exchange (FX) market [5].

The SABR model with constant parameters versus time dependent parameters

The SABR model introduced before has constant parameters. When Hagan's formula is used to calibrate to market data of several expiries, this usually leads to a different set of (constant) parameters for each expiry. This implies that the assumption of constant parameters needs to be relaxed in order to calibrate the model to data of multiple expiries accurately. The purpose of calibrating the SABR model to multiple expiries of European-type options could be to price and hedge an exotic derivative for which no market quote is available. The lack of flexibility of the SABR model with constant parameters is best illustrated with an example.

Suppose that the calibration of the constant SABR model to a market with two expiries, half a

year and one year, resulted in the effective parameters of table 4. Constant parameters and the constant SABR model will be called effective parameters and the effective SABR model from now on, the reason for this will be explained in section 4.3.

Expiry (Years)	γ	ρ	ω
0.5	1.0000	-0.09000	0.1000
1.0	0.9653	-0.08556	0.1101

Table 4: Market effective parameters

Figure 12a shows that when the effective SABR model is simulated with parameters $(\gamma, \rho, \omega) = (1.00, -0.09, 0.10)$ model implied volatilities match the market implied volatilities at the first expiry. In contrast, the model implied volatilities of the second expiry, shown in figure 12b, do not. The opposite is true when the model is simulated with $(\gamma, \rho, \omega) = (0.9653, -0.08556, 0.1101)$. The model implied volatilities of the first expiry do not match the market implied volatilities, but the model implied volatilities at the second expiry do, as can be seen in figures 12c and 12d respectively.

This example shows that the SABR model with constant parameters is not suited to calibrate to market data of multiple expiries. The model can be extended with time dependent parameters. For example, with piecewise constant parameters for example. A naive choice for the example above would be choosing time dependent parameters (81)

$$\gamma(t) := \begin{cases} 1.0000 & 0 \leq t \leq 0.5 \\ 0.9653 & 0.5 < t \leq 1 \end{cases}, \rho(t) := \begin{cases} -0.09000 & 0 \leq t \leq 0.5 \\ -0.08556 & 0.5 < t \leq 1 \end{cases} \text{ and } \omega(t) := \begin{cases} 0.1000 & 0 \leq t \leq 0.5 \\ 0.1101 & 0.5 < t \leq 1 \end{cases}. \quad (81)$$

This choice leads to an accurate calibration for the first expiry, as can be seen in figure 13a. In contrast, figure 13b shows that model and market implied volatilities do not match at the second expiry. Intuitively, piecewise constant parameters that are 'smartly averaged' from the market effective parameters might lead to a good fit. The method of finding those parameters is the core of the paper of Van der Stoep et al. [15]. It appears that when the time dependent parameters are chosen as in formula (82) the time dependent model fits well.

$$\gamma(t) := \begin{cases} 1.0 & 0 \leq t \leq 0.5 \\ 0.9 & 0.5 < t \leq 1 \end{cases}, \rho(t) := \begin{cases} -0.09 & 0 \leq t \leq 0.5 \\ -0.08 & 0.5 < t \leq 1 \end{cases} \text{ and } \omega(t) := \begin{cases} 0.10 & 0 \leq t \leq 0.5 \\ 0.12 & 0.5 < t \leq 1 \end{cases}, \quad (82)$$

This can be seen by comparing the market volatilities with those from the time dependent model in figure 12.

Clearly, a time dependent model is necessary for accurate calibration to different expiries at the same time. Before introducing the time dependent SABR model in section 4.2, the mathematical framework as well as details about the foreign exchange (FX) market are explained in section 4.1. These definitions need to be introduced before the time dependent SABR model and its calibration to the FX market can be explained.

4.1 FX products and mathematical framework

Let $r_d(t)$ and $r_f(t)$ be the deterministic domestic and foreign interest rates respectively. Furthermore, let B_d and B_f be the domestic and foreign bank accounts, determining the time-value of money, with dynamics

$$dB_d(t) = r_d(t)B_d(t)dt \text{ and } dB_f(t) = r_f(t)B_f(t)dt. \quad (83)$$

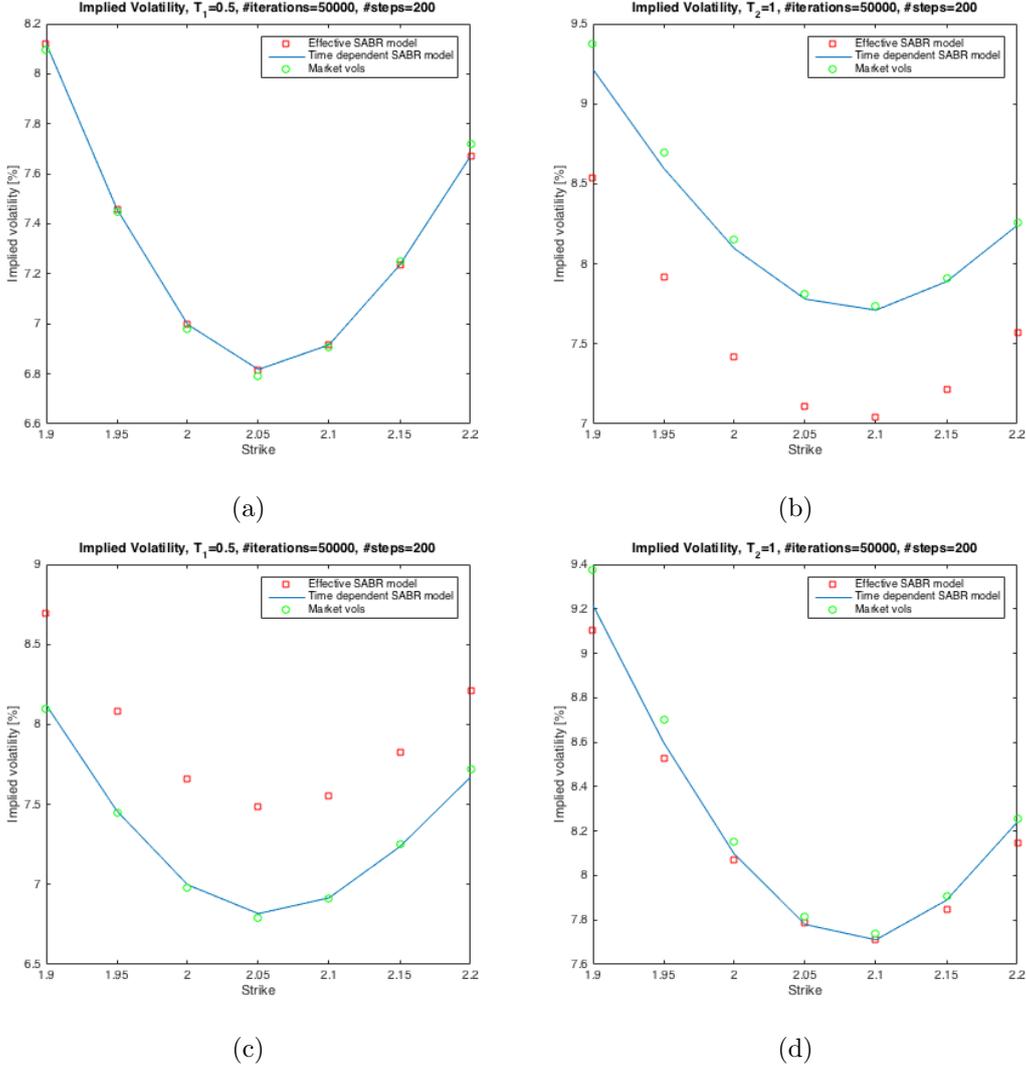


Figure 12: A comparison of the implied volatilities of simulated effective and time dependent SABR models with market implied volatilities. The effective parameters used in figures (a) and (b) are $(\gamma, \rho, \omega) = (1.00, -0.09, 0.10)$ and the effective parameters used in (c) and (d) are $(\gamma, \rho, \omega) = (0.9653, -0.08556, 0.1101)$. The time dependent parameters used in (a)-(d) are given in formula (82). The market implied volatilities are listed in table 4.

The domestic and foreign discount factors, or zero coupon bonds, are defined as

$$P_d(t, T) := \mathbb{E}^{\mathbb{S}} \left[\frac{B_d(t)}{B_d(T)} \mid \mathcal{F}_t \right] = e^{-\int_t^T r_d(s) ds} \quad \text{and} \quad P_f(t, T) := \mathbb{E}^{\mathbb{S}} \left[\frac{B_f(t)}{B_f(T)} \mid \mathcal{F}_t \right] = e^{-\int_t^T r_f(s) ds}, \quad (84)$$

where \mathbb{S} is the risk neutral (domestic) spot measure. The spot FX rate, expressed in units of domestic currency (DOM) per unit of foreign currency (FOR), is the current exchange rate at present time t and denoted by $S(t)$. The currency exchange rate is usually quoted as FOR-DOM. For example EUR-USD = 1.26 means that one EUR is worth 1.26 USD, the foreign currency is the EUR and the domestic currency is USD. By the interest rate parity, the relation between the spot FX $S(t)$ and forward FX $F^{(i)}(t)$ at time T_i is

$$F^{(i)}(t) = S(t) \frac{P_f(t, T_i)}{P_d(t, T_i)}. \quad (85)$$

The value $V(t)$ of an FX call option with expiry T_i prevailing at time $t \leq T_i$ with strike K is

$$\mathbb{E}^{\mathbb{F}^t} \left[\frac{P_d(t, T_i)}{P_d(T_i, T_i)} (F^{(i)}(T_i) - K)^+ \mid \mathcal{F}_t \right] = V(t) = \mathbb{E}^{\mathbb{S}} \left[\frac{B_d(t)}{B_d(T_i)} (S(T_i) - K)^+ \mid \mathcal{F}_t \right], \quad (86)$$

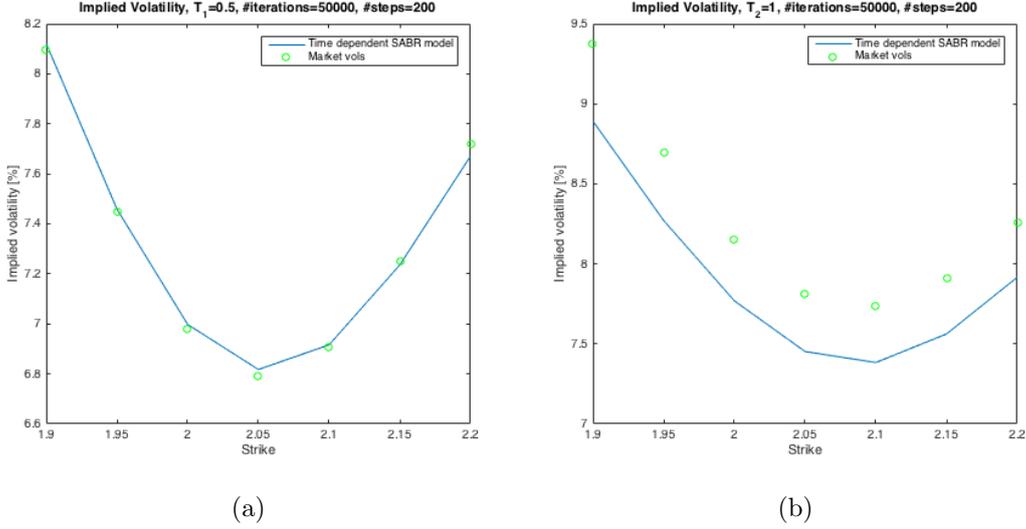


Figure 13: A comparison of the market implied volatilities belonging to market effective parameters listed in table 4 with implied volatilities obtained by simulating time dependent SABR with parameters in formula (81).

where \mathbb{F}^i is the risk neutral forward measure at time T_i . The Radon-Nikodym derivative [31] between both measures is equal to one:

$$\frac{d\mathbb{F}^i}{d\mathbb{S}} = \frac{P_d(T_i, T_i)B_d(t)}{B_d(T_i)P_d(t, T_i)} = 1, \quad (87)$$

due to deterministic rates $r_d(t)$ and $r_f(t)$. This implies $dW_{\mathbb{F}^i}(t) = dW_{\mathbb{S}}(t)$ and this knowledge is useful in the derivation of the time dependent SABR model in section 4.2. The goal of introducing these (mathematical) definitions is to price barrier options with the time dependent SABR model. Four different types of barrier FX options are priced in the experiments chapter 5: up-and-out (OU) calls (C) and puts (P), and down-and-out (DO) calls and puts. Suppose the strike is K , the expiry is T and the barrier is B , then the time t value $V(t)$ of these barrier options is:

$$V_{UO}^C(t) = \mathbb{E}^{\mathbb{S}} \left[\frac{B_d(t)}{B_d(T_i)} (S(T) - K)^+ \mathbf{1} \left(\max_{t \in [0, T]} S(t) < B \right) \mid \mathcal{F}_t \right], \quad (88a)$$

$$V_{UO}^P(t) = \mathbb{E}^{\mathbb{S}} \left[\frac{B_d(t)}{B_d(T_i)} (K - S(T))^+ \mathbf{1} \left(\max_{t \in [0, T]} S(t) < B \right) \mid \mathcal{F}_t \right], \quad (88b)$$

$$V_{DO}^C(t) = \mathbb{E}^{\mathbb{S}} \left[\frac{B_d(t)}{B_d(T_i)} (S(T) - K)^+ \mathbf{1} \left(\min_{t \in [0, T]} S(t) > B \right) \mid \mathcal{F}_t \right], \text{ and} \quad (88c)$$

$$V_{DO}^P(t) = \mathbb{E}^{\mathbb{S}} \left[\frac{B_d(t)}{B_d(T_i)} (K - S(T))^+ \mathbf{1} \left(\min_{t \in [0, T]} S(t) > B \right) \mid \mathcal{F}_t \right], \quad (88d)$$

respectively. These barrier options are valued by calibrating to market data later, the results are listed in the experiments chapter 5. Before the time dependent SABR model is introduced in section 4.2, FX market conventions and the construction of the implied volatility smile are explained. This knowledge is necessary for the calibration of the model to market data.

FX market quote conventions and the implied volatility smile construction

A transaction in the Foreign Exchange (FX) market is represented by an amount in foreign currency N_f and an amount in domestic currency N_d and a payment date T . Notionals N_f and N_d have different signs, where the sign depends on which side of the transaction the participant takes. The option on the FX transaction is represented by the underlying FX transaction, a strike K , expiry date T , and a call or put feature. The holder of an FX option has the right

(but not the obligation) to exchange a specific amount of money in foreign currency at an agreed exchange rate K at maturity time T . It is market practice to value the FX call and put options with Black's model, also called Garman-Kohlhagen in this context:

$$V^C(0) = P_f(0, T_i)S_0\Phi(d_1) - P_d(0, T_i)K\Phi(d_2), \quad (89a)$$

$$= P_d(0, T_i) \left(F_0^{(i)}\Phi(d_1) - K\Phi(d_2) \right), \quad (89b)$$

$$d_1 := \frac{\log\left(\frac{F_0^{(i)}}{K}\right)}{\sigma\sqrt{T_i}} + \frac{1}{2}\sigma\sqrt{T_i}, \text{ and } d_2 := d_1 - \sigma\sqrt{T_i}. \quad (89c)$$

Note that the value of the option is in domestic currency. The option position, however, may also be held in foreign currency. The option value formula applies to one unit of foreign notional, with the option value in units of domestic currency.

Deltas in the FX market can be calculated in several ways. It can be calculated with respect to the spot FX rate S_t or with respect to a forward FX rate $F^{(i)}(t)$. Furthermore, the standard 'unadjusted' delta is a quantity in percent of foreign currency. The actual hedge quantity must be changed if the premium is paid in foreign currency. This would be equivalent to paying stock options with stock shares. This type of delta is called the premium-adjusted delta. The markets considered in the experiments of this thesis are the EUR-USD and the GBP-USD markets with expiries up till one year (1Y). In these markets, the standard/unadjusted spot delta is used by default [29]. Therefore, the following analysis is limited to the standard spot delta:

$$\Delta_S^C = P_f(0, T_i)\Phi(d_1) \text{ and } \Delta_S^P = -P_f(0, T_i)\Phi(-d_1). \quad (90)$$

In the FX market the market prices or implied volatilities of call and put options are not quoted directly. The quotes of a particular expiry usually consist of an 'at-the-money' (ATM) implied volatility, and for one or more deltas a so called risk reversal (RR) and a straddle (STR). A different definition of ATM implied volatility is used than in previous sections. Before, the ATM volatility was the implied volatility where the strike K equals current underlying forward F_0 . In this situation, the ATM volatility is the volatility of a zero delta straddle, also called a risk neutral straddle. A straddle is a put and a call option with the same strike at a certain expiration. In this particular straddle the put and call have the same strike and the same delta (but the opposite sign). To make sure the call and put deltas are equal in absolute value

$$P_f(0, T_i)\Phi\left(\frac{\log\frac{F_0^{(i)}}{K_{\delta 0}^{(i)}} + \frac{1}{2}\sigma_{\delta 0}T_i}{\sigma_{\delta 0}\sqrt{T_i}}\right) = P_f(0, T_i)\Phi\left(-\frac{\log\frac{F_0^{(i)}}{K_{\delta 0}^{(i)}} + \frac{1}{2}\sigma_{\delta 0}T_i}{\sigma_{\delta 0}\sqrt{T_i}}\right), \quad (91)$$

the common strike $K_{\delta 0}$ has to be

$$K_{\delta 0}^{(i)} = F_0^{(i)} e^{\frac{1}{2}\sigma_{\delta 0}T_i}. \quad (92)$$

This means the zero delta strike does not equal the current forward, unless the volatility is zero. The (25-delta) risk reversal volatility

$$RR_{25} := \sigma_{C,\delta 25} - \sigma_{P,\delta 25}, \quad (93)$$

measures the skewness of the smile and the (25-delta) strangle volatility

$$STR_{25} := \frac{1}{2}(\sigma_{C,\delta 25} + \sigma_{P,\delta 25}) - \sigma_{\delta 0}, \quad (94)$$

measures curvature of the smile. The risk reversal and the strangle of different deltas, are calculated in the same way. Given the ATM volatility, the risk reversal(s) and the strangle(s), the market implied volatilities (at 25-delta) can be calculated with formulas

$$\sigma_{C,\delta 25} = \sigma_{\delta 0} + STR_{25} + RR_{25}, \quad (95a)$$

$$\sigma_{P,\delta 25} = \sigma_{\delta 0} + STR_{25} - RR_{25}. \quad (95b)$$

Volatility $\sigma_{C,\delta 25}$ is the implied volatility of a so called 25-delta call option. This is a call option with a strike chosen such that its delta is 0.25. A 25-delta put option is a put option with a strike chosen such that the delta is -0.25 . The strikes corresponding to these 25-delta (and 10-delta) put and call volatilities are not known explicitly. They are solved by using definition (91) of the spot delta:

$$K_{C,\delta 25}^{(i)} = F_0^{(i)} \exp \left(-\sigma_{C,\delta 25} \sqrt{T_i} \Phi^{-1}(0.25) + \frac{1}{2}(\sigma_{C,\delta 25})^2 T_i \right), \quad (96a)$$

$$K_{P,\delta 25}^{(i)} = F_0^{(i)} \exp \left(\sigma_{P,\delta 25} \sqrt{T_i} \Phi^{-1}(-(-0.25)) + \frac{1}{2}(\sigma_{P,\delta 25})^2 T_i \right). \quad (96b)$$

In the GBP-USD and EUR-USD markets, quotes of 25-delta as well as 10-delta are available. For each expiry, a five point volatility smile can be constructed with the techniques introduced here. Hagan's formula can be calibrated to these points to obtain a smooth volatility smile.

Now that FX products are introduced and the volatility smile at each expiry is constructed, the time dependent FX-SABR model that can be used to price exotic FX products is introduced in the next section.

4.2 Time dependent FX-SABR model

The classic SABR model can be extended with time dependent parameters to be able to calibrate to market data of several expiries. The following *time dependent FX-SABR model* in particular, is capable of calibration to FX options of different expiries:

$$dS(t) = S(t)(r_d(t) - r_f(t))dt + \left(\frac{P_d(t, T_N)}{P_f(t, T_N)} \right)^{1-\beta} \omega(t)\sigma(t)[S(t)]^\beta dW_{\mathbb{S}}(t), \quad (97a)$$

$$d\sigma(t) = \gamma(t)\sigma(t)dW_{\sigma}(t), \sigma(0) = 1, \quad (97b)$$

$$dW_{\mathbb{S}}(t)dW_{\sigma}(t) = \rho(t)dt. \quad (97c)$$

The model is introduced in Van der Stoep, Grzelak and Oosterlee [15].

Derivation

By the interest rate parity (85) the differential of $F^{(i)}(t)$ is

$$dF^{(i)}(t) = \frac{P_f(t, T_i)}{P_d(t, T_i)} dS(t) + (r_f(t) - r_d(t))S(t) \frac{P_f(t, T_i)}{P_d(t, T_i)} dt, \text{ or} \quad (98a)$$

$$dS(t) = \frac{P_d(t, T_i)}{P_f(t, T_i)} dF^{(i)}(t) + (r_d(t) - r_f(t))S(t)dt. \quad (98b)$$

In order to use Hagan's formula for calibration to market data, the forward rate should be of the form

$$dF^{(i)}(t) = \omega(t)\sigma(t)C(\cdot)[F^{(i)}(t)]^\beta dW_{\mathbb{F}^i}(t), \quad (98c)$$

where $C(\cdot)$ is some deterministic function. Function C should be chosen such that there is no dependency on a particular expiry date in the dynamics of the spot. Trivial candidate

$$C_1 := 1, \quad (98d)$$

leaves a dependency on expiry date T_i in the spot dynamics. Candidate function

$$C_2(t, i) := \left(\frac{P_f(t, T_i)}{P_d(t, T_i)} \right)^{1-\beta}, \quad (98e)$$

would eliminate this dependency on T_i in the spot dynamics, but it introduces a dependency on time t in the dynamics of the forward, which is undesirable since this makes using Hagan's formula for calibration impossible. Also candidate

$$C_3(t) := \left(\frac{P_f(T_i, T_N)}{P_d(T_i, T_N)} \right)^{1-\beta}, \quad (98f)$$

would eliminate the dependency on i of the spot dynamics. Furthermore, this candidate does not introduce an extra time dependency in the forward dynamics, enabling it to be calibrated with Hagan's formula. In this case the forward dynamics are

$$dF^{(i)}(t) = \omega(t)\sigma(t) \left(\frac{P_d(T_i, T_N)}{P_f(T_i, T_N)} \right)^{1-\beta} [F^{(i)}(t)]^\beta dW_{\mathbb{F}^i}(t). \quad (98g)$$

Normalizing $F^{(i)}(t)/f_0^{(i)}$, where $f_0^{(i)} := F^{(i)}(0)$ and $s_0 := S(0)$, leads to the model

$$dF^{(i)}(t) = \omega_1(t)\sigma(t)[F^{(i)}(t)]^\beta dW_{\mathbb{F}^i}(t), F^{(i)}(0) = 1, \quad (99a)$$

$$d\sigma(t) = \gamma(t)\sigma(t)dW_\sigma(t), \sigma(0) = 1, \quad (99b)$$

$$dW_{\mathbb{F}^i}(t)dW_\sigma(t) = \rho(t)dt, \quad (99c)$$

where

$$\omega_1(t) = \omega(t) \left(\frac{P_d(T_i, T_N)}{f_0^{(i)} P_f(T_i, T_N)} \right)^{1-\beta} = \omega(t) \left(\frac{P_d(0, T_N)}{s_0 P_f(0, T_N)} \right)^{1-\beta}. \quad (99d)$$

This model is in SABR form when ω , ρ and γ are chosen to be constant. Plugging (98g) in (98b) leads to the following spot dynamics of the time dependent SABR model:

$$dS(t) = S(t)(r_d(t) - r_f(t))dt + \frac{P_d(t, T_i)}{P_f(t, T_i)} \omega(t)\sigma(t) \left(\frac{P_d(T_i, T_N)}{P_f(T_i, T_N)} \right)^{1-\beta} \left(S(t) \frac{P_f(t, T_i)}{P_d(t, T_i)} \right)^\beta dW_{\mathbb{S}}(t), \quad (100a)$$

$$= S(t)(r_d(t) - r_f(t))dt + \left(\frac{P_d(t, T_N)}{P_f(t, T_N)} \right)^{1-\beta} \omega(t)\sigma(t) (S(t))^\beta dW_{\mathbb{S}}(t). \quad (100b)$$

The drift term of the resulting spot dynamics (97) is expressed in time dependent domestic and foreign interest rates $r_d(t)$ and $r_f(t)$. These rates can be bootstrapped from the domestic and foreign discount curves P_d and P_f , which are observable in the market.

Bootstrapping the rate function

The underlying model governing discount factors P_d as well as P_f in the time dependent SABR model is

$$P(s, t) = e^{-\int_s^t r(u)du}, \quad (101)$$

for some function $r(\cdot)$.

Suppose discount factors from t_0 till t_1, t_2, \dots, t_N are known: $P(t_0, t_1), P(t_0, t_2), \dots, P(t_0, t_N)$. Furthermore, assume that r is a piecewise constant function:

$$r(t) := \begin{cases} r_1 & t_0 \leq t \leq t_1 \\ r_2 & t_1 < t \leq t_2 \\ \vdots & \vdots \\ r_N & t_{N-1} < t \leq t_N \end{cases}. \quad (102)$$

The piecewise constant function r can then be extracted from the discount factors with the following formula:

$$r_j = -\frac{1}{t_j - t_{j-1}} \log \left(\frac{P(t_0, t_j)}{P(t_0, t_{j-1})} \right), \text{ for } j = 1, \dots, N, \quad (103)$$

since $P(t_{j-1}, t_j) = P(t_0, t_j)/P(t_0, t_{j-1})$.

The above procedure is used to construct a time dependent discount factor from discount factors observed in the market. Together with the constructed implied volatility smiles in section 4.1 and the introduced time dependent SABR model (97) there is almost enough information available to calibrate this model. Before the calibration procedure is explained in section 4.4, the mappings that are the core of this procedure are explained in section 4.3.

4.3 Mappings of time dependent to effective parameters

The calibration of the time dependent FX-SABR model is based on the different effects that the parameters have on the implied volatility smile. The calibration makes use of so called *effective parameters*, which are 'smart averages' of the time dependent parameters. This technique was introduced by Piterbarg [28]. The *effective FX-SABR model* under the T_i -forward measure \mathbb{F}^i , $i = 1, \dots, N$, is defined as

$$d\tilde{F}^{(i)}(t) = \tilde{\omega}_1 \tilde{\sigma}(t) [\tilde{F}^{(i)}(t)]^\beta dW_{\mathbb{F}^i}(t), \tilde{F}^{(i)}(0) = 1, \quad (104a)$$

$$d\tilde{\sigma}(t) = \tilde{\gamma} \tilde{\sigma}(t) dW_\sigma(t), \tilde{\sigma}(0) = 1, \quad (104b)$$

$$dW_{\mathbb{F}^i}(t) dW_\sigma(t) = \tilde{\rho} dt, \quad (104c)$$

where

$$\tilde{\omega}_1 = \tilde{\omega} \left(\frac{P_d(T_i, T_N)}{\tilde{f}_0^{(i)} P_f(T_i, T_N)} \right)^{1-\beta}. \quad (104d)$$

It is possible to calculate effective parameters such that the implied volatilities produced by the effective (104) and time dependent (99) models are (approximately) equal at time T_i . In the next paragraphs it is described how these effective parameters can be calculated. First, the effect of each parameter on the volatility smile is described.

The standard parameter setting in figure 14 is $F_0^{(i)} = 1.5$, $\beta = 0.5$, $\tilde{\gamma} = 0.4$, $\tilde{\omega}_1 = 0.14$, $\tilde{\rho} = 0$ and $T_i = 1$. In each subfigure a parameter is varied, as can be seen in the corresponding legend.

The volatility-of-volatility (volvol) parameter $\tilde{\gamma}$ in (104) influences curvature of the smile mainly. To a lesser extent, $\tilde{\gamma}$ influences the level of the volatility smile as well, as can be seen in figure 14a. It also has an almost negligible effect on the skew.

Term structure parameter $\tilde{\omega}$ mainly influences the level of the smile, as can be seen in figure 14b. Its effects on the curvature and skew are negligible. According to Hagan's formula (38), the ATM volatility is approximately equal to $\tilde{\omega}/(F_0^{(i)})^{1-\beta}$. Since $F_0^{(i)} = 1$ due to the normalisation in (99) and (104), the ATM volatility is approximately equal to the term structure parameter. This was actually one of the main reasons for using normalisation in the time dependent and effective models: to separate the different effects of different parameters as much as possible.

The correlation parameter $\tilde{\rho}$ has two effects, like the volvol parameter $\tilde{\gamma}$. It mainly influences the skew, as can be seen in figure 14c. It does not influence the level, but it does influence the

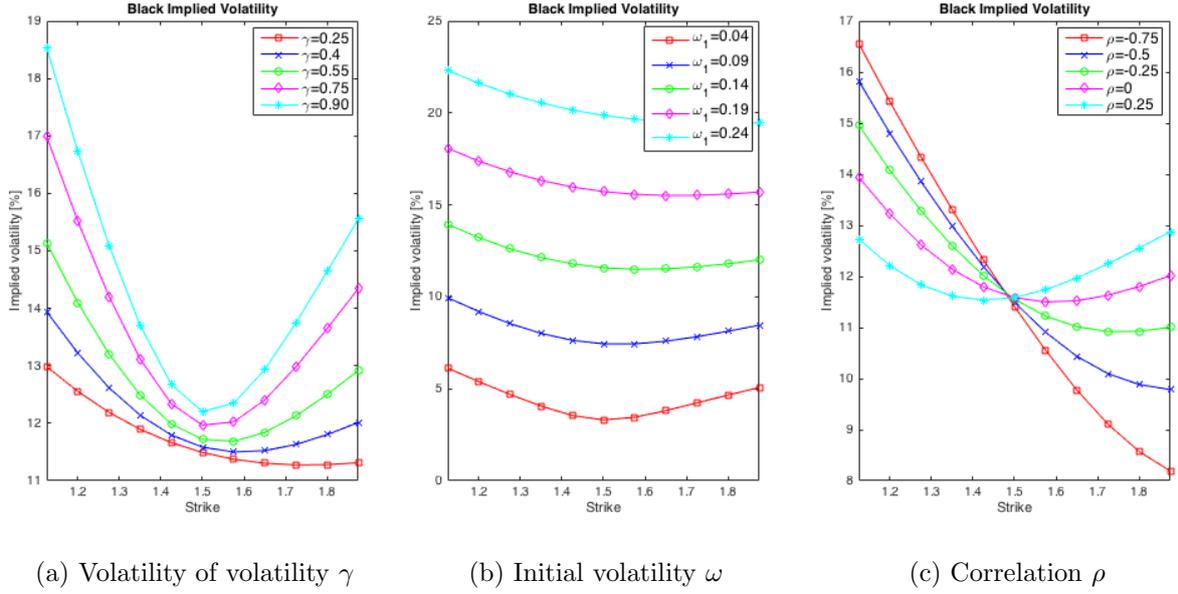


Figure 14: Figure (a), (b) and (c) illustrate the effect of the volatility-of-volatility $\tilde{\gamma}$, initial volatility $\tilde{\omega}$ and correlation $\tilde{\rho}$ parameters on the implied volatility smile respectively. The standard parameter setting is $F_0^{(i)} = 1.5$, $\beta = 0.5$, $\tilde{\gamma} = 0.4$, $\tilde{\omega}_1 = 0.14$, $\tilde{\rho} = 0$ and $T_i = 1$. In each of the figures one parameter is varied.

curvature. Hagan et al. [20] decompose the skew in *beta-skew* $-\frac{1}{2}(1 - \beta - \tilde{\rho}\frac{\tilde{\gamma}}{\tilde{\omega}_1})$ and *vanna-skew* $\frac{1}{2}\tilde{\rho}\frac{\tilde{\gamma}}{\tilde{\omega}_1}\log(K)$. The first is responsible for the skew caused by parameter β and the latter for the skew caused by correlation $\tilde{\rho}$. The calibration procedure of the skew, that is introduced later, is based on the calibration of vanna skew. The reason for this is that there are three degrees of freedom: level, skew and curvature, while there are four parameters. This means that after fixing one parameter there is still enough flexibility to model a wide variety of smiles. It is industry practice to fix β , since its effects on the smile can be easily replaced by the other parameters. The influence of β for a current forward smaller than one and larger than one can be seen in figures 15a and 15c respectively. The normalisation in (99) and (104) reduced the impact of the choice of β as can be seen in figure 15b.

The effective model (104) is indirectly used for the calibration of the time dependent model (97) via the calibration of time dependent parameters $\rho(t)$, $\gamma(t)$, and $\omega(t)$. The calculation methods of these time dependent parameters are described in the next paragraphs.

Volvol mapping

The mapping from a time dependent volvol parameter $\gamma(t)$ to its corresponding effective volvol parameter $\tilde{\gamma}$ is defined as

$$\int_0^{T_i} \omega_1^2(t) \left(\int_0^t \omega_1^2(s) e^{6 \int_0^s \gamma^2(u) du} + \int_s^t \gamma^2(u) du ds \right) dt \quad (105a)$$

$$= \frac{1}{5} \left(\frac{\int_0^{T_i} \omega_1^2(t) e^{\int_0^t \gamma^2(u) du} dt}{e^{\tilde{\gamma}^2 T_i} - 1} \right)^2 \left(\frac{1}{6} e^{6\tilde{\gamma}^2 T_i} - e^{\tilde{\gamma}^2 T_i} + \frac{5}{6} \right). \quad (105b)$$

The mapping provides an implicit relationship between the time dependent and effective volvol. It can be numerically solved for the effective parameter. Note that the mapping does not depend on the effective level $\tilde{\omega}_1$, in contrast to the time dependent level $\omega_1(t)$. It cancelled out during the derivation. The details of the derivation can be found in the proof of lemma 4.1 in Van der Stoep et al. [15, page 9-10]. Piecewise constant parameters yield analytic expressions for

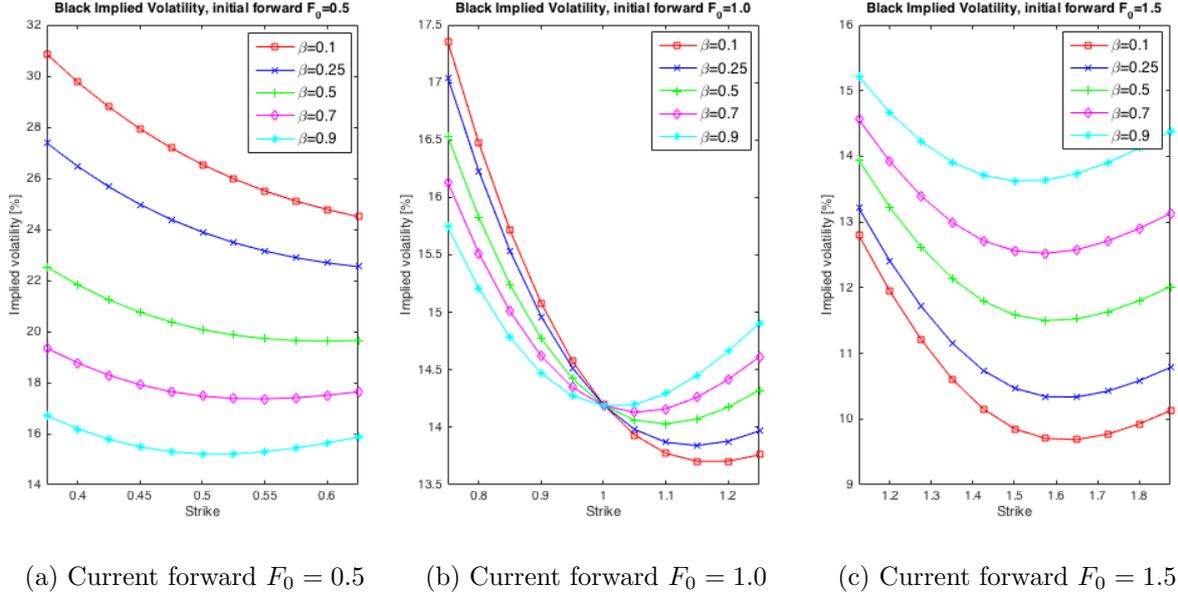


Figure 15: Figure (a), (b), and (c) illustrate the effect of the power parameter β on the shape of the volatility smile with a forward smaller than 1, equal to 1 and larger to 1 respectively. The standard parameter setting is $\tilde{\gamma} = 0.4$, $\tilde{\omega}_1 = 0.14$, $\tilde{\rho} = 0$ and $T_i = 1$. In each of the figures β is varied. The current forward equals 0.5 in (a) and 1.5 in (b).

the integrals in (105). The analytic expressions and their derivations can be found in appendix E.

The main influence of the volvol parameter on the volatility smile is on its curvature. Curvature can be measured in so called realised volatility. The idea behind the derivation is that the effective volvol parameter is determined such that the realised volatilities of the effective and time dependent models are equal in distribution:

$$\int_0^{T_i} \omega_1(t) \sigma(t) dW^{T_i}(t) \stackrel{d}{=} \int_0^{T_i} \tilde{\omega}_1(t) \tilde{\sigma}(t) dW^{T_i}(t). \quad (106)$$

The result in formula (105) is achieved by matching moments of the realised volatilities in (106). Since the first moment of the left and right hand sides are both zero, this provides little information and the second moment is matched instead.

Correlation mapping

Correlation ρ and power parameter β affect the skew similarly, so skew can be calibrated with ρ when β is fixed. Vanna skew in the SABR model is the part of the skew caused by ρ , the other part of the skew is mainly caused by β . In the context of the (effective) SABR model, vanna skew is defined by Hagan et al. [20] as

$$V(\tilde{\rho}) := \frac{1}{2} \tilde{\rho} \tilde{\omega} (f_0^{(i)})^{1-\beta} \log \left(\frac{K}{f_0^{(i)}} \right). \quad (107)$$

The effective correlation corresponding to expiry T_i is determined by matching the vanna skew of the effective SABR model with the average vanna skew of the time dependent SABR model over $[0, T_i]$:

$$\tilde{\rho} = \arg \min_x \left(\frac{x \tilde{\gamma}}{2 \tilde{\omega}} (f_0^{(i)})^{1-\beta} \log \left(\frac{K}{f_0^{(i)}} \right) - \frac{1}{T_i} \int_0^{T_i} \frac{\rho(t) \gamma(t)}{2 \omega(t)} (f_0^{(i)})^{1-\beta} \log \left(\frac{K}{f_0^{(i)}} \right) dt \right). \quad (108a)$$

The solution of above equation is

$$\tilde{\rho} = \frac{\tilde{\omega}}{\tilde{\gamma}T_i} \int_0^{T_i} \frac{\rho(t)\gamma(t)}{\omega(t)} dt. \quad (108b)$$

For piecewise constant parameters this reduces to

$$\tilde{\rho} = \frac{\tilde{\omega}}{\tilde{\gamma}T_i} \sum_{j=1}^i \frac{\rho_j \gamma_j}{\omega_j} (T_j - T_{j-1}). \quad (108c)$$

This formula will be referred to as the correlation mapping.

Term structure mapping

Term structure parameter ω mainly affects the level of the implied volatility smile. For this reason, effective parameter $\tilde{\omega}$ is derived by matching volatility smile levels of the effective and time dependent models. Furthermore, the level of the smile is completely determined by the value of the at-the-money (ATM) implied volatility. Therefore, the ATM price values of the time dependent (99) and effective models are matched

$$\mathbb{E} \left[(F^{(j)}(T_j) - 1)^+ \mid \mathcal{F}_0 \right] = \mathbb{E} \left[(\tilde{F}^{(j)}(T_j) - 1)^+ \mid \mathcal{F}_0 \right], \quad (109)$$

in order to calculate $\tilde{\omega}$. The expected ATM payoff of time dependent model (99) can be approximated by

$$\mathbb{E} \left[(F^{(j)}(T_j) - 1)^+ \mid \mathcal{F}_0 \right] = \frac{1}{\sqrt{2\pi}} \left(\omega_1(0) \sqrt{\Delta_t} \hat{\phi}_{Y_M} \left(-\frac{1}{2}i \right) \right), \quad (110)$$

where $\hat{\phi}_{Y_M}$ is an approximation of the characteristic function of $Y_M := \log \left(\sum_{k=1}^M \frac{\omega_1(t_k) \sigma^2(t_k)}{\omega_1(0)} \right)$ with $\Delta_t = T_j/M$ and $t_k = k\Delta_t$, for $k = 1, \dots, M$. The expected ATM payoff of the corresponding effective model is approximated in the same way with the approximated characteristic function of $\tilde{Y}_M := \log \left(\sum_{k=1}^M \tilde{\sigma}^2(t_k) \right)$. These payoffs should be equal by definition, leading to an expression for effective term structure $\tilde{\omega}$ in terms of time dependent term structure $\omega(t)$:

$$\tilde{\omega} = \tilde{\omega}_1 \left(\frac{f_0^{(i-1)} P_f(T_i, T_N)}{P_d(T_i, T_N)} \right)^{1-\beta}, \quad \text{with } \tilde{\omega}_1 = \frac{\omega_1(0) \hat{\phi}_{Y_M} \left(-\frac{1}{2}i \right)}{\hat{\phi}_{\tilde{Y}_M} \left(-\frac{1}{2}i \right)}. \quad (111)$$

Formula (111) is defined as the term structure mapping. The details of the derivation of (111), as well as an analysis of the approximation errors, can be found in Van der Stoep et al. [15, page 11-17]. The recovery procedure in appendix F can be used to calculate the characteristic function approximations.

4.4 Calibration procedure

The mappings from time dependent parameters to effective parameters introduced in the previous section are used to calibrate the time dependent SABR model (97). The calibration procedure is described in this section. First, Hagan's formula is used to calibrate a smooth volatility smile to the implied volatility market data of each expiry. The SABR parameters that produce such a smile are called the *market effective parameters* (of that particular expiry). In general, each expiry has different market effective parameters. In order to calibrate the time dependent SABR model the mappings

$$\tilde{\gamma}^{mod} = \Gamma(\gamma(t), \omega(t)), \quad \tilde{\omega}^{mod} = \Omega(\gamma(t), \omega(t); \tilde{\gamma}^{mod}) \quad \text{and} \quad \tilde{\rho}^{mod} = R(\gamma(t), \omega(t), \rho(t); \tilde{\gamma}^{mod}, \tilde{\omega}^{mod}), \quad (112)$$

are used. Mappings Γ , Ω and R represent formula (105), (111) and (108c) respectively. These functions map time dependent parameters $\gamma(t), \rho(t)$ and $\omega(t)$ to their model effective equivalents $\tilde{\gamma}^{mod}, \tilde{\rho}^{mod}$ and $\tilde{\omega}^{mod}$. They are denoted with superscript 'mod' to distinguish these *model effective parameters* from the market effective parameters.

Every mapping depends on several time dependent parameters, not just 'its own' time dependent parameter. The Γ mapping for instance, has $\gamma(t)$ as well as $\omega(t)$ as input variables. It is possible to use all three mappings at the same time in one optimization procedure. However, this is too computationally expensive due to the high dimensionality of the problem. Especially the evaluation of the Ω mapping is computationally expensive, due to the recovery procedure of the characteristic functions. Furthermore, generating a 'good' initial guess is computationally expensive since each three dimensional candidate function needs to be checked for accuracy. Another possibility is to alternate the different mappings in different stages and to perform an optimization procedure in each stage. Since $\gamma(t)$ and $\rho(t)$ both have curvature and skew effects these are calibrated in the same stage. An error function compares model implied volatilities with market implied volatilities, instead of model and market effective parameters, since we deal with two of those in this stage. Furthermore, the error function is extended with penalties for some parameter combinations: it is necessary that $\gamma^{mod} \geq 0$ and $\rho^{mod} \in [-1, 1]$ for example. The Ω mapping is used to calibrate the level of the time dependent SABR model. The error function in an ω stage should guarantee that $\tilde{\omega}^{mod} \geq 0$.

There is a chicken-or-the-egg problem in this method: the γ - ρ stage needs $\omega(t)$ as an input and the ω stage needs $\gamma(t)$ as an input. A solution to this problem is to use the market effective parameter as a proxy for the time dependent parameter in the first stage. If the first stage is a γ - ρ stage the following approximations are used:

$$\tilde{\gamma}^{mod} = \Gamma(\gamma(t), \omega(t)) \approx \Gamma(\gamma(t), \tilde{\omega}^{mar}), \quad (113a)$$

$$\tilde{\rho}^{mod} = R(\gamma(t), \omega(t), \rho(t); \tilde{\gamma}^{mod}, \tilde{\omega}^{mod}) \approx R(\gamma(t), \tilde{\omega}^{mar}, \rho(t); \tilde{\gamma}^{mod}). \quad (113b)$$

In the experiments the Levenberg-Marquardt algorithm (LMA) is used as the error minimization algorithm. It is used in both the calibration of market volatilities to Hagan's formula as well as in the calibration stages described above. The LMA solves nonlinear least squares problems. It finds a local minimum, not necessarily a global minimum. For this reason the first stage starts with a (random) global search for a 'good' initial guess. Subsequently this initial guess is used in a local search with LMA. Since no boundary conditions can be defined in the LMA, the boundary conditions should be included in the error function in the form of penalties. If the best candidate of the initial guess does not result in a satisfactory solution, the second best candidate of the global search process is taken as the initial guess and the local search process is repeated, and so forth. In summary, the calibration procedure consists of the following stages:

1. Calibrate smooth volatility smiles to market data with Hagan's formula and obtain market effective parameters for each expiry.
2. Calibrate curvature and skew simultaneously in the γ - ρ stage with (113). Market effective term structure $\tilde{\omega}^{mar}$ is used as a proxy for $\omega(t)$.
3. Calibrate level with term structure mapping Ω (111) in the ω stage.
4. Calibrate curvature and skew again, to compensate for the first γ - ρ stage where a proxy was used as input.

When the calibration accuracy is not satisfactory after these iterations, more ω and γ - ρ iterations should be performed. This was not necessary in experiments. In summary, this section provides an algorithm to calibrate the time dependent SABR model by using the level, skew and curvature properties of a volatility smile. The following section describes a method that guarantees that the calibration is perfect.

4.5 Improving calibration accuracy by adding a local volatility component

The calibration of the time dependent SABR model (97) may not be perfect for all expiries. Remaining calibration errors are eliminated by adding a local volatility component [15]. The local volatility component should provide a perfect fit in theory. In practice, there may still be a small error left after adding the local volatility component, due to rounding errors, randomness of the Monte Carlo simulation and interpolation choices in the construction of the implied volatility surface. The $S(t)$ dynamics with local volatility component σ_{SLV} are described by

$$dS(t) = S(t)(r_d(t) - r_f(t))dt + \sigma_{SLV}(t, S(t))\omega(t)\sigma(t) \left(\frac{P_d(t, T_N)}{P_f(t, T_N)} \right)^{1-\beta} [S(t)]^\beta dW_{\mathbb{S}}(t). \quad (114)$$

Define ψ as

$$\psi(S(t), \sigma(t)) := \omega(t)\sigma(t) \left(\frac{P_d(t, T_N)}{P_f(t, T_N)} \right)^{1-\beta} [S(t)]^{\beta-1}, \quad (115a)$$

and above spot dynamics are rewritten as

$$dS(t) = (r_d(t) - r_f(t))S(t)dt + \sigma_{SLV}(t, S(t))\psi(S(t), \sigma(t))S(t)dW_{\mathbb{S}}(t), \quad (115b)$$

$$d\sigma(t) = \gamma(t)\sigma(t)dW_{\sigma}(t), \sigma(0) = 1, \quad (115c)$$

$$dW_{\mathbb{S}}(t)dW_{\sigma}(t) = \rho(t)dt. \quad (115d)$$

System (115) has an explicit expression for the local volatility component according to [14]. The paper shows that the following relation holds between local volatility component σ_{SLV} and local volatility model σ_{LV} , or Dupire's formula [17]:

$$\sigma_{SLV}^2(t, K) = \frac{\sigma_{LV}^2(t, K)}{\mathbb{E}[\psi^2(S(t), \sigma(t)) \mid S(t) = K]}, \quad (116a)$$

$$= \frac{\sigma_{LV}^2(t, K)}{\omega^2(t) (P_d(t, T_N)/P_f(t, T_N))^{2-2\beta} K^{2\beta-2} \mathbb{E}[\sigma^2(t) \mid S(t) = K]}, \quad (116b)$$

see appendix G for more details. In the experiments later, conditional expectation $\mathbb{E}[\sigma^2(t) \mid S(t) = K]$ is estimated with a technique introduced in [14]. For practical purposes, it is more useful to express the Dupire local volatility in terms of implied volatility $\Sigma(T, K)$:

$$\sigma_{LV}^2(K, T) = \frac{\Sigma^2 + 2\Sigma T \left(\frac{\partial \Sigma}{\partial T} + (r_d(T) - r_f(T))K \frac{\partial \Sigma}{\partial K} \right)}{\left(1 - \frac{Ky}{\Sigma} \frac{\partial \Sigma}{\partial K} \right)^2 + K\Sigma T \left(\frac{\partial \Sigma}{\partial K} - \frac{1}{4}K\Sigma T \left(\frac{\partial \Sigma}{\partial K} \right)^2 + K \frac{\partial^2 \Sigma}{\partial K^2} \right)}. \quad (117)$$

The derivation of this expression is listed in appendix G0.2. Dupire's formula itself is derived in appendix G0.1. Dupire's local volatility is non-parametric, which is beneficial since it does not have model assumptions and therewith does not impose restrictions on the modelled spot process.

4.6 Normal version of the time dependent SABR model

In some markets or in some market circumstances, it might be more appropriate to work with normal instead of lognormal volatilities. FX rates are positive by definition and the market convention of quoting options in Black volatilities is therewith not likely to change. In the case of low or negative rates the traders in the interest rate market might prefer to switch from implied Black volatilities to implied Bachelier volatilities. In this case, path dependent interest options are priced by using a slightly different version of the time dependent SABR model. The following derivation is independent of the market to which the time dependent SABR model is applied, so the extension of the model to normal volatilities might be useful in future applications of the time dependent SABR model to markets different from FX.

When market quotes are expressed as Bachelier volatilities, the most natural version of time dependent SABR is the one with $\beta = 0$. Furthermore, Hagan's formula (37) is replaced with formula (62) in the first calibration step of obtaining the market effective parameters. The derivations of the mappings for γ and ρ are not influenced by switching to normal volatilities, so calibration steps based on these mappings are not affected. The derivation of the ω mapping [15, page 11-12] made use of a projection on lognormal dynamics. For $\beta = 0$ a more natural choice seems to project on normal dynamics. The following derivation shows that projection on normal dynamics makes the need for the Taylor approximation in [15, page 13] disappear. As a result, the obtained ω mapping has a zero (Taylor) truncation error.

Effective term structure for $\beta = 0$

The effective term structure mapping is derived by matching the at-the-money payoffs of the time dependent and the effective model:

$$\mathbb{E} \left[\left(F^{T_i}(T_i) - F_0^{T_i} \right)^+ \right] = \mathbb{E} \left[\left(\tilde{F}^{T_i}(T_i) - \tilde{F}_0^{T_i} \right)^+ \right], \text{ where } F_0^{T_i} = 1 = \tilde{F}_0^{T_i}. \quad (118)$$

In the derivation of the effective term structure for $\beta = 0$ it makes more sense to assume normal instead of lognormal dynamics [15, page 11-12] for the forward. Let's start with projecting the forward dynamics of the time-dependent model (99) on normal dynamics:

$$dF^{T_i}(t) = \omega_1(t)\sigma(t)dW_{\mathbb{F}^i}(t), \quad (119a)$$

$$\text{with solution } F^{T_i}(T_i) = \int_0^{T_i} \omega_1(t)\sigma(t)dW_{\mathbb{F}^i}(t) + F_0^{T_i}. \quad (119b)$$

The projection of the effective model is similar. Note that no projection takes place actually, since the SABR model (99) with $\beta = 0$ has the same forward dynamics as (119). In order to match the payoffs of the time dependent and effective model, several identities have to be derived first. Observe that for $I(T_i) := \int_0^{T_i} \omega_1(t)\sigma(t)dW_{\mathbb{F}^i}(t)$ and $J(T_i) := \int_0^{T_i} \omega_1^2(t)\sigma^2(t)dt$ it holds that

$$\mathbb{E} [I(T_i)] = 0 \text{ and} \quad (120a)$$

$$\mathbb{E} [I^2(T_i)] = \mathbb{E} \left[\int_0^{T_i} \omega_1^2(t)\sigma^2(t)dt \right] = \mathbb{E} [J(T_i)], \quad (120b)$$

by martingality of Ito integrals and the Ito isometry respectively. Furthermore, it holds that

$$\mathbb{E} \left[\sqrt{\frac{J(T_i)}{T_i}} W_{\mathbb{F}^i}(T_i) \mid \omega_1(t)\sigma(t), 0 \leq t \leq T_i \right], \quad (120c)$$

$$= \mathbb{E} \left[\sqrt{\frac{J(T_i)}{T_i}} \mid \omega_1(t)\sigma(t), 0 \leq t \leq T_i \right] \mathbb{E} [W_{\mathbb{F}^i}(T_i) \mid \omega_1(t)\sigma(t), 0 \leq t \leq T_i], \text{ and} \quad (120d)$$

$$\mathbb{E} \left[\frac{J(T_i)}{T_i} (W_{\mathbb{F}^i}(T_i))^2 \mid \omega_1(t)\sigma(t), 0 \leq t \leq T_i \right], \quad (120e)$$

$$= \mathbb{E} [J(T_i) \mid \omega_1(t)\sigma(t), 0 \leq t \leq T_i] \mathbb{E} \left[\frac{(W_{\mathbb{F}^i}(T_i))^2}{T_i} \mid \omega_1(t)\sigma(t), 0 \leq t \leq T_i \right], \quad (120f)$$

since $J(T_i)$ and $W_{\mathbb{F}^i}(T_i)$ are independent given $\{\omega_1(t)\sigma(t), 0 \leq t \leq T_i\}$. Applying the Tower property yields

$$\mathbb{E} \left[\sqrt{\frac{J(T_i)}{T_i}} W_{\mathbb{F}^i}(T_i) \right] = \mathbb{E} \left[\sqrt{\frac{J(T_i)}{T_i}} \right] \mathbb{E} [W_{\mathbb{F}^i}(T_i)] = 0, \text{ and} \quad (120g)$$

$$\mathbb{E} \left[\frac{J(T_i)}{T_i} (W_{\mathbb{F}^i}(T_i))^2 \right] = \mathbb{E} [J(T_i)] \mathbb{E} \left[\frac{(W_{\mathbb{F}^i}(T_i))^2}{T_i} \right] = \mathbb{E} [J(T_i)]. \quad (120h)$$

In conclusion, the following equation

$$I(T_i) \stackrel{d}{=} \sqrt{\frac{J(T_i)}{T_i}} W_{\mathbb{F}^i}(T_i), \quad (120i)$$

holds in distribution, since both left and right hand sides are normally distributed with the same mean and variance. As a consequence

$$F^{T_i}(T_i) = \sqrt{\frac{J(T_i)}{T_i}} W_{\mathbb{F}^i}(T_i) + F_0^{T_i}, \text{ with } F_0^{T_i} = 1. \quad (121)$$

Applying Bachelier's option pricing formula (60a) to the above equation results in

$$\mathbb{E} \left[(F^{T_i}(T_i) - 1)^+ \mid \omega_1(t)\sigma(t), 0 \leq t \leq T_i \right] = \sqrt{\frac{J(T_i)}{T_i}} \cdot \sqrt{T_i} \cdot \frac{1}{\sqrt{2\pi}} = \sqrt{\frac{J(T_i)}{2\pi}} \text{ and} \quad (122a)$$

$$\mathbb{E} \left[(F^{T_i}(T_i) - 1)^+ \right] = \frac{1}{\sqrt{2\pi}} \mathbb{E} \left[\sqrt{J(T_i)} \right], \quad (122b)$$

by the Tower property. So in the normal case there is no need for a Taylor series approximation. For the forward dynamics of the effective model

$$d\tilde{F}^{T_i}(t) = \tilde{\omega}_1 \tilde{\sigma}(t) dW_{\mathbb{F}^i}(t), \quad (123)$$

similar reasoning leads to

$$\mathbb{E} \left[\left(\tilde{F}^{T_i}(T_i) - \tilde{F}_0^{T_i} \right)^+ \right] = \frac{1}{\sqrt{2\pi}} \mathbb{E} \left[\sqrt{\tilde{J}(T_i)} \right], \text{ with } \tilde{J}(T_i) := \tilde{\omega}_1^2 \int_0^{T_i} \tilde{\sigma}^2(t) dt. \quad (124)$$

Matching the at-the-money-payoffs of the time dependent and effective models results in

$$\frac{1}{\sqrt{2\pi}} \mathbb{E} \left[\sqrt{\int_0^{T_i} \omega_1^2(t) \sigma^2(t) dt} \right] = \frac{\tilde{\omega}_1}{\sqrt{2\pi}} \mathbb{E} \left[\sqrt{\int_0^{T_i} \tilde{\sigma}^2(t) dt} \right]. \quad (125)$$

The integral of the left hand side of (125) is approximated by

$$\int_0^{T_i} \omega_1^2(t) \sigma^2(t) dt = \Delta_t \sum_{j=1}^M \omega_1^2(t_j) \sigma^2(t_j), \text{ where } \Delta_t = T_i/M. \quad (126a)$$

The left hand side of (125) reduces to the analytically solvable expression

$$\mathbb{E} \left[\sqrt{\int_0^{T_i} \omega_1^2(t) \sigma^2(t) dt} \right] = \omega_1(0) \sqrt{\Delta_t} \mathbb{E} \left[\left(\sum_{j=1}^M \frac{\omega_1^2(t_j) \sigma^2(t_j)}{\omega_1^2(0)} \right)^{1/2} \right], \quad (126b)$$

$$= \omega_1(0) \sqrt{\Delta_t} \mathbb{E} \left[e^{\frac{1}{2} Y_M} \right] = \omega_1(0) \sqrt{\Delta_t} \phi_{Y_M} \left(-\frac{1}{2} i \right), \quad (126c)$$

where ϕ_{Y_M} is the characteristic function of

$$Y_M := \log \left(\sum_{j=1}^M \frac{\omega_1^2(t_j) \sigma^2(t_j)}{\omega_1^2(0)} \right). \quad (126d)$$

Matching the ATM-payoffs of the time-dependent and effective models results in expression

$$\tilde{\omega}_1 = \frac{\omega_1(0) \phi_{Y_M} \left(-\frac{1}{2} i \right)}{\phi_{\tilde{Y}_M} \left(-\frac{1}{2} i \right)}, \quad (126e)$$

where

$$\tilde{Y}_M := \log \left(\sum_{j=1}^M \tilde{\sigma}^2(t_j) \right), \quad (126f)$$

is used in the calculation of the approximation of the right hand side of (125).

In conclusion, term structure mapping (126e) derived by projection on normal forward dynamics is equal to term structure mapping (111) derived by projection on lognormal forward dynamics together with a Taylor approximation. Both techniques result in the same mapping, but (126e) does not suffer from truncation errors caused by a Taylor approximation. Furthermore, when implied Bachelier volatilities are used with (126e) the term structure is calibrated more accurately than with (111) and implied Black volatilities. Moreover, this derivation shows that the same mapping can be used when traders want to use Bachelier volatilities to allow a negative underlying asset.

5 Experiments

In this section foreign exchange barrier options are priced with the time dependent SABR model, with and without local volatility component, as well as with the constant parameter SABR model and the local volatility model. These outcomes are compared with the valuation of ABN Amro’s proprietary trading model.

EUR-USD and GBP-USD foreign exchange market data

In order to price foreign exchange barrier options a pricing model needs to be calibrated to European-type foreign exchange call and put options. In these experiments two datasets are used: a EUR-USD and a GBP-USD dataset, both from 27/07/2015. Both markets are very liquid and all three currencies are among the most traded in the world [29]. The data is obtained from the data vendor of ABN Amro. The datasets contain information of options with expiries of roughly 1 day, 1 week, 2 weeks, 1 month, 2 months, 3 months, half a year and one year from ‘now’: 27/07/2015. As described before, the prices or implied volatilities of these options are not quoted directly. The implied volatilities of each expiry are extracted from the at-the-money volatility together with the risk-reversals and strangles of 10 and 25 delta. The market quotes of the EUR-USD market and the GBP-USD market are listed in table 5. The risk reversals and strangles are both calculated with a spot delta, since this is the default delta definition in the EUR-USD and GBP-USD markets [29]. Furthermore, the delta is not adjusted to the premium. The EUR-USD and GBP-USD spot rates at 27/07/2015 are 1.0988 and 1.55175 respectively.

Date	EUR-USD market quotes					GBP-USD market quotes				
	$\sigma_{\delta 0}$	RR ₂₅	RR ₁₀	STR ₂₅	STR ₁₀	$\sigma_{\delta 0}$	RR ₂₅	RR ₁₀	STR ₂₅	STR ₁₀
28/07/15	7.8650	7.89	0.14	-3.33	0.82	4.8800	0.00	-0.20	-4.88	1.29
03/08/15	11.5975	0.15	0.24	0.19	0.63	7.7700	-0.10	-0.17	0.17	0.54
10/08/15	11.2275	-0.06	-0.10	0.18	0.57	8.1775	-0.16	-0.28	0.16	0.50
26/08/15	10.3575	-0.28	-0.45	0.19	0.54	7.4950	-0.28	-0.50	0.17	0.52
25/09/15	10.7025	-0.51	-0.86	0.21	0.65	7.6550	-0.44	-0.80	0.20	0.61
27/10/15	10.5025	-0.65	-1.09	0.24	0.73	7.6175	-0.61	-1.10	0.23	0.71
27/01/16	10.4350	-1.07	-1.84	0.27	0.88	7.7975	-0.85	-1.52	0.26	0.83
27/07/16	10.3825	-1.35	-2.32	0.31	1.03	8.0425	-1.11	-2.03	0.30	0.98

Table 5: EUR-USD and GBP-USD market quotes retrieved at 27/07/2015 by ABN Amro. The EUR-USD spot FX rate at 27/07/2015 is 1.0988 and the GBP-USD spot rate is 1.55175.

The domestic currency in both currency pairs is the USD. For each expiry, the risk-reversal quote provides information about the implied volatility skew of the currency pair. The 25-delta risk reversal of the EUR-USD currency pair with an expiry one day in the future is large in magnitude and positive. This means that there is much more demand for a call position than for a put position on this currency pair. Since the FX rate is expressed in units of domestic currency per unit of foreign currency, this means that the demand for domestic currency USD is much larger than foreign currency EUR in this short time horizon. The 25-delta risk reversal of the EUR-USD currency pair for larger expiries shows that the expected demand for both currencies is roughly equal in one week and the demand for EUR, relative to USD, increases in time thereafter. The 10-delta risk reversals show similar behaviour, but less strong for a short time horizon. The risk reversals of the GBP-USD currency pair shows that the demand for GBP, relative to USD, is equal ‘currently’, but is expected to increase over time. Such preferences are likely to change over time according to [29]; the skew of a currency pair varies much over time.

The strangle quotes provide information about the curvature, or convexity, of the implied volatility smile. For both currencies, the curvature is large in a short time horizon of one day. For larger expiries the curvature decreases quickly and stabilizes. The level of the implied volatility

smile, measured by at-the-money volatility $\sigma_{\delta 0}$, increases after the first expiry, and stays quite constant thereafter.

Maturity		USD	EUR	GBP
28/07/15	O/N	0.30	-0.12	0.30
29/07/15	T/N	0.30	-0.12	0.30
30/07/15	S/N	0.30	-0.13	0.30
05/08/15	1W	0.30	-0.14	0.30
12/08/15	2W	0.30	-0.15	0.30
19/08/15	3W	0.30	-0.16	0.30
31/08/15	1M	0.30	-0.16	0.30
29/09/15	2M	0.30	-0.19	0.30
29/10/15	3M	0.30	-0.22	0.30
30/11/15	4M	0.33	-0.22	0.33
29/12/15	5M	0.35	-0.21	0.35
29/01/16	6M	0.37	-0.24	0.37
29/02/16	7M	0.40	-0.24	0.40
29/03/16	8M	0.43	-0.24	0.43
29/04/16	9M	0.45	-0.25	0.45
31/05/16	10M	0.48	-0.25	0.48
29/06/16	11M	0.51	-0.25	0.51
29/07/16	1Y	0.54	-0.26	0.54

Table 6: USD, EUR and GBP discount rate data, denoted in percentages (%), of maturities until one year (1Y). The data is retrieved at 27/07/2016 by ABN Amro.

A discount curve is constructed for each currency, in order to price foreign exchange options. Discount rates observed in the market are listed in table 6. The EUR and GBP discount rates are quoted in continuous compounding format and the USD discount rates are quoted in simple compounding format. This means that discount factors P are calculated in the first case with $P(0, T) = e^{-rT/100}$ and in the latter case with $P(0, T) = 1/(1 + r/100)^T$. Expiry T is expressed in years and discount rate r is expressed as a percentage. The rates in table 6 are expressed in percentages as well. These obtained discount factors are 'bootstrapped' to construct a time dependent discount function for each currency, as explained in section 4.2.

At this point the EUR-USD and GBP-USD market data is analyzed and (time dependent) discount curves are constructed. Before starting to price barrier options, the market quotes are transformed to implied volatility smiles for each expiry as explained in section 4.1.

EUR-USD and GBP-USD implied volatility surfaces

First, Hagan's formula is calibrated to the five implied volatilities of each expiry to obtain the market effective parameters for that particular expiry. The market effective parameters of the EUR-USD and GBP-USD markets are listed in tables 7a and 7b respectively. The parameters provide information about these markets and are useful to analyze.

Level representing term structure parameter $\tilde{\omega}$ is stable in time for both currencies, just like the at-the-money volatility $\sigma_{\delta 0}$. Skew representing parameter $\tilde{\rho}$ starts slightly positive and turns negative for larger expiries in the EUR-USD market. In the GBP-USD market this parameter starts slightly negative and the negative skew effect increases for larger expiries. This is consistent with the observations of the risk reversals in table 5. For both currencies, the curvature representing volatility-of-volatility parameter $\tilde{\gamma}$ is large initially, and decays (quickly) over time. In contrast, the strangles in table 5 are very stable. This difference is explained by the fact that $\tilde{\rho}$ and $\tilde{\gamma}$ have similar effects on the implied volatility smile. When $\tilde{\rho}$ increases relatively fast in magnitude compared to the strangles, this is compensated by a reduced $\tilde{\gamma}$.

Table 7: Tables (a) and (b) show market effective parameters obtained from calibrating the effective SABR model to the market quotes of the EUR-USD and the GBP-USD markets in 5 respectively.

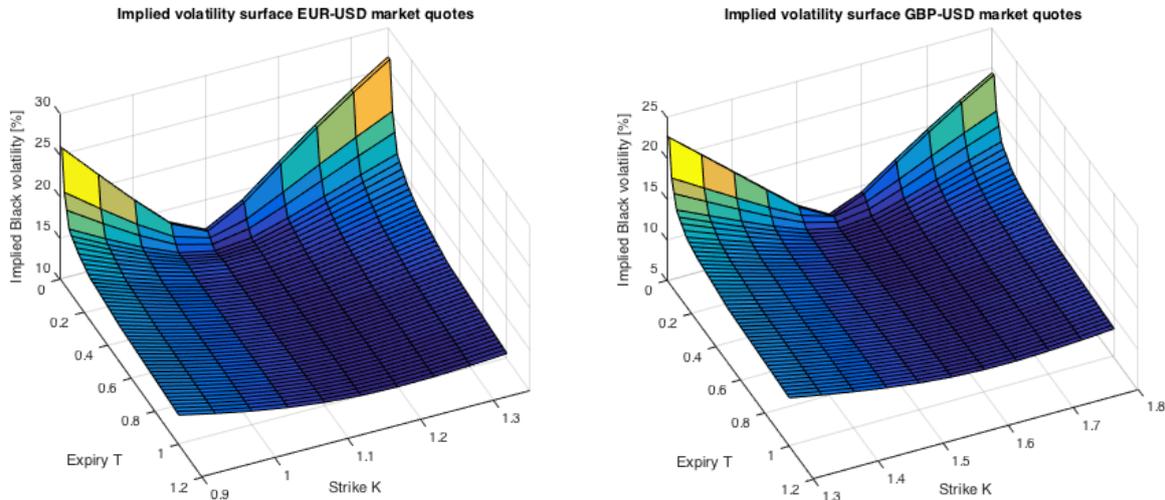
(a) EUR-USD market effective parameters

Date	$\tilde{\gamma}$	$\tilde{\omega}$	$\tilde{\rho}$
03/08/15	3.1228	0.1198	0.0546
10/08/15	2.1266	0.1162	0.0078
26/08/15	1.4533	0.1074	-0.0488
25/09/15	1.1097	0.1107	-0.0937
27/10/15	0.9518	0.1086	-0.1171
27/01/16	0.7357	0.1078	-0.1934
27/07/16	0.5592	0.1075	-0.2206

(b) GBP-USD market effective parameters

Date	$\tilde{\gamma}$	$\tilde{\omega}$	$\tilde{\rho}$
03/08/15	3.4919	0.0951	-0.0246
10/08/15	2.3194	0.1003	-0.0415
26/08/15	1.6771	0.0919	-0.0847
25/09/15	1.2630	0.0937	-0.1257
27/10/15	1.0970	0.0930	-0.1649
27/01/16	0.8247	0.0950	-0.2069
27/07/16	0.6206	0.0978	-0.2449

The market effective parameters in table 7 are used to construct implied volatility smiles for each expiry listed in the same table, but the input of (Van der Kamp's [12] adjustment of) Dupire's local volatility formula is a set of implied volatility smiles for a continuum of expiries. The smiles of table 7 are interpolated between the expiries to obtain an *implied volatility surface*. Interpolation of volatility smiles is more art than science. In these experiments the market effective parameters are linearly interpolated between every two expiries. A smile inbetween those two expiries is calculated by inserting those interpolated parameters in Hagan's formula. The resulting volatility surfaces are shown in figure 16 for illustration purposes.



(a) EUR-USD implied volatility surface

(b) GBP-USD implied volatility surface

Figure 16: Figure (a) and (b) show the implied volatility surfaces of the EUR-USD and GBP-USD market data respectively. Expiry T is expressed in years.

EUR-USD and GBP-USD barrier option data and experiment outcomes

At this point all information is available to start with the calibration and simulation of the models to price barrier options. The valuations of four different models are compared with the valuation of ABN Amro's proprietary trading model. It is compared with the local volatility model, the constant parameter SABR model and the time dependent SABR model with and without local volatility component. The simulation uses $5 \cdot 10^5$ simulation paths and $4 \cdot 365 = 1460$ steps per year. The number of steps is large because the barrier level of some products is close to the

current spot FX rate. In the EUR-USD market valuations of several call and put up-and-out (UO) barrier options are compared. There are also valuations of down-and-out (DO) barriers in the GBP-USD market. The EUR-USD and GBP-USD barrier option valuations are listed in table 8 and 9 respectively. The time dependent SABR model is calibrated to two expiries. When the expiry of the barrier is 27/01/16 (6M) the model is calibrated to expiries 27/10/15 (3M) and 27/01/16 (6M). If the barrier's expiry is 27/07/16 (1Y), the time dependent model is calibrated to market data of 27/01/16 (6M) and 27/07/16 (1Y). It is possible to calibrate the time dependent model to more than two expiries, fewer expiries speed up the calibration process though. Here is chosen for two expiries, since this was sufficient for the AUD-USD market data in [15] and the EUR-USD and GBP-USD are expected to be similar. Firstly, the AUD-USD, GBP-USD and EUR-USD currency pairs all have the USD as the domestic rate. Moreover, all the currency pairs are among the most traded currency pairs in the world [29].

				Model Barrier Option Value (USD)					
Expiry	Strike	Barrier		C/P	Prop	LV	C	TD	TD+LV
27/01/16	1.15	1.20	UO	Call	1,504.94	1,578.90	1,772.45	1,713.99	1,912.00
27/01/16	1.05	1.15	UO	Put	12,644.97	13,053.32	12,224.06	11,871.61	11,696.56
27/01/16	1.05	1.15	UO	Call	10,877.67	10,347.13	11,870.63	13,192.51	12,717.81
27/07/16	1.10	1.25	UO	Call	19,465.73	18,794.05	20,219.19	20,341.85	20,745.76
27/07/16	1.09	1.20	UO	Put	36,052.25	36,807.21	35,662.96	35,012.86	34,915.28
27/07/16	1.09	1.20	UO	Call	9,915.04	8,934.60	10,333.52	11,271.97	10,892.82

Table 8: EUR-USD barrier option valuations of the proprietary trading model of ABN Amro (Prop), the local volatility (LV) model, the constant parameter SABR (C) model and the time dependent SABR model with (TD+LV) and without (TD) a local volatility component. The models were calibrated to market data from table 5. The notional of every barrier option is 1 million EUR and the currency of every valuation is USD. All barriers are up-and-out (OU) barriers.

				Model Barrier Option Value (USD)					
Expiry	Strike	Barrier		C/P	Prop	LV	C	TD	TD+LV
27/01/16	1.55	1.45	DO	Call	33,259.16	32,211.66	31,592.77	31,025.54	31,265.27
27/01/16	1.60	1.50	DO	Put	9,165.60	8,871.49	10,500.28	12,223.26	11,293.41
27/01/16	1.60	1.50	DO	Call	12,694.63	12,770.43	11,657.36	11,324.16	11,132.90
27/07/16	1.55	1.55	UO	Call	6,435.68	5,567.38	6,568.87	7,155.15	7,028.77
27/07/16	1.50	1.50	UO	Put	22,384.52	25,868.98	23,801.51	22,997.73	22,865.21
27/07/16	1.50	1.50	UO	Call	4,165.06	3,636.34	4,730.38	5,859.41	5,445.44

Table 9: GBP-USD barrier option valuations of the proprietary trading model of ABN Amro (Prop), the local volatility (LV) model, the constant parameter SABR (C) model and the time dependent SABR model with (TD+LV) and without (TD) a local volatility component. The models were calibrated to market data from table 5. The notional of every barrier option is 1 million GBP and the currency of every valuation is USD. The barriers are either up-and-out (OU) or down-and-out (DO) barriers.

The local volatility model and the time dependent SABR model with local volatility component are both perfectly calibrated to the volatility surface by construction. In contrast, table 8 and 9 show that these models value path dependent options differently. The reason is that the transition probability density function of one future state to another, is different for these models. These probability density functions are not observable in the market and therewith cannot be calibrated, they must be assumed.

6 Conclusion

Black's model assumes a constant volatility parameter. The volatility that should be inserted in Black's formula to obtain the market price is called the implied volatility. It can be observed in the market that the implied volatility is a function of both the strike and the expiry; the implied volatility surface. Black's model can be extended with a local volatility function to be able to calibrate to an implied volatility surface. The dynamics of the implied volatility smile with respect to a change in the underlying are not in line with observed market dynamics under the local volatility model. This leads to incorrect risk metrics. The SABR model is a two factor model with dynamics in line with observed market dynamics. Hagan's formula is an approximation of the implied volatility of the SABR model. It can be used to calibrate a volatility smile to a set of market data of a certain expiry. The formula breaks down for rates close to zero, so it is not reliable in a low interest rate environment. Finally, conventional methods in interest rate option pricing assume a zero probability on negative rates.

Three different solution methods were discussed to cope with negative interest rates: normal models, displaced models and free boundary models. The forward rate has a normal distribution under Bachelier's model and it therewith allows negative forward rates. A great benefit of Bachelier's model is that it does not introduce an additional (shift) parameter. Furthermore, analytic expressions are available for the call and put price valuations and for its risk metrics. The normal SABR model can be used to price path dependent options in a negative rate context. It is calibrated with a special version of Hagan's formula for implied Bachelier volatilities with $\beta = 0$. The only disadvantage of the normal models is that they assume a positive probability on large negative rates. In contrast, displaced models limit the largest negative rate at $-s$, where $s > 0$ is the (constant) displacement parameter. It could be argued that there is a limit to how low rates can go, since storing (and transporting) cash becomes cheaper than the interest at a certain point. For financial institutions this is different however, because they are obliged by regulation to have a certain part of their assets stored at the central bank. The disadvantage of a displaced model is that the displacement parameter has to be chosen a priori. This choice is more of an art than a science. When the rate decreases sufficiently, a larger displacement parameter needs to be introduced and all implied volatilities, valuations and risk metrics have to be adjusted. Finally, the free boundary models were proposed as a solution. These models can model rates from the entire real line and do not introduce an additional parameter. Nevertheless, the models have an uncontrollable spike around zero in their probability distribution, while this is the most crucial area in a low or negative rate environment. Furthermore, there is no accurate analytic (calibration) formula available for the free boundary SABR model. Moreover, the free boundary models do not have an obvious model interpretation and are therewith not intuitive for traders.

When Hagan's formula is calibrated to market data of different expiries this results in a different set of SABR parameters for each expiry. This implies that the constant parameter SABR model does not have enough flexibility to calibrate to multiple expiries simultaneously. The pricing and hedging of path dependent financial products requires an accurate calibration to prices of European-type options with different expiries that contain information about market behaviour through time. Extending the SABR model with time dependent parameters solves this problem. The initial research goal was to price path dependent interest rate options with the time dependent SABR model in a negative interest rate environment. However, it appeared that there is no obvious transformation of the time dependent FX-SABR model to the interest rate market. An important assumption of the time dependent FX-SABR model is deterministic discount curves (due to deterministic interest rates), while a realistic model for pricing interest rate derivatives requires stochastic interest rates. The research goal was adjusted to the pricing and hedging of foreign exchange barrier options. Time dependent SABR has been proven to work well in this application area and ABN Amro was eager to gain knowledge about how this method performs compared to their own proprietary trading model. The model can be

extended with a local volatility component to improve the calibration accuracy. Furthermore, it can be adjusted such that it can model negative underlyings by calibrating to implied Bachelier volatilities. In the experiments section, different foreign exchange barrier options were valued by the local volatility model, the constant parameter SABR model, the time dependent SABR model with and without local volatility component and with ABN Amro's proprietary trading model. Using the valuations of ABN Amro's proprietary trading model and the (perfectly calibrated) local volatility model as a benchmark, the various SABR models produce realistic valuations. Without market quotes of foreign exchange barrier options, there can be no conclusion drawn about what model produces the most accurate valuations.

6.1 Further research

The following issues or possibilities are recommended for future research:

1. The displacement parameter s influences the level of displaced volatility as well as risk metrics such as delta and vega. An operational transition from a nondisplaced model to its displaced version should be implemented consistently throughout the complete system of a financial institution. The impact of plugging a nondisplaced volatility in a displaced model on the risk metrics and prices may be large. This impact should be analyzed.
2. The displacement parameter s influences the value of a path-dependent option, since it influences the underlying transition probability density function. There is a zero probability on an underlying smaller than $-s$ for instance. When a displaced model is used for pricing such a path-dependent option, the impact of the displacement on the valuation should be examined.
3. There were no barrier option market quotes available to test the quality of the barrier option valuations in the experiments in chapter 5. When this market data would be available, it would be possible to determine what model performs best. Without this data the model valuations can be compared with each other only. However, the EUR-USD and GBP-USD markets are comparable to the AUD-USD market used in [15]. The USD, EUR, GBP and AUD are all among the most traded currencies in the world according to [29]. These markets are highly liquid and they are all currencies of Western countries. In the AUD-USD market of [15] the time dependent SABR model with local volatility component had the smallest error in most cases. Therefore, the time dependent SABR model with local volatility component is also likely to give accurate valuations of barrier options in the EUR-USD and GBP-USD markets. Nevertheless, this should be examined with actual EUR-USD and GBP-USD barrier option market quotes.
4. The time dependent SABR model could possibly be applied in different markets than the foreign exchange market, it would be interesting to research these possibilities.

References

- [1] Andersen and Piterbarg. *Interest Rate Modeling*. 1 edition, 2010. Atlantic Financial Press.
- [2] Antonov, Konikov, and Spector. SABR spreads its wings. *Risk*, 26(8):58, August 2013.
- [3] Antonov, Konikov, and Spector. The Free Boundary SABR: Natural Extension to Negative Rates. *Available at SSRN 2557046*, January 2015. Numerix.
- [4] Bachelier. Théorie de la spéculation. *Annales scientifiques de l'École Normale Supérieure*, 3(17):21–86, 1900.
- [5] Baker, Bener, and Zilber. FX Barriers with Smile Dynamics. *Available at SSRN 964627*, December 2004.
- [6] Bartlett. Hedging under SABR Model. *Willmott Magazine*, pages 2–4, July 2006.
- [7] Black. The pricing of commodity contracts. *Journal of Financial Economics*, 3:167–179, 1976.
- [8] Breeden and Litzenberger. Prices of State-Contingent Claims Implicit in Option Prices. *The Journal of Business*, 51(4):621–651, October 1978.
- [9] Carr and Madan. A note on sufficient conditions for no arbitrage. *Finance Research Letters*, 2(3):125–130, June 2005.
- [10] Chen, Grzelak, and Oosterlee. Calibration and Monte Carlo Pricing of the SABR-Hull-White Model for Long-Maturity Equity Derivatives. *The Journal of Computational Finance (79–113) Volume*, 15, December 2011.
- [11] Bermudez de Castro and Lukic. Validation of interest rate caplets and floorlets construction in the displaced diffusion model. January 2015. ABN AMRO Internal Report (VAL14-04).
- [12] Van der Kamp. Local Volatility Modelling. July 2009. MSc thesis University of Twente.
- [13] Van der Stelt. Interest Rate Curve Building and Vanilla Products without Optionality. November 2013. ABN AMRO Internal Report (RVL13-03).
- [14] Van der Stoep, Grzelak, and Oosterlee. The Heston Stochastic-Local Volatility Model: Efficient Monte Carlo Simulation. *International Journal of Theoretical and Applied Finance*, 17(7):1450045, 2014.
- [15] Van der Stoep, Grzelak, and Oosterlee. The Time-Dependent FX-SABR Model: Efficient Calibration based on Effective Parameters. *International Journal of Theoretical and Applied Finance*, 18(6):1550042, 2015.
- [16] Derman and Kani. The Volatility Smile and Its Implied Tree. January 1994. Goldman Sachs.
- [17] Dupire. Pricing with a Smile. *Risk*, 7(1):18–20, January 1994.
- [18] Feller. Two Singular Diffusion Problems. *Annals of Mathematics*, 54(1):173–182, July 1951.

- [19] Bank for International Settlements. Amounts outstanding of over-the-counter (otc) derivatives. <http://www.bis.org/statistics/dt1920a.pdf>, March 2015. BIS Quarterly Review.
- [20] Hagan, Kumar, Lesniewski, and Woodward. Managing Smile Risk. *Willmott Magazine*, July 2002.
- [21] Hagan and Lesniewski. LIBOR market model with SABR style stochastic volatility. *JP Morgan Chase and Ellington Management Group*, 32, May 2008.
- [22] Hagan and Woodward. Equivalent Black Volatilities. *Applied Mathematical Finance*, (6):147–157, 1999. Numerix and Bank of Tokyo-Mitsubishi.
- [23] Henry-Labordère. *Analysis, geometry, and modeling in finance: Advanced methods in option pricing*. CRC Press, 2008.
- [24] Hull. *Options, Futures and Other Derivates*. 7 edition, 2009. University of Toronto.
- [25] Hull and White. LIBOR vs. OIS: The Derivatives Discounting Dilemma. *Journal of Investment Management*, 11, March 2013.
- [26] Oblój. Fine-tune your smile: correction to Hagan et al. March 2008. Imperial College London.
- [27] Oosterlee, Grzelak, and Fang. *Efficient Valuation and Computation in Finance (unpublished)*. 2015. TU Delft and CWI.
- [28] Piterbarg. Time to smile. *Risk*, 19(5):71–75, May 2005.
- [29] Reiswich and Wystup. FX Volatility Smile Construction. *CPQF Working Paper Series*, (20), September 2009. Frankfurt School of Finance and Management.
- [30] Rouah. T-Forward Measure.
- [31] Shreve. *Stochastic Calculus for Finance II: Continuous-Time Models*. 1 edition, 2004. Springer-Verlag New York.
- [32] Söderlind. Lecture Notes - Econometrics: Some Statistics. July 2013. University of St. Gallen.
- [33] Wandl and Li. Interest Rate Volatility Cube. May 2014. ABN AMRO Internal Report (VAL14-08).
- [34] West. Calibration of the SABR Model in Illiquid Markets. *Applied Mathematical Finance*, 12(4):371–385, December 2005.
- [35] Wilmott. *Quantitative Finance*. 2 edition, 2006. Wiley.
- [36] Zhang and Oosterlee. Efficient Pricing of European-Style Asian Options under Exponential Lévy Processes Based on Fourier Cosine Expansions. *SIAM Journal on Financial Mathematics*, 4:399–426, 2013. TU Delft and CWI.

Appendix A

Solution to Black's model

Define transformation $Y_t = \log(F_t)$ on lognormal SDE (13), then according to Ito's formula

$$dY_t = \frac{1}{F_t} dF_t + \frac{1}{2} \cdot -\frac{1}{F_t^2} (dF_t)^2 = \frac{1}{F_t} (\sigma F_t dW_t) - \frac{1}{2} \frac{1}{F_t^2} \sigma^2 F_t^2 dt = \sigma dW_t - \frac{1}{2} \sigma^2 dt, \quad (\text{A.1a})$$

$$Y_t = Y_0 + \sigma W_t - \sigma^2 t/2. \quad (\text{A.1b})$$

This implied that $Y_t \sim \mathcal{N}(Y_0 - \sigma^2 t/2, \sigma^2 t)$ and that F_t is lognormally distributed:

$$F_t = \exp(Y_t) = F_0 \exp(\sigma W_t - \sigma^2 t/2). \quad (\text{A.1c})$$

Note that if some stochastic variable $Y \sim \mathcal{N}(\mu, \xi^2)$ and $X = \exp(Y)$, then truncated mean $\mathbb{E}[Y | Y > K] = \exp(\mu + \xi^2) \Phi(\xi - K_0) / \Phi(-K_0)$, where $K_0 = (\log K - \mu) / \xi$ [32]. Furthermore, $\mathbb{E}[X] = \exp(\mu + \xi^2/2)$. Forward F_t is a martingale under Black's model, since

$$\mathbb{E}[F_t | \mathcal{F}_0] = \exp((\log(F_0) - \sigma^2 t/2) + (\sigma^2 t)/2) = F_0. \quad (\text{A.1d})$$

Assuming a deterministic zero coupon bond $P(0, t)$, the value $V^c(0)$ of a European-type call option at time 0 under Black's model is given by

$$\frac{V^c(0)}{P(0, t)} = \mathbb{E}[\max(F_t - K, 0) | \mathcal{F}_0] = \mathbb{E}[F_t - K | F_t > K] P(F_t > K) + 0 \cdot P(F_t < K), \quad (\text{A.1e})$$

$$= \left[1 - \Phi\left(\frac{\log(K) - (Y_0 - \sigma^2 t/2)}{\sigma\sqrt{t}}\right) \right] \cdot \left[\exp\left(\left(Y_0 - \frac{\sigma^2 t}{2}\right) + \frac{\sigma^2 t}{2}\right), \quad (\text{A.1f}) \right.$$

$$\left. \cdot \Phi\left(\sigma\sqrt{t} - \frac{\log(K) - (Y_0 - \sigma^2 t/2)}{\sigma\sqrt{t}}\right) \right] / \left[\Phi\left(-\frac{\log(K) - (Y_0 - \sigma^2 t/2)}{\sigma\sqrt{t}}\right) - K \right], \quad (\text{A.1g})$$

$$= \exp(Y_0) \Phi\left(\frac{\sigma\sqrt{t}}{2} - \frac{\log(K) - Y_0}{\sigma\sqrt{t}}\right) - K \Phi\left(-\frac{\sigma\sqrt{t}}{2} - \frac{\log(K) - Y_0}{\sigma\sqrt{t}}\right), \quad (\text{A.1h})$$

$$= F_0 \Phi(d_1) - K \Phi(d_2), \text{ where } d_1 \equiv \frac{\sigma\sqrt{t}}{2} + \frac{\log(F_0/K)}{\sigma\sqrt{t}} \text{ and } d_2 \equiv d_1 - \sigma\sqrt{t}. \quad (\text{A.1i})$$

By the put-call parity (16), the value $V^p(0)$ of a European-type put option at time 0 under Black's model is

$$V^p(0) = V^c(0) - P(0, t)(F_0 - K) = P(0, t) [F_0 (\Phi(d_1) - 1) - K (\Phi(d_2) - 1)], \quad (\text{A.1j})$$

$$= P(0, t) [K \Phi(-d_2) - F_0 \Phi(-d_1)]. \quad (\text{A.1k})$$

Appendix B

Risk metrics of Black's model

The delta Δ_c , vega Λ , and gamma Γ of a European call option under Black's model are derived as follows:

$$\Delta_c \equiv \frac{\partial V^c}{\partial F} = \frac{\partial}{\partial F} (P(0, T)[F\Phi(d_1) - K\Phi(d_2)]), \quad (\text{B.1a})$$

$$= P(0, T) \left[\Phi(d_1) + F\Phi'(d_1) \frac{\partial d_1}{\partial F} - K\Phi'(d_2) \frac{\partial d_2}{\partial F} \right], \quad (\text{B.1b})$$

$$\stackrel{(\text{B.1e})}{=} P(0, T) \left[\Phi(d_1) + \frac{\partial d_1}{\partial F} (F\phi(d_1) - K\phi(d_2)) \right] \stackrel{(\text{B.1g}), (\text{B.1i})}{=} P(0, T)\Phi(d_1),$$

$$\Lambda \equiv \frac{\partial V^{c,p}}{\partial \sigma} = \frac{\partial}{\partial \sigma} (P(0, T)[F\Phi(d_1) - K\Phi(d_2)]) = P(0, T) \left[F\phi(d_1) \frac{\partial d_1}{\partial \sigma} - K\phi(d_2) \frac{\partial d_2}{\partial \sigma} \right], \quad (\text{B.1c})$$

$$\stackrel{(\text{B.1e})}{=} P(0, T) \left[F\phi(d_1) \left(\frac{\partial d_2}{\partial \sigma} + \sqrt{T} \right) - K\phi(d_2) \frac{\partial d_2}{\partial \sigma} \right] \stackrel{(\text{B.1g}), (\text{B.1i})}{=} P(0, T)F\sqrt{T}\phi(d_1),$$

$$\Gamma \equiv \frac{\partial^2 V^{c,p}}{\partial F^2} = \frac{\partial \Delta_{c,p}}{\partial F} = P(0, T) \frac{\partial \Phi(d_1)}{\partial F} = P(0, T)\phi(d_1) \frac{\partial d_1}{\partial F} = P(0, T)\phi(d_1) \frac{1}{\sigma\sqrt{T}F}, \quad (\text{B.1d})$$

where the following identities were used:

$$d_2 = d_1 - \sigma_B\sqrt{T} \Rightarrow \frac{\partial d_1}{\partial F} = \frac{\partial d_2}{\partial F} \frac{\partial d_1}{\partial d_2}, \quad \frac{\partial d_1}{\partial \sigma} = \frac{\partial d_2}{\partial \sigma} + \sqrt{T}, \quad (\text{B.1e})$$

$$0 = F\phi(d_1) - K\phi(d_2) \Leftrightarrow \frac{F}{K} = \frac{\phi(d_2)}{\phi(d_1)} = \exp\left(-\frac{1}{2}(d_2^2 - d_1^2)\right), \quad (\text{B.1f})$$

$$\Leftrightarrow -2 \log(F/K) = d_2^2 - d_1^2, \quad (\text{B.1g})$$

$$d_2^2 - d_1^2 = (d_1 - \sigma\sqrt{T})^2 - d_1^2 = \sigma^2 T - 2d_1\sigma\sqrt{T}, \quad (\text{B.1h})$$

$$= \sigma^2 T - 2(\log(F/K) + \sigma^2 T/2) = -2 \ln(F/K). \quad (\text{B.1i})$$

Appendix C

Risk metrics under Bachelier's model

Note that if $X \sim \mathcal{N}(\mu, \xi^2)$, then $\mathbb{E}[X \mid X > K] = \mu + \xi\phi(u)/(1 - \Phi(u))$, where $u = (K - \mu)/\xi$ [32]. Define zero coupon bond $P(0, T)$, expiry T , strike K , T -forward rate $F^T(t)$ at time t , current forward $F := F^T(0)$, constant volatility σ , and $d \equiv \frac{F-K}{\sigma\sqrt{t}}$, then the following expressions for the current value of the call V^c and put V^p options hold:

$$V^c(0)/P(0, T) = \mathbb{E}[\max(F^T(T) - K, 0) \mid \mathcal{F}_0], \quad (\text{C.1a})$$

$$= \mathbb{E}[F^T(T) - K \mid F^T(T) > K] P(F^T(T) > K) + 0 \cdot P(F^T(T) < K), \quad (\text{C.1b})$$

$$= [1 - P(F^T(T) \leq K)] (\mathbb{E}[F^T(T) \mid F^T(T) > K] - K), \quad (\text{C.1c})$$

$$= \left[1 - \Phi\left(\frac{K - F}{\sigma\sqrt{t}}\right)\right] (\mathbb{E}[F^T(T) \mid F^T(T) > K] - K), \quad (\text{C.1d})$$

$$= [1 - \Phi(-d)] \left(F + \sigma\sqrt{t} \frac{\phi(-d)}{1 - \Phi(-d)} - K\right) = (F - K) \Phi(d) + \sigma\sqrt{t}\phi(d).$$

$$V^p(0) = V^c(0) - P(0, T)(F - K) = P(0, T) [(F - K) (\Phi(d) - 1) + \sigma\sqrt{t}\phi(d)], \quad (\text{C.1e})$$

$$= P(0, T) [(K - F) \Phi(-d) + \sigma\sqrt{t}\phi(d)]. \quad (\text{C.1f})$$

An option's call delta Δ_c , put delta Δ_p , vega Λ , and gamma Γ under Bachelier's model are derived as follows

$$\Delta_c \equiv \frac{\partial V^c}{\partial F} = P(0, T) \frac{\partial}{\partial F} [(F - K)\Phi(d) + \sigma\sqrt{t}\phi(d)], \quad (\text{C.1g})$$

$$= P(0, T) [(F - K)\phi(d)d' + \Phi(d) + \sigma\sqrt{t}\phi'(d)d'], \quad (\text{C.1h})$$

$$= P(0, T) \left[(F - K)\phi(d) \frac{1}{\sigma\sqrt{t}} + \Phi(d) + \sigma\sqrt{t}\phi(d) \cdot -d \cdot \frac{1}{\sigma\sqrt{t}}\right], \quad (\text{C.1i})$$

$$= P(0, T) [d\phi(d) + \Phi(d) - d\phi(d)] = P(0, T) [\Phi(d)], \quad (\text{C.1j})$$

$$\Delta_p = \frac{\partial V^p}{\partial F} = \frac{\partial}{\partial F} [V^c - D_t(F - K)] = \frac{\partial V^c}{\partial F} - P(0, T) = -P(0, T)\Phi(-d), \quad (\text{C.1k})$$

$$\Lambda \equiv \frac{\partial V^c}{\partial \sigma} = P(0, T) \frac{\partial}{\partial \sigma} [(F - K)\Phi(d) + \sigma\sqrt{t}\phi(d)], \quad (\text{C.1l})$$

$$= P(0, T) \left[(F - K)\phi(d) \frac{\partial d}{\partial \sigma} + \sqrt{t}\phi(d) + \sigma\sqrt{t}\phi'(d) \frac{\partial d}{\partial \sigma}\right], \quad (\text{C.1m})$$

$$= P(0, T) \left[(F - K)\phi(d) \cdot -\frac{1}{\sigma}d + \sqrt{t}\phi(d) + \sigma\sqrt{t} \cdot -d\phi(d) \cdot -\frac{1}{\sigma}d\right], \quad (\text{C.1n})$$

$$= P(0, T) [-\sqrt{t}\phi(d)d^2 + \sqrt{t}\phi(d) + \sqrt{t}\phi(d)d^2] = P(0, T)\sqrt{t}\phi(d), \quad (\text{C.1o})$$

$$\Gamma \equiv \frac{\partial^2 V^c}{\partial F^2} = \frac{\partial \Delta_c}{\partial F} = P(0, T) \frac{\partial}{\partial F} \Phi(d) = P(0, T) \frac{\phi(d)}{\sigma\sqrt{t}}. \quad (\text{C.1p})$$

$$(\text{C.1q})$$

Appendix D

Transition probability density of the CEV process

The CEV process (31) can be transformed to a time-changed squared Bessel process, for which the transition density is known [27]. The transition density of the CEV process can be derived from the inverse transformation. To start with the derivation, first apply the transformation $X_t = F_t^{1-\beta}/(1-\beta)$ to (31), for $\beta \neq 1$

$$dX_t = F_t^{-\beta} dF_t + \frac{1}{2} \cdot -\beta F_t^{-1-\beta} (dF_t)^2, \quad (\text{D.1a})$$

$$= F_t^{-\beta} \sigma F_t^\beta dW_t - \frac{\beta}{2} F_t^{-1-\beta} \sigma^2 F_t^{2\beta} dt, \quad (\text{D.1b})$$

$$= \sigma dW_t - \frac{\beta}{2} F_t^{\beta-1} \sigma^2 dt, \quad (\text{D.1c})$$

$$= \sigma dW_t - \frac{\beta}{2(1-\beta)X_t} \sigma^2 dt. \quad (\text{D.1d})$$

Applying transformation $Y_t = X_t^2$, results in

$$dY_t = 2X_t dX_t + \frac{1}{2} \cdot 2(dX_t)^2, \quad (\text{D.1e})$$

$$= 2\sqrt{|Y_t|} \left(\sigma dW_t - \frac{\beta}{2(1-\beta)\sqrt{|Y_t|}} \sigma^2 dt \right) + \sigma^2 dt, \quad (\text{D.1f})$$

by (D.1),

$$= 2\sqrt{|Y_t|} \sigma dW_t + \frac{1-2\beta}{1-\beta} \sigma^2 dt, \quad (\text{D.1g})$$

$$= 2\sqrt{|Y_t|} \sigma dW_t + \delta \sigma^2 dt. \quad (\text{D.1h})$$

where $\delta \equiv (1-2\beta)/(1-\beta)$. We define the time-change $\nu(t) = \sigma^2 t$. Then, $Y_t = Z_{\nu(t)}$, where $\{Z_t\}_{t \geq 0}$ is a δ -dimensional squared Bessel process:

$$dZ_t = 2\sqrt{|Z_t|} dW_t + \delta dt. \quad (\text{D.1i})$$

Solutions to the squared Bessel process and its transition probability density

For a (standard) squared Bessel process $\{Z_t\}_{t \geq 0}$, governed by SDE (D.1i), the following statements hold for solutions to this SDE [27]:

1. For $\delta < 2$, $Z_t = 0$ is an attainable boundary for the squared Bessel process.
2. For $\delta \leq 0$, the SDE has a unique strong solution and the boundary condition at zero is absorbing.
3. For $0 < \delta < 2$, the SDE does not have a unique solution, unless a boundary condition is specified at $Z_t = 0$.
4. For $\delta \geq 2$, the SDE has a unique solution and zero is not attainable.

For the cases $\delta \leq 0$ and $0 < \delta < 2$ the transition probability density $f_\delta(t, Z_0, Z_t)$ of the squared Bessel process (D.1i) is known [27]:

1. For $\delta \leq 0$ and for $0 < \delta < 2$ with an absorbing boundary at $Z_t = 0$:

$$f_\delta(t, Z_0, Z_t) = \frac{1}{2t} \left(\frac{Z_t}{Z_0} \right)^{\frac{\delta-2}{4}} \exp\left(-\frac{Z_0 + Z_t}{2t}\right) I_{|\frac{\delta-2}{2}|} \left(\frac{\sqrt{Z_0 Z_t}}{t} \right). \quad (\text{D.2})$$

2. For $0 < \delta < 2$ and a reflecting boundary at $Z_t = 0$:

$$f_\delta(t, Z_0, Z_t) = \frac{1}{2t} \left(\frac{Z_t}{Z_0} \right)^{\frac{\delta-2}{4}} \exp\left(-\frac{Z_0 + Z_t}{2t}\right) I_{\frac{\delta-2}{2}} \left(\frac{\sqrt{Z_0 Z_t}}{t} \right), \quad (\text{D.3})$$

where

$$I_a(x) \equiv \sum_{j=0}^{\infty} \frac{(x/2)^{2j+a}}{j! \Gamma(a+j+1)}, \quad (\text{D.4})$$

is the modified Bessel function of the first kind and where

$$\Gamma(x) \equiv \int_0^{\infty} u^{x-1} e^{-u} du, \quad (\text{D.5})$$

is the gamma function. A detailed derivation of the solution to the Fokker-Planck equation related to a class of diffusion processes that contains (D.1i) can be found in Feller [18].

Transition probability density of the CEV process

The transition probability density of the CEV process is determined by a mapping from the transition density of the squared Bessel process. Note that F_t in (31) and Z_t in (D.1i) are related via the equation

$$F_t = \left[(1 - \beta) \sqrt{|Z_{\nu(t)}|} \right]^{\frac{1}{1-\beta}}, \text{ or equivalently via } Z_{\nu(t)} = \frac{F_t^{2(1-\beta)}}{(1 - \beta)^2}. \quad (\text{D.6a})$$

Define the map T as

$$T : y \mapsto [(1 - \beta) \sqrt{y}]^{\frac{1}{1-\beta}}, \text{ for } y \geq 0. \quad (\text{D.6b})$$

Since T is strictly increasing for $0 < \beta < 1$, and thereby one-to-one, its inverse T^{-1}

$$T^{-1} : f \mapsto \frac{f^{2(1-\beta)}}{(1 - \beta)^2}, \text{ for } f \geq 0, \quad (\text{D.6c})$$

exists and is strictly increasing as well. The mapping allows equation (D.6a) to be rewritten as $F_t = T(Z_{\nu(t)})$ and $Y_t = Z_{\nu(t)} = T^{-1}(F_t)$. With this change of variables, it follows that the transition density $p(F_t | F_0)$ of the CEV process is given by

$$p(F_t | F_0) = f_{\delta}(\nu(t), Z_0, Z_{\nu(t)}) \cdot \left[\frac{dT^{-1}(f)}{df} \right]^{f=F_t} = f_{\delta}(\nu(t), T^{-1}(F_0), T^{-1}(F_t)) \cdot \left[\frac{dT^{-1}(f)}{df} \right]^{f=F_t}. \quad (\text{D.6d})$$

In more detail, the transition probability density of the CEV process (31) is of the following form for $0 < \beta < 1$:

1. For $0 < \beta < \frac{1}{2}$ with an absorbing boundary at $F_t = 0$ and for $\frac{1}{2} \leq \beta < 1$ without (the need of) applying a boundary condition:

$$\begin{aligned} p_A(t, f, F_0) &:= P(F_t = f | F_0) & (\text{D.7a}) \\ &= \frac{1}{\nu(t)} \left(\frac{f}{F_0} \right)^{-\frac{1}{2}} \exp \left(-\frac{f^{2(1-\beta)} + F_0^{2(1-\beta)}}{2(1-\beta)^2 \nu(t)} \right) I_{|\frac{\delta-2}{2}|} \left(\frac{(F_0 f)^{1-\beta}}{\nu(t)(1-\beta)^2} \right) \frac{f^{1-2\beta}}{1-\beta}. & (\text{D.7b}) \end{aligned}$$

2. For $0 < \beta < \frac{1}{2}$ with a reflecting boundary at $F_t = 0$:

$$\begin{aligned} p_R(t, f, F_0) &:= P(F_t = f | F_0) & (\text{D.8a}) \\ &= \frac{1}{\nu(t)} \left(\frac{f}{F_0} \right)^{-\frac{1}{2}} \exp \left(-\frac{f^{2(1-\beta)} + F_0^{2(1-\beta)}}{2(1-\beta)^2 \nu(t)} \right) I_{\frac{\delta-2}{2}} \left(\frac{(F_0 f)^{1-\beta}}{\nu(t)(1-\beta)^2} \right) \frac{f^{1-2\beta}}{1-\beta}. & (\text{D.8b}) \end{aligned}$$

Appendix E

Integrals in the volvol mapping for piecewise constant parameters

Effective volatility-of-volatility (volvol) parameter $\tilde{\gamma}$ can be obtained from time-dependent volvol parameter $\gamma(t)$ with mapping (105) introduced in section 4.3. When $\gamma(t)$ is piecewise-constant, there are analytical expressions available for the integrals used in the calculations of the effective parameter $\tilde{\gamma}$. These expressions are derived in this section.

For expiries $0 =: T_0 < T_1 < T_2 < \dots < T_{N-1} < T_N$ and $i \in \{1, 2, \dots, N\}$ define piecewise constant volvol γ and term structure ω parameters as

$$\gamma(t) := \begin{cases} \gamma_1 & \text{when } T_0 \leq t \leq T_1 \\ \gamma_2 & \text{when } T_1 < t \leq T_2 \\ \vdots & \vdots \\ \gamma_N & \text{when } T_{N-1} < t \leq T_N \end{cases} \quad \text{and } \omega(t) := \begin{cases} \omega_1 & \text{when } T_0 \leq t \leq T_1 \\ \omega_2 & \text{when } T_1 < t \leq T_2 \\ \vdots & \vdots \\ \omega_N & \text{when } T_{N-1} < t \leq T_N \end{cases}. \quad (\text{E.1a})$$

Define left L_i and right R_i integrals till expiry T_i for $i \in \{1, 2, \dots, N\}$ as

$$L_i := \int_0^{T_i} \omega^2(t) \left[\int_0^t \omega^2(s) \exp \left(6 \int_0^s \gamma^2(u) du + \int_s^t \gamma^2(u) du \right) ds \right] dt, \quad \text{and} \quad (\text{E.1b})$$

$$R_i := \int_0^{T_i} \omega^2(t) \exp \left(\int_0^t \gamma^2(u) du \right) dt. \quad (\text{E.1c})$$

Both can be analytically solved due to the assumption of piecewise constant parameters. Let's start with right integral R_i . When $T_{j-1} \leq t \leq T_j$, the integrand can be expressed as

$$\omega^2(t) \exp \left(\int_0^t \gamma^2(u) du \right) = \omega_j^2 \exp \left(\sum_{u=1}^{j-1} \gamma_u^2 (T_u - T_{u-1}) + \gamma_j^2 (t - T_{j-1}) \right). \quad (\text{E.1d})$$

So right integral R_i becomes

$$R_i = \int_0^{T_i} \omega^2(t) \exp \left(\int_0^t \gamma^2(u) du \right) dt, \quad (\text{E.1e})$$

$$= \sum_{j=1}^i \int_{T_{j-1}}^{T_j} \omega_j^2 \exp \left(\sum_{u=1}^{j-1} \gamma_u^2 (T_u - T_{u-1}) + \gamma_j^2 (t - T_{j-1}) \right) dt, \quad (\text{E.1f})$$

$$= \sum_{j=1}^i \omega_j^2 \exp \left(\sum_{u=1}^{j-1} \gamma_u^2 (T_u - T_{u-1}) - \gamma_j^2 T_{j-1} \right) \int_{T_{j-1}}^{T_j} e^{\gamma_j^2 t} dt, \quad (\text{E.1g})$$

$$= \sum_{j=1}^i \frac{\omega_j^2}{\gamma_j^2} \exp \left(\sum_{u=1}^{j-1} \gamma_u^2 (T_u - T_{u-1}) - \gamma_j^2 T_{j-1} \right) \left(e^{\gamma_j^2 T_j} - e^{\gamma_j^2 T_{j-1}} \right). \quad (\text{E.1h})$$

For $T_{p-1} \leq s \leq T_p \leq T_j$ and $T_{j-1} \leq t \leq T_j$ the inner integrand of L_i becomes

$$\omega^2(s) \exp \left(6 \int_0^s \gamma^2(u) du + \int_s^t \gamma^2(u) du \right) \quad (\text{E.1i})$$

$$= \omega_p^2 \exp \left(6 \left(\sum_{u=1}^{p-1} \gamma_u^2(T_u - T_{u-1}) + \gamma_p^2(s - T_{p-1}) \right) + \gamma_p^2(T_p - s) + \sum_{u=p+1}^{j-1} \gamma_u^2(T_u - T_{u-1}) + \gamma_j^2(t - T_{j-1}) \right), \quad (\text{E.1j})$$

$$= \omega_p^2 e^{5\gamma_p^2 s} \exp \left(6 \left(\sum_{u=1}^{p-1} \gamma_u^2(T_u - T_{u-1}) - \gamma_p^2 T_{p-1} \right) + \gamma_p^2 T_p + \sum_{u=p+1}^{j-1} \gamma_u^2(T_u - T_{u-1}) + \gamma_j^2(t - T_{j-1}) \right). \quad (\text{E.1k})$$

So the inner integral till T_j of L_i becomes

$$\sum_{p=1}^j \int_{T_{p-1}}^{T_p} \omega_p^2 e^{5\gamma_p^2 s} \exp \left(6 \left(\sum_{u=1}^{p-1} \gamma_u^2(T_u - T_{u-1}) - \gamma_p^2 T_{p-1} \right) + \gamma_p^2 T_p + \sum_{u=p+1}^{j-1} \gamma_u^2(T_u - T_{u-1}) + \gamma_j^2(t - T_{j-1}) \right) ds$$

$$= \sum_{p=1}^j \omega_p^2 \int_{T_{p-1}}^{T_p} e^{5\gamma_p^2 s} ds \quad (\text{E.1l})$$

$$\cdot \exp \left(6 \left(\sum_{u=1}^{p-1} \gamma_u^2(T_u - T_{u-1}) - \gamma_p^2 T_{p-1} \right) + \gamma_p^2 T_p + \sum_{u=p+1}^{j-1} \gamma_u^2(T_u - T_{u-1}) + \gamma_j^2(t - T_{j-1}) \right), \quad (\text{E.1m})$$

$$= \sum_{p=1}^j \frac{\omega_p^2}{5\gamma_p^2} \left(e^{5\gamma_p^2 T_p} - e^{5\gamma_p^2 T_{p-1}} \right) \quad (\text{E.1n})$$

$$\cdot \exp \left(6 \left(\sum_{u=1}^{p-1} \gamma_u^2(T_u - T_{u-1}) - \gamma_p^2 T_{p-1} \right) + \gamma_p^2 T_p + \sum_{u=p+1}^{j-1} \gamma_u^2(T_u - T_{u-1}) + \gamma_j^2(t - T_{j-1}) \right), \quad (\text{E.1o})$$

$$= \sum_{p=1}^j \frac{\omega_p^2}{5\gamma_p^2} e^{\gamma_j^2 t} \left(e^{5\gamma_p^2 T_p} - e^{5\gamma_p^2 T_{p-1}} \right) \quad (\text{E.1p})$$

$$\cdot \exp \left(6 \left(\sum_{u=1}^{p-1} \gamma_u^2(T_u - T_{u-1}) - \gamma_p^2 T_{p-1} \right) + \gamma_p^2 T_p + \sum_{u=p+1}^{j-1} \gamma_u^2(T_u - T_{u-1}) - \gamma_j^2 T_{j-1} \right). \quad (\text{E.1q})$$

And finally L_i is

$$L_i = \sum_{j=1}^i \frac{\omega_j^2}{\gamma_j^2} \left(e^{\gamma_j^2 T_j} - e^{\gamma_j^2 T_{j-1}} \right) \quad (\text{E.1r})$$

$$\cdot \sum_{p=1}^{j-1} \frac{\omega_p^2}{5\gamma_p^2} e^{6(\sum_{u=1}^{p-1} \gamma_u^2(T_u - T_{u-1}) - \gamma_p^2 T_{p-1}) + \gamma_p^2 T_p + \sum_{u=p+1}^{j-1} \gamma_u^2(T_u - T_{u-1}) - \gamma_j^2 T_{j-1}} \left(e^{5\gamma_p^2 T_p} - e^{5\gamma_p^2 T_{p-1}} \right) \quad (\text{E.1s})$$

$$+ \sum_{j=1}^i \frac{\omega_j^4}{35\gamma_j^2} \left(e^{7\gamma_j^2 T_j} - e^{7\gamma_j^2 T_{j-1}} \right) e^{6(\sum_{u=1}^{j-1} \gamma_u^2(T_u - T_{u-1}) - \gamma_j^2 T_{j-1}) - \gamma_j^2 T_{j-1}} \quad (\text{E.1t})$$

$$- \sum_{j=1}^i \frac{\omega_j^4}{10\gamma_j^2} \left(e^{2\gamma_j^2 T_j} - e^{2\gamma_j^2 T_{j-1}} \right) e^{5\gamma_j^2 T_{j-1}} e^{6(\sum_{u=1}^{j-1} \gamma_u^2(T_u - T_{u-1}) - \gamma_j^2 T_{j-1}) - \gamma_j^2 T_{j-1}}. \quad (\text{E.1u})$$

Appendix F

Recovery procedure of the characteristic function

In order to calibrate time-dependent term structure parameter $\omega(t)$ the characteristic function of stochastic variable Y_M needs to be determined. This characteristic function is obtained by an iterative procedure [15, 36]. For completion, this procedure is explained here in the appendix, after introducing some necessary definitions and deriving properties of those definitions. Define

$$R_j := \log \left(\frac{\omega_1^2(t_j)\sigma^2(t_j)}{\omega_1^2(t_{j-1})\sigma^2(t_{j-1})} \right) \text{ and} \quad (\text{F.1a})$$

$$Y_j := \log \left(\frac{\omega_1^2(t_{M-j+1})\sigma^2(t_{M-j+1})}{\omega_1^2(t_{M-j})\sigma^2(t_{M-j})} + \frac{\omega_1^2(t_{M-j+2})\sigma^2(t_{M-j+2})}{\omega_1^2(t_{M-j})\sigma^2(t_{M-j})} + \dots + \frac{\omega_1^2(t_M)\sigma^2(t_M)}{\omega_1^2(t_{M-j})\sigma^2(t_{M-j})} \right),$$

$$= \log(\exp(Y_{j-1}) + 1) + R_{M-j+1}, \quad (\text{F.1b})$$

$$= Z_{j-1} + R_{M-j+1}, \text{ where } Z_j := \log(\exp(Y_j) + 1). \quad (\text{F.1c})$$

Since Z_{j-1} and R_{M-j+1} are independent, the following relation for their characteristic functions ϕ hold:

$$\phi_{Y_j}(u) = \phi_{Z_{j-1}}(u)\phi_{R_{M-j+1}}(u). \quad (\text{F.1d})$$

By the definition of the characteristic function the following holds

$$\phi_{Z_{j-1}}(u) := \mathbb{E} \left[e^{iu \log(\exp(Y_{j-1})+1)} \right] = \mathbb{E} \left[(\exp(Y_{j-1}) + 1)^{iu} \right], \quad (\text{F.1e})$$

$$= \int_{-\infty}^{\infty} (e^y + 1)^{iu} f_{Y_{j-1}}(y) dy. \quad (\text{F.1f})$$

The integration range is truncated to $[a, b]$ to apply the Fourier cosine series expansion to approximate the characteristic function

$$\hat{\phi}_{Z_{j-1}}(u) = \int_a^b (e^y + 1)^{iu} f_{Y_{j-1}}(y) dy. \quad (\text{F.1g})$$

It will be explained later what a good choice for interval $[a, b]$ is. The Fourier cosine expansion of $f_{Y_{j-1}}$ is

$$f_{Y_{j-1}}(y) = \sum_{k=0}^{\infty} A_k \cos((y-a)u_k), \quad (\text{F.1h})$$

$$\text{where } A_k := \frac{2}{b-a} \int_a^b f_{Y_{j-1}}(y) \cos((y-a)u_k) dy \text{ and } u_k := \frac{k\pi}{b-a}. \quad (\text{F.1i})$$

The prime indicates that the first term in the sum is multiplied by a half. Calculating the infinite sum in (F.1h) is not feasible in practice. Plugging in a truncated version of this sum leads to

$$\hat{\phi}_{Z_{j-1}}(u) = \int_a^b (e^y + 1)^{iu} \sum_{k=0}^{N-1} A_k \cos((y-a)u_k) dy, \quad (\text{F.1j})$$

$$= \sum_{k=0}^{N-1} A_k \int_a^b (e^y + 1)^{iu} \cos((y-a)u_k) dy, \quad (\text{F.1k})$$

$$= \sum_{k=0}^{N-1} A_k \int_a^b (e^y + 1)^{iu} \cos((y-a)u_k) dy, \quad (\text{F.1l})$$

$$= \frac{2}{b-a} \sum_{l=0}^{N-1} \Re \left(\hat{\phi}_{Y_{j-1}}(u_l) e^{-iau_l} \right) \cdot \int_a^b (e^y + 1)^{iu} \cos((y-a)u_l) dy, \quad (\text{F.1m})$$

where $\hat{\phi}_{Y_{j-1}}$ is the approximation of $\phi_{Y_{j-1}}$. The last equation results from the following observations

$$\phi_{Y_{j-1}}(u) := \mathbb{E} [e^{iuY_{j-1}}] = \int_a^b e^{iuy} f_{Y_{j-1}}(y) dy, \text{ so} \quad (\text{F.1n})$$

$$e^{-iau_k} \phi_{Y_{j-1}}(u_k) = e^{-iau_k} \int_a^b e^{iyu_k} f_{Y_{j-1}}(y) dy = \int_a^b e^{i(y-a)u_k} f_{Y_{j-1}}(y) dy, \quad (\text{F.1o})$$

$$= \int_a^b \cos((y-a)u_k) f_{Y_{j-1}}(y) dy + i \int_a^b \sin((y-a)u_k) f_{Y_{j-1}}(y) dy, \quad (\text{F.1p})$$

by Euler's formula. So

$$\Re (e^{-iau_k} \phi_{Y_{j-1}}(u_k)) = \int_a^b \cos((y-a)u_k) f_{Y_{j-1}}(y) dy = \frac{b-a}{2} \cdot A_k. \quad (\text{F.1q})$$

Formula (F.1m) is used in the calculation of characteristic function $\hat{\phi}_{Z_{j-1}}$.

Integral $\int_a^b (e^y + 1)^{iu} \cos((y-a)u_l) dy$ in this characteristic function can be approximated by using the Clenshaw-Curtis quadrature method [36] which is suited for oscillating functions especially, or it can be calculated with simple trapezoidal integration.

Calculating characteristic function $\hat{\phi}_{Z_{j-1}}$ is clear now. Also ϕ_{R_j} needs to be known before iterative relation (F.1d) is used in an algorithm to determine the characteristic function of Y_j . The characteristic function of R_j is known analytically, since R_j is normally distributed:

$$R_j \sim \mathcal{N}(\mu_j, \sigma_j^2), \mu_j := \log \left(\frac{\omega_1^2(t_j)}{\omega_1^2(t_{j-1})} \right) - \int_{t_{j-1}}^{t_j} \gamma^2(s) ds \text{ and } \sigma_j^2 = 4 \int_{t_{j-1}}^{t_j} \gamma^2(s) ds. \quad (\text{F.1r})$$

So, its characteristic function is

$$\phi_{R_j}(u) = e^{iu\mu_j - u^2\sigma_j^2/2}. \quad (\text{F.1s})$$

Iterative relation (F.1d), formula (F.1m) and (F.1s) are combined to form an algorithm that calculates characteristic function $\hat{\phi}_{Y_M}$.

Recovery procedure

The characteristic function ϕ_{Y_M} is obtained by executing the following recovery procedure:

1. Since $Y_1 = R_M$, characteristic function $\phi_{Y_1}(u_k)$ can be calculated analytically.
2. For $j \in \{1, 2, \dots, M-2\}$ determine $\phi_{Y_{j+1}}$ in points u_k , $k \in \{0, 1, \dots, N-1\}$, by using relation (F.1d): $\phi_{Y_{j+1}}(u_k) = \phi_{R_{M-j}}(u_k)\phi_{Z_j}(u_k)$. Iterate $\hat{\phi}_{Y_{j+1}}(u_k) := \phi_{R_{M-j}}(u_k)\hat{\phi}_{Z_j}(u_k)$.
3. The previous iteration led to $\hat{\phi}_{Y_{M-1}}$. Now obtain $\hat{\phi}_M(-\frac{1}{2}i)$ by calculating $\hat{\phi}_{Y_M}(-\frac{1}{2}i) := \phi_{R_1}(-\frac{1}{2}i)\hat{\phi}_{Z_{M-1}}(-\frac{1}{2}i)$.

Characteristic function $\hat{\phi}_{Y_M}$ is obtained by executing this algorithm. The expression of the term structure mapping also contains the characteristic function $\phi_{\tilde{Y}_M}$ of \tilde{Y}_M , where

$$\tilde{Y}_M := \log \left(\sum_{j=1}^M \tilde{\sigma}^2(t_j) \right). \quad (\text{F.2a})$$

The recovery procedure is applied to certain definitions of \tilde{Y}_j , \tilde{R}_j and \tilde{Z}_j to obtain $\phi_{\tilde{Y}_M}$. For general $j = 1, \dots, M$ define \tilde{Y}_j and \tilde{R}_j as

$$\tilde{Y}_j := \log \left(\frac{\tilde{\sigma}^2(t_{M-j+1})}{\tilde{\sigma}^2(t_{M-j})} + \dots + \frac{\tilde{\sigma}^2(t_M)}{\tilde{\sigma}^2(t_{M-j})} \right) \text{ and } \tilde{R}_j := \log \left(\frac{\tilde{\sigma}^2(t_j)}{\tilde{\sigma}^2(t_{j-1})} \right). \quad (\text{F.2b})$$

Then

$$\tilde{Y}_j = \tilde{R}_{M-j+1} + \log \left(1 + \exp(\tilde{Y}_{j-1}) \right) =: \tilde{R}_{M-j+1} + \tilde{Z}_{j-1} \text{ and} \quad (\text{F.2c})$$

$$\tilde{R}_j \sim \mathcal{N} \left(-\tilde{\gamma}^2(t_j - t_{j-1}), 4\tilde{\gamma}^2(t_j - t_{j-1}) \right), \text{ so } \phi_{\tilde{R}_j}(u) = \exp \left(-\tilde{\gamma}^2(t_j - t_{j-1})(iu + 2u^2) \right), \quad (\text{F.2d})$$

where \tilde{R}_{M-j+1} and \tilde{Z}_{j-1} are independent for all j , due to independent increments of the underlying Brownian motion. So the application of the recovery procedure to \tilde{Y}_j and \tilde{Z}_{j-1} results in $\phi_{\tilde{Y}_M}$. At this point the method to calculate ϕ_{Y_M} and $\phi_{\tilde{Y}_M}$ is clear, except for the choice of truncated integration range $[a, b]$. This is explained in the rest of this appendix.

Integration range

In order to be able to apply the Fourier series expansion to approximate the characteristic function, the integration was truncated from the real line to range $[a, b]$ in (F.1m). This truncation introduces an error, that can be controlled by choosing the boundaries a, b wisely. In [36] the integration range for each Y_j is determined by its cumulants:

$$\left[\zeta_1(Y_j) - L\sqrt{\zeta_2(Y_j) + \sqrt{\zeta_4(Y_j)}}, \zeta_1(Y_j) + L\sqrt{\zeta_2(Y_j) + \sqrt{\zeta_4(Y_j)}} \right], \quad (\text{F.3a})$$

where ζ_n is the n -th cumulant. Determining the cumulant of Y_j is a computationally expensive operation, therefore the boundaries are expressed in terms of R_j instead of Y_j by using the following identities

$$\zeta_1 \left(\log \left(j \frac{S_{M-j+1}}{S_{M-j}} \right) \right) \leq \zeta_1(\exp(Y_j)) \leq \zeta_1 \left(\log \left(j \frac{S_M}{S_{M-j}} \right) \right), \quad (\text{F.3b})$$

$$0 \leq \zeta_2(\exp(Y_j)) \leq \zeta_2 \left(\log \left(j \frac{S_M}{S_{M-j}} \right) \right), \quad (\text{F.3c})$$

$$0 \leq \zeta_4(\exp(Y_j)) \leq \zeta_4 \left(\log \left(j \frac{S_M}{S_{M-j}} \right) \right), \quad (\text{F.3d})$$

where $S_i := \tilde{\sigma}_1^2 \omega^2(t_i), i = 1, \dots, M$. Define

$$a_j := \zeta_1 \left(\log \left(j \frac{S_{M-j+1}}{S_{M-j}} \right) \right) - L \sqrt{\zeta_2 \left(\log \left(j \frac{S_M}{S_{M-j}} \right) \right) + \zeta_4 \left(\log \left(j \frac{S_M}{S_{M-j}} \right) \right)}, \quad (\text{F.3e})$$

$$b_j := \zeta_1 \left(\log \left(j \frac{S_M}{S_{M-j}} \right) \right) + L \sqrt{\zeta_2 \left(\log \left(j \frac{S_M}{S_{M-j}} \right) \right) + \zeta_4 \left(\log \left(j \frac{S_M}{S_{M-j}} \right) \right)}. \quad (\text{F.3f})$$

$$(\text{F.3g})$$

Although $\zeta(\log_n(\cdot)) \neq \log(\zeta_n(\cdot))$, when $L \rightarrow \infty$ the truncation error goes to zero. Note that

$$\zeta_1 \left(\log \left(j \frac{S_{M-j+1}}{S_{M-j}} \right) \right) = \zeta_1 \left(\log \left(\frac{S_{M-j+1}}{S_{M-j}} \right) + \log(j) \right) = \zeta_1(R_1) + \log(j), \quad (\text{F.3h})$$

$$\zeta_1 \left(\log \left(j \frac{S_M}{S_{M-j}} \right) \right) = \zeta_1 \left(\log \left(\frac{S_M}{S_{M-j}} \right) \right) + \log(j), \quad (\text{F.3i})$$

$$= \zeta_1 \left(\log \left(\frac{S_{M-j+1}}{S_{M-j}} \frac{S_{M-j+2}}{S_{M-j+1}} \dots \frac{S_M}{S_{M-1}} \right) \right) + \log(j) = j \zeta_1(R_1) + \log(j). \quad (\text{F.3j})$$

Furthermore, for all $n \geq 2$ we have

$$\zeta_n \left(\log \left(j \frac{S_{M-j+1}}{S_{M-j}} \right) \right) = \zeta_n(R_1) \quad \text{and} \quad \zeta_n \left(\log \left(j \frac{S_M}{S_{M-j}} \right) \right) = j \zeta_n(R_1). \quad (\text{F.3k})$$

Since

$$R_1 = \log \left(\frac{S_1}{S_0} \right) = \log \left(\frac{\sigma^2(t_1)}{\sigma^2(t_0)} \right) = \log(\sigma^2(\Delta_t)), \quad \text{with} \quad d\sigma^2(t) = \sigma^2(t) (\tilde{\gamma}^2 dt + 2\tilde{\gamma} dW_t),$$

$$\text{so } \zeta_1(R_1) = \zeta_1 \left(\left(\tilde{\gamma}^2 - \frac{(2\tilde{\gamma})^2}{2} \right) \Delta_t + 2\tilde{\gamma} W_{\Delta_t} \right) = \zeta_1(-\tilde{\gamma}^2 \Delta_t + 2\tilde{\gamma} W_{\Delta_t}) = -\tilde{\gamma}^2 \Delta_t, \quad (\text{F.3l})$$

$$\text{and } \zeta_2(R_1) = 4\tilde{\gamma}^2 \Delta_t, \zeta_n(R_1) = 0, \forall n > 2. \quad (\text{F.3m})$$

This results in

$$a_j = -\tilde{\gamma}^2 \Delta_t + \log(j) - L \sqrt{j \cdot 4\tilde{\gamma}^2 \Delta_t}, \quad (\text{F.3n})$$

$$b_j = -j\tilde{\gamma}^2 \Delta_t + \log(j) + L \sqrt{j \cdot 4\tilde{\gamma}^2 \Delta_t}. \quad (\text{F.3o})$$

To make sure the integration has to be performed just once for all timesteps, one range is used that includes all other ranges:

$$[a, b] := [\min_j a_j, \max_j b_j] \approx [-\tilde{\gamma}^2 \Delta_t - L \sqrt{M \cdot 4\tilde{\gamma}^2 \Delta_t}, -M\tilde{\gamma}^2 \Delta_t + \log(M) + L \sqrt{M \cdot 4\tilde{\gamma}^2 \Delta_t}]. \quad (\text{F.3p})$$

Appendix G

Local and stochastic local volatility

This appendix contains the derivation of an expression of the local volatility component $\sigma_{SLV}(t, K)$ in terms of Dupire local volatility $\sigma_{LV}(t, K)$ and a conditional expectation. First, define (deterministic) zero coupon bond

$$P(0, T) := e^{-\int_0^T r(s)ds}, \quad (\text{F.3aa})$$

and European-type call option

$$C(T, K) = \mathbb{E} [P(0, T)(S(T) - K)^+ | \mathcal{F}_0] = P(0, T) \int_K^\infty (s - K)\phi(T, s)ds, \quad (\text{F.3ab})$$

where $S(t)$ is the spot at time t . The differential of this call option is

$$dC(t, K) = dP(0, T)\mathbb{E} [(S(t) - K)^+ | \mathcal{F}_0] + P(0, T)d\mathbb{E} [(S(t) - K)^+ | \mathcal{F}_0], \quad (\text{F.3ac})$$

$$= -r(t)P(0, T)\mathbb{E} [(S(t) - K)^+ | \mathcal{F}_0] dt + P(0, T)\mathbb{E} [d(S(t) - K)^+ | \mathcal{F}_0], \quad (\text{F.3ad})$$

by Fubini. Since $(S(t) - K)^+$ is not differentiable in $S(t) = K$, Ito's lemma cannot be applied to this expression. The generalization of Ito's lemma, the Tanaka-Meyer formula, can be used to solve this expression. Applying Tanaka-Meyer yields:

$$d(S(t) - K)^+ = \mathbb{1}_{S(t) > K} dS(t) + \frac{1}{2} \delta(S(t) - K) (dS(t))^2. \quad (\text{F.3ae})$$

Substitute the following dynamics for $S(t)$:

$$dS(t) = r(t)S(t)dt + \sigma_{SLV}(t, S(t))\psi(S(t), V(t))S(t)dW(t), \quad (\text{F.3af})$$

where $\sigma_{SLV}(t, S(t))$ is the local volatility component and $\psi(S(t), V(t))$ is a function of spot $S(t)$ and a stochastic volatility process $V(t)$, which is correlated with $S(t)$. This choice for differential $dS(t)$ leads to an expression for the differential of $(S(t) - K)^+$:

$$d(S(t) - K)^+ = \mathbb{1}_{S(t) > K} \cdot [r(t)S(t)dt + \sigma_{SLV}(t, S(t))\psi(S(t), V(t))S(t)dW(t)] \quad (\text{F.3ag})$$

$$+ \frac{1}{2} \delta(S(t) - K) \sigma_{SLV}^2(t, S(t)) \psi^2(S(t), V(t)) S^2(t) dt. \quad (\text{F.3ah})$$

Function δ is Dirac's delta function. So

$$\begin{aligned} dC(t, K) &= -r(t)P(0, T)\mathbb{E} [(S(t) - K)^+ | \mathcal{F}_0] dt \\ &+ P(0, T)\mathbb{E} [\mathbb{1}_{S(t) > K} \cdot [r(t)S(t)dt + \sigma_{SLV}(t, S(t))\psi(S(t), V(t))S(t)dW(t)] | \mathcal{F}_0] \\ &+ \frac{1}{2} P(0, T)\mathbb{E} [\delta(S(t) - K) \sigma_{SLV}^2(t, S(t)) \psi^2(S(t), V(t)) S^2(t) dt | \mathcal{F}_0], \end{aligned} \quad (\text{F.3ai})$$

$$\begin{aligned} &= -r(t)P(0, T)\mathbb{E} [(S(t) - K)^+ | \mathcal{F}_0] dt + r(t)P(0, T)\mathbb{E} [\mathbb{1}_{S(t) > K} S(t) | \mathcal{F}_0] dt \\ &+ \frac{1}{2} P(0, T)\mathbb{E} [\delta(S(t) - K) \sigma_{SLV}^2(t, S(t)) \psi^2(S(t), V(t)) S^2(t) | \mathcal{F}_0] dt, \end{aligned} \quad (\text{F.3aj})$$

where the last step is due to martingality of the Ito integral. When using $(S(t) - K)^+ = \mathbf{1}_{S(t) > K}(S(t) - K)$, above reduces to

$$\begin{aligned} dC(t, K) &= r(t)K P(0, T) \mathbb{E} [\mathbf{1}_{S(t) > K} | \mathcal{F}_0] dt \\ &\quad + \frac{1}{2} P(0, T) \mathbb{E} [\delta(S(t) - K) \sigma_{SLV}^2(t, S(t)) \psi^2(S(t), V(t)) S^2(t) | \mathcal{F}_0] dt. \end{aligned} \quad (\text{F.3ak})$$

To simplify this equation further, note that

$$\frac{\partial C(t, K)}{\partial K} = -P(0, T) \mathbb{E} [\mathbf{1}_{S(t) > K} | \mathcal{F}_0] \quad \text{and} \quad \frac{\partial^2 C(t, K)}{\partial K^2} = P(0, T) \phi(K). \quad (\text{F.3al})$$

Also note that

$$\mathbb{E} [\delta(S(t) - K) \sigma_{SLV}^2(t, S(t)) \psi^2(S(t), V(t)) S^2(t) | \mathcal{F}_0] \quad (\text{F.3am})$$

$$= \int_{\mathbb{R}} \int_{\mathbb{R}} \delta(s - K) \sigma_{SLV}^2(t, s) \psi^2(s, v) y^2 f_{V,y}(v, s) ds dv, \quad (\text{F.3an})$$

$$= \int_{\mathbb{R}} \sigma_{SLV}^2(t, K) \psi^2(K, v) K^2 f_{V,s}(v, K) dv, = K^2 \sigma_{SLV}^2(t, K) \int_{\mathbb{R}} \psi^2(K, v) f_{V,s}(v, K) dv, \quad (\text{F.3ao})$$

$$= K^2 \sigma_{SLV}^2(t, K) \int_{\mathbb{R}} \psi^2(K, v) f_{V|s=K}(v, K) \phi(K) dv, \quad (\text{F.3ap})$$

$$= K^2 \sigma_{SLV}^2(t, K) \phi(K) \mathbb{E} [\psi^2(S(t), V(t)) | y(t) = K]. \quad (\text{F.3aq})$$

In conclusion,

$$dC(t, K) = -r(t)K \frac{\partial C(t, K)}{\partial K} dt + \frac{1}{2} K^2 \sigma_{SLV}^2(t, K) \frac{\partial^2 C(t, K)}{\partial K^2} \mathbb{E} [\psi^2(S(t), V(t)) | S(t) = K] dt.$$

Reordering terms leads to an expression for the local volatility component σ_{SLV} :

$$\sigma_{SLV}^2(t, K) = \frac{\frac{\partial C(t, K)}{\partial t} + r(t)K \frac{\partial C(t, K)}{\partial K}}{\frac{1}{2} K^2 \frac{\partial^2 C(t, K)}{\partial K^2} \mathbb{E} [\psi^2(S(t), V(t)) | S(t) = K]}, \quad (\text{F.3ar})$$

$$= \frac{\sigma_{LV}(t, K)}{\mathbb{E} [\psi^2(S(t), V(t)) | S(t) = K]}, \quad (\text{F.3as})$$

where σ_{LV} is the Dupire local volatility. Conditional expectation $\mathbb{E} [\psi^2(S(t), V(t)) | S(t) = K]$ can be estimated with a technique introduced in [14].

0.1 Dupire's formula

Consider the generalization of a geometric Brownian motion, where the volatility σ_{LV} is a function of the spot $S(t)$ as well as time t and where the drift $r(t)$ is a time dependent function:

$$dS_t = r(t)S_t dt + \sigma_{LV}(t, S_t) S_t dW_t. \quad (\text{F.3ba})$$

Function σ_{LV} is known as the local volatility function in the literature. Dupire [17] derived an expression for the local volatility in terms of (derivatives of) call prices. This derivation is shown here. In order to do this, the (deterministic) zero coupon bond P and the European-type call price are again defined as

$$P(0, T) = e^{-\int_0^T r(s) ds}, \quad \text{and} \quad (\text{F.3bb})$$

$$C(T, K) = \mathbb{E} [P(0, T)(S(T) - K)^+ | \mathcal{F}_0] = P(0, T) \int_K^\infty (s - K) \phi(T, s) ds. \quad (\text{F.3bc})$$

Function $\phi(T, s)$ is the risk-neutral probability density function of the spot at expiry T . Its time evolution is described by the Fokker-Planck equation

$$\frac{\partial \phi(t, s)}{\partial t} = \frac{1}{2} \frac{\partial^2}{\partial s^2} [s^2 \sigma^2(t, s) \phi(t, s)] - \frac{\partial}{\partial s} [sr(t) \phi(t, s)]. \quad (\text{F.3bd})$$

This result can be used in the derivative $\frac{\partial C}{\partial T}$ to obtain an expression for local volatility function σ_{LV} in terms of derivatives of call prices:

$$\frac{\partial C}{\partial T} = -r(T)C + P(0, T) \int_K^\infty (s - K) \frac{\partial \phi(T, s)}{\partial T} ds, \quad (\text{F.3be})$$

$$= -r(T)C + P(0, T) \int_K^\infty (s - K) \left(\frac{1}{2} \frac{\partial^2}{\partial s^2} [s^2 \sigma_{LV}^2(T, s) \phi(T, s)] - r(T) \frac{\partial}{\partial s} [s \phi(T, s)] \right) ds, \quad (\text{F.3bf})$$

$$= -r(T)C + \frac{1}{2} P(0, T) \left(\left[(s - K) \frac{\partial}{\partial s} [s^2 \sigma_{LV}^2(T, s) \phi(T, s)] \right]_{s=K}^{s=\infty} \right. \quad (\text{F.3bg})$$

$$\left. - \int_K^\infty \frac{\partial}{\partial s} [s^2 \sigma_{LV}^2(T, s) \phi(T, s)] ds \right) - P(0, T) r(T) \left(\left[(s - K) s \phi(T, s) \right]_{s=K}^{s=\infty} - \int_K^\infty s \phi(T, s) ds \right),$$

$$= -r(T)C - \frac{1}{2} P(0, T) \int_K^\infty \frac{\partial}{\partial s} [s^2 \sigma_{LV}^2(T, s) \phi(T, s)] ds + P(0, T) r(T) \int_K^\infty s \phi(T, s) ds, \quad (\text{F.3bh})$$

$$= -r(T)C - \frac{1}{2} P(0, T) [s^2 \sigma_{LV}^2(T, s) \phi(T, s)]_{s=K}^{s=\infty} + P(0, T) r(T) \int_K^\infty s \phi(T, s) ds, \quad (\text{F.3bi})$$

$$= -r(T)C + \frac{1}{2} P(0, T) K^2 \sigma_{LV}^2(T, K) \phi(T, K) + P(0, T) r(T) \int_K^\infty ((s - K) + K) \phi(T, s) ds, \quad (\text{F.3bj})$$

$$= -r(T)C + \frac{1}{2} P(0, T) K^2 \sigma_{LV}^2(T, K) \phi(T, K) + r(T) \left(C + P(0, T) K \int_K^\infty \phi(T, s) ds \right), \quad (\text{F.3bk})$$

$$= -r(T)C + \frac{1}{2} K^2 \sigma_{LV}^2(T, K) \frac{\partial^2 C}{\partial K^2} + r(T) \left(C - K \frac{\partial C}{\partial K} \right). \quad (\text{F.3bl})$$

Rewriting this result leads to an expression for the local volatility:

$$\sigma_{LV}^2(T, K) = \frac{\frac{\partial C}{\partial T} + r(T) K \frac{\partial C}{\partial K}}{\frac{1}{2} K^2 \frac{\partial^2 C}{\partial K^2}}. \quad (\text{F.3bm})$$

0.2 Dupire's formula expressed in implied volatilities

Above Dupire equation suffers from numerical problems in practice because the divisor can be unstable due to the finite difference approximation of the second order derivative. Van der Kamp [12] derived a more practical expression of the local volatility in terms of implied volatilities. The derivation is applied to the FX context here. First, define parameters

$$y := \ln \left(\frac{K}{F_{0,T}} \right) \text{ and } w = \Sigma^2 T, \text{ where } F_{0,T} = S_0 e^{\int_0^T (r_d(s) - r_f(s)) ds}, \quad (\text{F.3ca})$$

is the current FX forward rate and S_0 the current spot FX rate. P_d is the (deterministic) domestic discount curve $P_d(0, T) = e^{-\int_0^T r_d(s) ds}$. A European-type call price observed in the market, is equal to the call price produced by Black's model when the right implied volatility $\Sigma(T, K)$ is inserted.

$$C(K, T) = P_d(0, T) [F_{0,T} \Phi(d_1) - K \Phi(d_2)], \text{ where} \quad (\text{F.3cb})$$

$$d_1 = -\frac{y}{\sqrt{w}} + \frac{\sqrt{w}}{2} \text{ and } d_2 = -\frac{y}{\sqrt{w}} - \frac{\sqrt{w}}{2}. \quad (\text{F.3cc})$$

In order to express the Dupire equations in implied volatilities, we need to calculate the derivatives with respect to the parameters of Blacks model:

$$\frac{\partial C}{\partial K} = \frac{\partial C}{\partial y} \frac{\partial y}{\partial K} + \frac{\partial C}{\partial w} \frac{\partial w}{\partial K} = \frac{1}{K} \frac{\partial C}{\partial y} + \frac{\partial w}{\partial K} \frac{\partial C}{\partial w}, \quad (\text{F.3cd})$$

$$\frac{\partial^2 C}{\partial K^2} = -\frac{1}{K^2} \frac{\partial C}{\partial y} + \frac{1}{K} \frac{\partial}{\partial K} \left(\frac{\partial C}{\partial y} \right) + \frac{\partial^2 w}{\partial K^2} \frac{\partial C}{\partial w} + \frac{\partial w}{\partial K} \frac{\partial}{\partial K} \left(\frac{\partial C}{\partial w} \right), \quad (\text{F.3ce})$$

$$\begin{aligned} &= -\frac{1}{K^2} \frac{\partial C}{\partial y} + \frac{1}{K} \left(\frac{1}{K} \frac{\partial^2 C}{\partial y^2} + \frac{\partial w}{\partial K} \frac{\partial^2 C}{\partial w \partial y} \right) + \frac{\partial^2 w}{\partial K^2} \frac{\partial C}{\partial w} + \frac{\partial w}{\partial K} \left(\frac{1}{K} \frac{\partial^2 C}{\partial y \partial w} + \frac{\partial w}{\partial K} \frac{\partial^2 C}{\partial w^2} \right), \\ &= \frac{1}{K^2} \left(\frac{\partial^2 C}{\partial y^2} - \frac{\partial C}{\partial y} \right) + \frac{2}{K} \frac{\partial w}{\partial K} \frac{\partial^2 C}{\partial w \partial y} + \frac{\partial^2 w}{\partial K^2} \frac{\partial C}{\partial w} + \left(\frac{\partial w}{\partial K} \right)^2 \frac{\partial^2 C}{\partial w^2}. \end{aligned} \quad (\text{F.3cf})$$

Plugging in these derivatives results in a Dupire equation expressed in call prices according to Black's model with parameters w and y :

$$\sigma_{LV}^2(T, K) = 2 \frac{-r_f(T)C - r(T) \frac{\partial C}{\partial y} + \frac{\partial w}{\partial T} \frac{\partial C}{\partial w} + r(T) \frac{\partial C}{\partial y} + r(T)K \frac{\partial w}{\partial K} \frac{\partial C}{\partial w} + r_f(T)C}{\left(\frac{\partial^2 C}{\partial y^2} - \frac{\partial C}{\partial y} \right) + 2K \frac{\partial w}{\partial K} \frac{\partial^2 C}{\partial w \partial y} + K^2 \frac{\partial^2 w}{\partial K^2} \frac{\partial C}{\partial w} + K^2 \left(\frac{\partial w}{\partial K} \right)^2 \frac{\partial^2 C}{\partial w^2}}, \quad (\text{F.3cg})$$

where $r(T) := r_d(T) - r_f(T)$. Above equation can be simplified by making use of the following identities:

$$\frac{\partial^2 C}{\partial w^2} = \left(-\frac{1}{8} - \frac{1}{2w} + \frac{y^2}{2w^2} \right) \frac{\partial C}{\partial w}, \quad \frac{\partial^2 C}{\partial w \partial y} = \left(\frac{1}{2} - \frac{y}{w} \right) \frac{\partial C}{\partial w}, \quad \frac{\partial^2 C}{\partial y^2} - \frac{\partial C}{\partial y} = 2 \frac{\partial C}{\partial w}, \quad (\text{F.3ch})$$

resulting in the following expression of the local volatility function:

$$\sigma_{LV}^2(T, K) = \frac{\frac{\partial w}{\partial T} + (r_d(T) - r_f(T))K \frac{\partial w}{\partial K}}{1 + K \frac{\partial w}{\partial K} \left(\frac{1}{2} - \frac{y}{w} \right) + \frac{1}{2} K^2 \frac{\partial^2 w}{\partial K^2} - \frac{1}{4} K^2 \left(\frac{\partial w}{\partial K} \right)^2 \left(\frac{1}{4} + \frac{1}{w} - \frac{y^2}{w^2} \right)}. \quad (\text{F.3ci})$$

The partial derivatives of w can be expressed in terms of implied volatilities as

$$\frac{\partial w}{\partial K} = 2\Sigma T \frac{\partial \Sigma}{\partial K}, \quad \frac{\partial^2 w}{\partial K^2} = 2T \left(\frac{\partial \Sigma}{\partial K} \right)^2 + 2\Sigma T \frac{\partial^2 \Sigma}{\partial K^2}, \quad \frac{\partial w}{\partial T} = \Sigma^2 + 2\Sigma T \frac{\partial \Sigma}{\partial T}. \quad (\text{F.3cj})$$

Plugging in above equations results in

$$\sigma_{LV}^2(T, K) = \frac{\Sigma^2 + 2\Sigma T \left(\frac{\partial \Sigma}{\partial T} + (r_d(T) - r_f(T))K \frac{\partial \Sigma}{\partial K} \right)}{\left(1 - \frac{Ky}{\Sigma} \frac{\partial \Sigma}{\partial K} \right)^2 + K\Sigma T \left(\frac{\partial \Sigma}{\partial K} - \frac{1}{4} K\Sigma T \left(\frac{\partial \Sigma}{\partial K} \right)^2 + K \frac{\partial^2 \Sigma}{\partial K^2} \right)}, \quad (\text{F.3ck})$$

$$\text{where } \Sigma = \Sigma(T, K). \quad (\text{F.3cl})$$