SABR: A Stochastic Volatility Model in Practice

Thesis · June 2019

DOI: 10.13140/RG.2.2.35290.16325

4 authors, including:

Natalya Bogatyreva
Centre for Genomic Regulation
51 PUBLICATIONS **2,865** CITATIONS
SEE PROFILE

Sergio Rodriguez-Apolinar
Universidad EAN
5 PUBLICATIONS **2** CITATIONS
SEE PROFILE

Abraham Soldevilla
Universidad Peruana de Ciencias Aplicadas (UPC)
1 PUBLICATION **0** CITATIONS
SEE PROFILE

SABR: A Stochastic Volatility Model in Practice

Master Project

Authors:
Natalia Bogatyreva,
Rodrigo Grandez,
Sergio Rodriguez Apolinar,
Abraham Soldevilla
Supervisor: Dr. Elisa Alos

Master in Finance
Barcelona Graduate School of Economics
Pompeu Fabra University

June 28, 2019

Acknowledgements

We would like to express our gratitude to our adviser Elisa Alos for all her support during this master project as well as David García-Lorite who gave us the observed market implied volatility of vital importance in our work.

Natalia Bogatyreva,
Rodrigo Grandez,
Sergio Rodriguez Apolinar,
Abraham Soldevilla

Barcelona, June 28, 2019

Abstract

The Black and Scholes model (BS) assumes that the volatility of an asset is constant over the trading period. As a result, BS returns a flat volatility surface. This assumption fails to capture the asset's volatility dynamics (smile), which is particularly important if we want to price complex derivatives. Local Volatility (LV) models captures the volatility smile, but not the price dynamics. In this project, we study the SABR (Stochastic Alpha, Beta, Rho) model, a stochastic volatility (SV) model designed to describe the implied volatility (IV) surface capturing both the smile and price dynamics. We calibrate its parameters, compute the IV using its closed-form solution and Monte Carlo Methods (MCM), and compare our results with real market data. Finally, we discuss our results and their implementations to practitioners.

Contents

List of Figures

List of Tables

1 Introduction

In the Black and Scholes's (BS) model the implied volatility is assumed to be a constant, however if we use the fact that implied volatility can be obtained by matching the observed option prices from the market with the theoretical value under BS model, we get that options with different strikes require different volatilities to match their market prices which suggests that volatility is in general not a constant but is dependent on the strike and it forms the so called volatility smile or skew. There are several works that have addressed this problem. For example, Dupire (1994) developed a local volatility model to match precisely the observed market volatility smiles.
However, several papers have tested that the dynamics behaviour of the market smile predicted by local volatility models and it turned out to be opposite to the observed market behaviour which brings about hedging problem.(see Hagan et al. (2002)). As an alternative approach to compute the unobserved implied volatility, a number of other methods were subsequently developed that took into account the stochastic nature of volatility such as the Hull-White, Heston and finally the SABR model.
The SABR model is very popular and widely used among practitioners because it has a tractable form and allows modeling the market price and the market risks, including vanna risk and volga risk (introduced by SABR model and discussed in Subsection 3.1) to be computed immediately from BS formula. Furthermore, this model provides a close form solution for implied volatility that captures the correct dynamics of the market smile volatility, and thus yields stable hedges. The capabilities of this model is what motivates us in this work.
Our purpose is to study in detail the SABR model and determine its superior performance over alternative estimation techniques. In particular, this is what we intend to do:

- First, to calibrate the SABR model for each maturity and to investigate which calibration method works better for market data.
- Second, to compute the implied volatility under the SABR model for different maturity time given a set of market prices.
- Third, to compare this implied volatility with the market volatility.
- Four, to compute the implied volatility under Conditional Monte Carlo method for a single set of parameters obtained in the previous step.
- Fifth, to compare the volatility surface obtained by SABR and Monte Carlo method.
- Finally, to discuss some extensions of the SABR model

A review of literature and a brief account of local and stochastic volatility necessary for our purpose are given in Section 2. In Section 3 we give the theoretical framework of the SABR model, we do calibration under the SABR model and examine how implied

volatilities behave compared to implied volatility obtained from the market. Section 4 we explain the conditional montecarlo method to compute the implied volatility surface under the SABR model. Section 5 contains comparisons between implied volatility obtained from the Hagan's formula and the one obtained from Monte carlo method. Finally, in Section 6 we discuss an extension to deal with the issues found in the SABR model.

2 From Black-Scholes to Stochastic Volatility

2.1 Implied Volatility

For a European call option with strike K, maturity T, and underlying current price S_0, we can define The Black-Scholes solution for the option premium (C_0) as follows:

$$C_0(K, T, S_0) = C_{BS}(K, T, S_0, \sigma),$$

where C_{BS} denotes the Black-Scholes formula.
Since σ represents the volatility of the underlying asset, the BS model imposes this to be a constant across strikes and maturities, another parameter of the model, because the underlying asset is the same for all specifications of the options.
Volatility in this case is unobservable, but with market option prices it is possible to retrieve the implied volatility (IV) using the BS formula. Given BS assumptions one should expect this implicit value to be almost equal for different pairs of K and T. This is, assuming that asset prices follow the log-normal distribution of the BS model, which implies no heavy tails.
This assumption used to work well before the crash of October 1987 when most of the trading world both didn't know about the presence of heavy tails in asset prices distribution and there wasn't a generalized fear of potential crashes in the equity market (see chapter 20 of Hull (2018)). After this, the market realized that in fact, implied volatility was not constant across strikes (for a certain maturity). In fact, empirical results found that volatility exhibits either smiles (decreasing for low strikes, increasing for high strikes, IV is almost the same at the same distance from the at-the-money strike) or skews (IV is higher for low strikes) [Figure 1].

[Insert Figure 1 here]

If this is the case, then, the BS formula can be rewritten to take this phenomenon into account: given an observed European call option price, the implied volatility (σ_{C_0}) is defined to be the value of the volatility parameter that must go into the Black-Scholes formula to match this price. So, we can write,

$$C_0(K, T, S_0) = C_{BS}(K, T, S_0, \sigma(K, T, S_0)),$$

The implied volatility is then defined as

$$\sigma_{C_0} = C_{BS}^{-1}(K, T, S_0, C_0(K, T, S_0)) \tag{1}$$

Depending on the type of option, the C_{BS}^{-1} function has not a closed form solution for implied volatility so we need to use numerical methods to find the zero of the function or a mathematical model that succeeds in capturing the behavior of market volatility.
As highlighted before, the importance of modeling the behavior of the IV lies not only in its role in option pricing but also in risk management strategies (hedging). In particular,

think about delta-hedging with an option whose IV is not constant:

$$\Delta = \frac{\partial C}{\partial S} + \frac{\partial C}{\partial \sigma}\frac{\partial \sigma}{\partial S} \tag{2}$$

In this case, failing at modeling the right dynamics of the IV σ with respect to the price of the underlying asset would result in wrong hedges (they could be worse performing than the ones given by the BS model). This issue can be extended to other types of risk that involve changes in the volatility with respect to some of the parameters of the model.

Additionally, some operational restrictions should be considered (calibration errors, computing time, stability of the model) when choosing the appropriate model.

2.2 Local vs Stochastic Volatility Models [1]

The local volatility (LV) models developed by Dupire (1994) and later Derman and Kani (1994) represented a breakthrough in volatility modeling mainly because IV could be computed only by using market option prices and inputs like strike prices and maturities. This proved particularly useful in the case of exotic options for which it was difficult to find a closed-form solution without recurring to expensive numerical methods. The main result in Dupire's work is what is now called the Dupire's Formula which gives a tractable expression for the local variance:

$$\sigma^2_{LV}(t,S) = 2\frac{\frac{dC}{dT} + rK\frac{dC}{dK}}{K^2\frac{d^2C}{dK^2}}\bigg|_{K=S,\ T=t} \tag{3}$$

For forward prices, we can use the same expression minus the drift

$$\sigma^2_{LV}(t,F) = 2\frac{\frac{dC}{dT}}{K^2\frac{d^2C}{dK^2}}\bigg|_{K=F,\ T=t}$$

The BS formula can then be used to relate the local variance to the implied variance. As expected, the relationship between the local variance and the implied variance is positive.

By discretizing (3), sets of strike prices, option prices, maturities and asset prices can be used as inputs to compute the local variance. Since none of these steps requires the assumption on a parameter and it only uses market inputs, the LV is a market model.

The main problem with LV models is that they are meant to be static models. They fail to capture the price dynamics of the volatility skew and therefore they are not suitable for hedging. Specifically, in a LV environment, when the price of the underlying increases, the IV smile/skew shifts left (i.e. lower volatility at the same strike price)[2].

[Insert Figure 2 here]

An alternative to the LV model, Stochastic volatility (SV) models propose a two-factor framework in which a stochastic process is defined not only for the underlying asset price

[1] This subsection is based on Gatheral (2006), Bergomi (2016), and Majmin (2005)

[2] This is something pointed out by Hagan et al. (2002)

but also for the volatility which also involves the modeling of two different Brownian motions as sources of randomness. Depending on the model, these Brownian motions could be correlated or not.
In the Hull and White model (Hull and White (1987)) the variance follows a lognormal stochastic process. This was the first diffusin SV model that became popular in the financial industry. It can be summarized in the following system of differential equations:

$$\begin{aligned} dS &= \phi S dt + \sigma S dW_1 \\ d\sigma^2 &= \mu\sigma^2 dt + \psi\sigma^2 dW_2 \\ dW_1 dW_2 &= \rho dt \end{aligned} \tag{4}$$

Where ϕ, μ, ψ are deterministic functions of S, σ and t. For the zero-correlation case, this model reduces to a BS model with a time-dependent volatility and possesses a closed-form solution. However, for non-zero correlation, which is more empirically correct, the distributional properties of the solution depend on the path of the implied variance between t and T which prompts the need of numerical methods to find a solution of the partial differential equations. Hull and White (1987) proposes the Antithetic Variable Technique to implement in the Monte Carlo simulation.
The Heston model (Heston (1993)) proposes a mean-reverting model in which the two Brownian motions are allowed to be correlated. It became popular due to its well-known analytical solution to retrieve option prices

$$\begin{aligned} dS &= \phi S dt + \sqrt{\sigma^2} S dW_1 \\ d\sqrt{\sigma^2} &= -\mu\sqrt{\sigma^2} dt + \psi dW_2 \\ dW_1 dW_2 &= \rho dt \end{aligned} \tag{5}$$

In order to solve for the volatility and the option price of this model, Heston uses Fourier transform techniques to arrive to an expression that then can be simulated using Monte Carlo numerical methods. Bergomi (2016) discusses that the Heston model lacks flexibility when handling with exotic options (see Chapter 6).
In sum, some of the most popular models that came before the SABR model by Hagan et al. (2002) were either bad at predicting the dynamics of the volatility or too demanding in terms of the computational methods they require. The SABR model is going to provide a closed-form formula that not only is able to reproduce the dynamics of the volatility smile/skew but also offers a relatively easy analytical expression to compute the implied volatility almost instantaneously. We shall now describe the main characteristics of this model.

3 The SABR model

The Stochastic alpha, beta, rho (from now on, SABR) model, introduced by Hagan et al. (2002), is a stochastic volatility model whose dynamics is given by the following system of stochastic differential equations:

$$\begin{aligned} dF_t &= \sigma_t F_t^{\beta} dW_{1,t} \\ d\sigma_t &= \alpha \sigma_t dW_{2,t} \end{aligned} \tag{6}$$

Where, F_t is a forward price, σ_t is the volatility of the process and $W_{1,t}$ and $W_{2,t}$ are two correlated standard brownian motions, such that $Corr\big(W_{1,t}, W_{2,t}\big) = \rho$. The initial values of these processes are denoted as f and σ_0. The parameter α is the volatility of the volatility process (known as vol-of-vol), and $\beta \in [0, 1]$ determines how the at-the-money volatility changes when the forward price changes. For the log-normal version of the SABR model ($\beta = 1$), it is expected that the ATM volatility remains constant with movements in forward prices.

The main advantage of the SABR model is the closed form solution (Hagan's formula from now on) that allows to get an analytical expression from which we can obtain the implied volatility and then plug it into the Black and Scholes's (BS) equation for pricing options. In contrast to the BS model, the implied volatility from the SABR model is going to be a function of both the strike price and the time to maturity which is expected to reproduce better the behavior of market smiles or skews.

3.1 Hagan's Formula

The formula for implied volatility $\sigma(K, f)$ has been derived in Hagan et al. (2002) by solving a differential equation for the joint probability on the values of the underlying asset and its volatility, computing the price of a european option integrating the payoff times the probability density and then equating the price found in the previous step with the price of the same option under the BS model and solving for the constant volatility obtained from BS model.

Let σ be implied volatility of an option under the SABR model defined as a function of today's price of a forward f maturing at T and the strike K. Then:

$$\sigma(K, f) = \sigma_0 \frac{z}{\chi(z)} \frac{1 + \tau \left(\frac{(1-\beta)^2}{24} \frac{\sigma_0^2}{(fK)^{1-\beta}} + \frac{1}{4} \frac{\rho \beta \alpha \sigma_0}{(fK)^{\frac{1-\beta}{2}}} + \frac{2-3\rho^2}{24} \alpha^2 \right)}{(fK)^{\frac{1-\beta}{2}} \left(1 + \frac{(1-\beta)^2}{24} \log^2 \frac{f}{K} + \frac{(1-\beta)^4}{1920} \log^4 \frac{f}{K} \right)} \tag{7}$$

Where:

$$\begin{aligned}\tau &= T - t \\ z &= \frac{\alpha}{\sigma_0}(fK)^{(1-\beta)/2}\log\frac{f}{K} \\ \chi(z) &= \log\left(\frac{\sqrt{1-2z\rho+z^2}+z-\rho}{1-\rho}\right)\end{aligned}$$

For the log-normal case, $\beta = 1$, which is our chosen specification[3], the equation (7) reduces to:

$$\sigma(K,f) = \sigma_0\frac{z}{\chi(z)}\left[1+\frac{1}{8}\tau\left(2\alpha\rho\sigma_0+\alpha^2\left(\frac{2}{3}-\rho^2\right)\right)\right] \tag{8}$$

Where:

$$\begin{aligned}z &= \frac{\alpha}{\sigma_0}\log\left(\frac{f}{K}\right) \\ \chi(z) &= \log\left(\frac{\sqrt{1-2\rho z+z^2}+z-\rho}{1-\rho}\right)\end{aligned}$$

It is then easy to plug the implied volatility into a Black pricing formula to obtain option prices which is the first desirable property of a volatility pricing model.
Under the SABR model, the value of a call is

$$C_{call} = C_{BS}(K,T,f,\sigma_B(K,f)) \tag{9}$$

for our case of analysis, $\beta = 1$, we have $\sigma_B(K,f) = \sigma_B(K,f,\sigma_0,\alpha,\rho)$ is given by equation (8). If we differentiate the value of the call in (9) with respect to σ_0 we get the vega risk, the sensibility of option value to changes in volatility. We note that, when $K = f$, $z/\chi(z) \to 1$, this source of risk will remain constant as described in Hagan et al. (2002). Furthermore, the SABR model has two additional sources of risk ,since it has two parameters ρ and α that are stochastic. These has been named as vanna that is the risk to ρ and as volga that is the risk to α. Vanna represents the risk to the skew increasing and volga represents the risk to the smile becoming more pronounced.
One of the things to remark at this point is that Hagan's formula might not be able to reproduce the entire volatility surface because the short-dated slope of the skew produced by 8 is smaller than the one observed from market prices. According to Lee (2005) the slope of observed skews has the following characteristics:

$$\begin{aligned}\frac{\partial\sigma}{\partial x} &\propto \left(\frac{1}{\sqrt{\tau}}\right), \text{ if } \tau \to 0 \\ \frac{\partial\sigma}{\partial x} &\propto \left(\frac{1}{\tau}\right), \text{ if } \tau \to \infty\end{aligned} \tag{10}$$

[3]During a presentation at MIT, Andrew Lesniewski (Lesniewski (2014)) pointed out that the usual assumptions of $\beta \to 0$ had led to extreme calibration results during the peak of the 2008-2009 crisis which is why most practitioners have turned to $\beta \approx 1$ in the post-crisis period.

Where x represents the log-moneyness, $\log(K/f)$. However, following Alós et al. (2007), the short skew slope of the SABR is given by:

$$\frac{\partial \sigma}{\partial x} \to -\frac{1}{2}\alpha\rho, \text{ if } \tau \to 0 \tag{11}$$

which is $O(1)$, strictly smaller than $O(\tau^{-1/2})$ which means that the SABR model is not able to capture the behavior of empirical surfaces at short-maturities[4] if calibrated for a whole (K,T) surface.
Moreover, Hagan's formula fails when $K \to 0$ if this makes the probability density function implied by the option prices become negative, so it is advisable to calibrate the surface in reasonable ranges for K.
Figure 3 depicts our observed volatility surface which is going to be our starting and reference point for the rest of the document.

[Insert Figure 3 here]

Our observed data represents the typical example of a volatility skew in which, at short-time maturities, the slope of the skew tends to infinity when the strike prices are low, while flattens for higher strikes. The (K,T) surface for 51 strike prices and 18 different maturities offers a view of the long-dated behavior of this implied volatility: for large maturities the volatility tends to resemble more the constant volatility result of the Black-Scholes model.

3.2 Calibration

While the SABR model offers an easy way to compute the implied volatility from market prices, it has the shortcoming, like other diffusion SV models, that it fails to reproduce the short-time behavior of the volatility surface. For this reason, fitting a single set of parameters to the whole strike-maturity surface might prove ineffectual. In fact, what practitioners recommend in order to almost perfectly reproduce the market volatility surface is to calibrate a set of parameters for each maturity in the book which is our preferred calibration method.
Given a set of strike prices and a forward price, we can estimate parameters α, ρ and σ_0 for each maturity by minimizing the following sum of squared errors:

$$(\hat{\alpha}, \hat{\rho}, \hat{\sigma_0})_\tau = \underset{\alpha,\rho,\sigma_0}{argmin} \sum_{i}^{N} \left(\sigma_{i,\tau}^{m} - \sigma_i^{SABR}(f, \tau, K_i, \alpha, \rho, \sigma_0)\right)^2 \tag{12}$$

Where σ^{SABR} is given by (8) and σ^m is the observed implied volatility from market prices. We estimate (12) by running a Nonlinear Least Squares (NLS) regression applying the "curvefit" function in the Scipy Python program on the implied market volatility generated by (1). The results from this calibration and the behavior of the short-time skew are shown in Figure 4. Afterwards, we will check if the calibrated parameters and

[4] Lee (2005) also states that for long-dated maturities, the theoretical slope of the skew also behaves as to $O(\tau^{-1})$.

the implied volatility obtained with them are coherent with both market data as well as theoretical results.

[Insert Figure 4 here]

For our set of maturities, we can see that the behavior of the calibrated parameters (Panel 4a) is consistent with market behavior: (i) the vol-of-vol α diminishes almost exponentially with respect to the time to maturity (in Figure 3, the IV surface does indeed flatten as T increases); (ii) while increasing, the initial volatility parameter σ_0 seems fairly stable with respect to τ as does (iii) the correlation parameter ρ, which gets closer to zero with the largest maturities.[5] It is also interesting to see that the short skew (Panel 4b) is proportional to T^{-1}. This is then consistent with the fact that Hagan's implied volatility is better suited for long maturities.
Having checked that our calibration gives coherent values for the SABR parameters, we will now use them to estimate the implied volatility.

3.3 Implied volatility: Hagan's formula and market data

We want now to evaluate the performance of our calibration of Hagan's formula in terms of its ability to reproduce the observed data.
We compute the implied volatility surface by applying the set of calibrated parameters for each maturity. Alternatively, we have estimated the implied volatility with Hagan's formula applying a single set of parameters[6] to the whole (K,T) surface to illustrate the difference in results with respect to the observed volatility skews and surface.
Our results, are shown in Figures 5 and 6. Applying the calibration for each maturity generates a volatility skew and surface that almost perfectly overlaps the observed ones. In contrast, when we use a single set of parameters, Hagan's surface undershoots the observed surface for low maturities.

[Insert Figure 5 here]

[Insert Figure 6 here]

With this in mind, we have computed the mean squared errors for our estimations by comparing them with the volatility surface implied by the market data (Figure 7). We thought it was interesting to see how this error measure behaved for different expiration times and strike prices. It is evident from both panels in Figure 7 that by using a different set of calibrated parameters for each maturity ("HaganT") our deviations are virtually zero for the entire surface (across K and T). In contrast, when using a single set of estimated parameters for the entire surface, Hagan's formula performs relatively

[5] This behavior of the SABR parameters is similar to those found in Skantzos (2016) which performed the same type of calibration exercise for 6-month EURIBOR forward caps on different valuations dates. They found evidence of the stability of the parameters over a significant period of time.

[6] We took the calibrated parameters for a medium maturity ($\tau \approx 1.7$). Specifically, $(\alpha, \rho, \sigma_0) \approx (0.598, -0.589, 0.152)$.

worse (with MSE up to 14 and 18 percent) for short maturities and low strikes (as predicted above).
The fact that the fit of Hagan's implied volatility is almost perfect when we use different parameters for each maturity date may explain why this practice is so popular among practitioners even if it is not mathematically ideal.[7]

[Insert Figure 7 here]

Additionally, we checked that our resulting calibration was able to reproduce market price dynamics. As explained in Section 2, the main reason why SV models topped over LV models was the fact that the latter failed to reproduce the dynamics of the volatility smile with changes of the underlying forward price. The SABR model we calibrated is able to reproduce the right price dynamics (Figure 8): as the price of the forward increases (decreases) the volatility smile shifts to higher (lower) values for each strike price. An additional thing to highlight here is that the ATM volatility behavior is consistent with our choice for the parameter β: as predicted, the ATM volatility remains virtually unchanged with changes in the forward price.

[Insert Figure 8 here]

[7] Were it the case that Hagan's formula could be calibrated for the entire volatility surface.

4 Conditional Monte Carlo

In this section, we are going to explain another popular method to compute the Implied Volatility Surface, the Conditional Monte Carlo approach. The main advantage of this method is its generality, that is, we can apply it in every stochastic volatility model. However, Conditional Monte Carlo (CMC, from now) is not to popular among practitioners, because of the high computational time it demands.

The idea behind CMC is to use the Law of Iterated Expectation to re-write our problem. Given an appropriate choice of conditioning paths, we can reduce the dimension of uncertainty in our problem and come up with a tractable expression for simulations.

4.1 Willard's transformation

In general, if we compute a call option price (V_0), the following is true:

$$\begin{aligned} V_0 &= E\Big[e^{-rT}(F_T - K)_+ \Big] \\ &= E\Big[E\Big[e^{-rT}(F_T - K)_+ \mid \{W_{2,t_i}\}_0^T, \{\sigma_{t_i}\}_0^T\Big] \Big] \end{aligned} \tag{13}$$

The main issue we face when we want to use CMC to compute an option price from a SV model is the correlation between $W_{1,t}$ and $W_{2,t}$. Normally, we need uncorrelated uncertainty process to apply this method. However, we can solve this issue doing a simple transformation (Willard, 1997).

First, note that, for any pair of correlated brownian motions $(W_{1,t}, W_{2,t})$ with ρ as correlation, the following decomposition is true:

$$W_{1,t} = \sqrt{1-\rho^2} Z_t + \rho W_{2,t}$$

with $W_{2,t}$ and Z_t independent brownian motion. Then the Forward price dynamics in the Black Model can be re-write as:

$$dF_t = F_t \sigma_t (\sqrt{1-\rho^2} dZ_t + \rho dW_{2,t}) \tag{14}$$

After usual algebra, problem (13) can be re-write as the Black formula's expectation:

$$V_0 = E\Big[e^{-rT}\big(F_0\xi_0 Prob(x < d_1) - K Prob(x < d_2)\big) \Big] = E\Big[Black(F_0\xi_0, K, \bar{\sigma_\rho}, T, r)) \Big]$$

Where:

$$\xi_0 = \exp\Big\{\rho \int_0^T \sigma_s dW_s^{(2)} - \frac{1}{2}\rho^2 \int_0^T \sigma_s^2 ds\Big\}$$

$$\bar{\sigma_\rho} = \sqrt{\frac{1}{T}\Big(\int_0^T \sigma_s^2 ds - \rho^2 \int_0^T \sigma_s^2 ds\Big)}$$

$$d_2 = \frac{ln\big(\frac{F_0\xi_0}{K}\big)}{\bar{\sigma_\rho}\sqrt{T}} - \frac{1}{2}\bar{\sigma_\rho}\sqrt{T}; \quad d_1 = d_2 + \bar{\sigma_\rho}\sqrt{T}$$

Is important to mention here, that this result is independent to any volatility model choice. In consequence, if we model the Forward price as a Geometric Brownian Motion, we always can re-define a vanilla call option price as:

$$V_0 = E\Big[Black(F_0\xi_0, K, \bar{\sigma_\rho}, T, r)\ \Big] \tag{15}$$

This is quite important, because we do not need to be concerned about $W_{1,t}$ (this is why we say at the beginning that we can reduce the dimension of uncertainty in our problem). As a result, the expectation in (15) can be computed easily, thanks to the Law of Large Numbers, as the average of several realizations of the Black Formula.

4.2 From CMC prices to Implied Volatility

We start this chapter describing the CMC approach to compute a derivatives price. In this section we are going to explain how to compute the implied volatility (I) given this price. Following the definition of implied volatility by (Fouque et al., 2000).

We can compute I solving the following optimization problem:

$$I = \arg\min_{I>0} \Big(e^{-rT}\big(F_0 Prob(x < d2 + I\sqrt{T}) - K Prob(x < d2)\big) - V_0\Big)^2 \tag{16}$$

Where:

$$d2 = \frac{ln\big(\frac{F_0}{K}\big)}{I\sqrt{T}} - \frac{1}{2}I\sqrt{T}$$
$$x \sim \mathcal{N}(0,1)$$

Note that this problem is highly non-linear, this is why it is usually solved using numerical methods. In the following section, we are going to explain some of the most popular numerical methods to solve optimization problems, their advantages and their disadvantages.

4.3 Numerical methods to solve optimization problems

The dependence of option price on the implied volatility is given by the Black-Scholes formula, see the section 2, to get the values of implied volatility using known values of option price, we have to use inverse function, see equation (1). However, as already mentioned in the section 2, the analytically expression is not possible, and, therefore, numerical methods have to be used for the problem at hand.

The required values of the implied volatility are roots of the function representing difference between the right part of the Black-Scholes formula and the option price. Thus, we can apply all the power of root-finding numerical methods developed for continuous functions. All the methods are iterative. Among most popular, simple and effective methods are bisection method and Newton-Raphson method.

4.3.1 Bisection method

Every iteration of the bisection method starts from the segment $x \in [x_{left}, x_{right}]$ of domain of the investigated function f, where function values $f(x_{left})$ and $f(x_{right})$ have different signs at the ends of the segment:

$$\text{sign}(f(x_{left})) \neq \text{sign}(f(x_{right}))$$

Then we check the sign of the function at

$$x = (x_{left} + x_{right})/2$$

and for the next iteration set $x_{left} = x$ if $\text{sign}(f(x)) = \text{sign}(f(x_{left}))$ and $x_{right} = x$ otherwise.

The advantage of the bisection method is its simplicity and low computational cost of individual iteration. However, it converges slowly compared to some other methods.

4.3.2 Newton-Raphson method

A more complicated method is Newton-Raphson algorithm. The essence of the Newton method is as follows. According to Taylor's expansion, at the vicinity of x_0 we have

$$f(x) = f(x_0) + f^{'}(x_0)(x - x_0) + h(x)(x - x_0)$$

where function $h(x)$ is infinitely small at $x \to x_0$. The equation

$$f(x) = f(x_0) + f^{'}(x_0)(x - x_0)$$

defines the linear tangent function at point x_0. Note that we want to find values of x where $f(x) = 0$. Therefore, if to set $f(x) = 0$ in the equation above, we have

$$f(x_0) + f^{'}(x_0)(x - x_0) \to 0$$

from where we have

$$x = x_0 - \frac{f(x_0)}{f'(x_0)}$$

which sets the general iterative procedure lying in the heart of the Newton method:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

It could turn out that steps are not optimal. They can be tuned by introducing scalar factor λ. This generalization is called Newton-Raphson method:

$$x_{n+1} = x_n - \lambda\frac{f(x_n)}{f'(x_n)}$$

The advantage of Newton-Raphson method is its faster convergence. But the individual iterations are more computationally expensive. Another disadvantage of this method is that if the first derivative is close to zero ($f'(x_n) \to 0$) then this procedure behaves unstably.

4.3.3 Brent–Dekker method

The Brent-Dekker method is a numerical iteration method of finding zero of a given function [Dekker (1969), Brent (1971)]. It utilizes combination of bisection and secant methods. At each iteration, it computes two candidate values for the next value of the argument, one according to secant method and the other according to bisection method and chooses the best value as the next iteration.

More detailed, the Brent-Dekker method at each iteration considers three values of the argument, b_{k-1}, b_k, and a_k. Values b_{k-1} and b_k are two last iterates and are used for the application of the secant method. Value a_k are chosen from the list of previous iterates (or arbitrarily at the first iteration) so that $f(a_k)$ and $f(b_k)$ were of different signs: this guarantees the existence of the root between a_k and b_k. Values a_k and b_k are used for the application of the bisection method. Normally, b_k should be better guess than a_k, that is, $f(b_k) <= f(a_k)$. At the first iteration b_{k-1} is set equal to a_k. If the secant method gives value of iterate between current iterate b_k and value given by bisection method, then the method takes the iterate from the secant method. Otherwise, the iterate from bisection method is used.

The method demonstrates convergence superior to that for the secant and bisection methods. If the function f is sufficiently smooth around the root, the convergence of the method is more than linear. In this work we used *Python* function *optimize.fminbound()* from *scipy* library, which implements Brent-Dekker method.

[Insert Table 1 here]

In the table we compare three different numerical methods. It is presented with the number of steps and time of convergence in seconds in the corresponding cells. Easy to see that the Newton-Raphson test converges always faster than others and just for a few steps. One step of the bisection method is faster than others (as you can see from line 3) but problems of this method are: (1) needs to use a lot of iteration until getting results with reasonable accuracy; (2) needs to specify borders very precisely to optimize count time; (3) needs a stop criterion for the algorithm. Brent-Dekker method is easy to use but relatively slow.

4.4 Implied volatility applying a CMC on the SABR Model

In this section we explain how we compute CMC call option prices in the SABR given a set of the SABR model parameters $(\alpha, \beta, \rho, \sigma_0)$[8] and a set of Forward market prices given a range of maturities ($F_0 = F_0(T)$). Then, we compute the implied volatility, as

[8]This set of parameters were estimated in the previous chapter from the Hagan Formula. We use the same set of parameters for the sake of comparison.

the solution of the problem describe in (16), using a numerical optimization method for each of the option prices previously computed. After these two steps we obtain the CMC implied volatility surface.

4.4.1 Computing CMC call option prices

In this section, we explain how to compute CMC call option prices for the SABR model. From now on, lets assume $\beta = 1$ as in the previous chapter. Following equation (14), we can re-write the SABR model describe in (3) as:

$$\begin{aligned} dF_t &= F_t\sigma_t(\sqrt{1-\rho^2}dZ_t + \rho dW_{2,t}) \\ d\sigma_t &= \alpha\sigma_t dW_{2,t} \end{aligned} \tag{17}$$

Note that, in this model, the volatility process $\{\sigma_t\}_0^T$ is the solution of a Geometric Brownian Motion without drift:

$$\sigma_t = \sigma_0 exp\Big\{-0.5\alpha^2 t + \alpha W_{2,t}\Big\},\ \forall\ t \in [0,T] \tag{18}$$

In consequence, to compute $V_0(T,K)$, as in (15), we need to generate several paths of $\{W_{2,t}\}_0^T$ and then, using (18), generate several paths of $\{\sigma_t\}_0^T$. Recall that we need these paths to define a conditional expectation as in (13).

[Insert Figure 9 here]

After doing this, we compute $V_0(T,K)$ as in (15), that means:[9]

1. Compute $\int_0^T \sigma_t^2 dt \approx \frac{T}{n}\sum_{i=0}^{n-1}\sigma_{t_i}^2,\ st: t_0 = 0\ and\ t_n = T$
2. Compute $\int_0^T \sigma_t dW_{2,t} \approx \sum_{i=0}^{n-1}\sigma_{t_i}(W_{2,t_{i+1}} - W_{2,t_i})$
3. Compute ξ_0 and $\bar{\sigma_\rho}$
4. Compute the call option price of this path using the Black Formula and save it
5. Repeat the last three steps for each simulated Brownian Motion path
6. Average all the save option prices, this result is our CMC option price

We use 10000 time steps, which we denote by n, and 10000 simulated paths, which we denote by m. We fix $m = 10000$ because, with a lower number of simulations the average of the following stochastic integral $\int_0^T \sigma_s dW_{2,t}$ would not be close to zero, and with a higher number of simulations the algorithm spends too much time to give us a result. The following table summarizes what we observed about this:

[Insert Table 2 here]

[9] It is important to remember that, the initial forward price given for the market is a function of the maturity.

$$F_0(T) = S_0 e^{rT}$$

Then, we need to change it for each maturity when we calculate the option price given (15).

4.4.2 Implied volatility from CMC option prices

To compute the whole implied volatility surface we need to solve the problem described in (16) for each pair (K,T). To do this we use the Brent-Dekker method, for the reasons we discuss in the previous section. Also, we made a constrain, $I \in [0.005, 5]$, to guaranty stability in the implied volatility surface.[10] In Figure 10 we compare our CMC implied volatility surface with the implied volatility surface calibrated from market data.

[Inser Figure 10 here]

We can observe a high difference between both of them in the short size (low K and low T) as expected.

[Insert Figure 11 here]

[10] If we do not put a lower bound, given that (16) could have several different roots, $I \to 0$ could be a solution. However, this solution does not make sense in an uncertain world.

The most clear example is observed when F_0 is significantly higher than K and T is close to zero, in our case $max\left(\frac{F_0}{K}\right) \approx 30$ and $min(T) \approx 0.1$, lets see what happen in this case:

$d1(I) = \frac{ln(30)}{I\sqrt{0.1}} + 0.5I\sqrt{0.1}$

$d2(I) = \frac{ln(30)}{I\sqrt{0.1}} - 0.5I\sqrt{0.1}$

If we compare $I = 3$ and $I = 0$ there is not a significant difference in the resulting option prices (2898.686 and 2898.666, respectively) because the probability density function has almost the same values.

5 Comparison between Hagan's Formula and CMC Implied Volatility

Figure 12 plots the four surfaces discussed so far. While the Hagan estimation calibrated for each maturity completely overlaps the observed volatility surface (in blue), both the Hagan surface computed with a single parameter set (red) and the MC-simulated surface fail to capture the behavior of the observed volatility for low strike prices and short times to maturity. As mentioned before, the CMC simulation overstimates the low-strike explosion of the volatility while the alternative Hagan computation underestimates it.

[Insert Figure 12 here]

In fact, Figure 13 shows that most of the difference in the Mean Squared Estimation error is produced at maturities lower than 2 years and strikes lower than 40% of the forward price. Once these thresholds are surpassed, there is not too much difference in MSE between the three methods. Regarding the reasons why there are differences between the Hagan and the CMC methods when a single set of parameters is applied one can think of two possibilities: (i) Hagan's formula fails out of the money or (ii) the CMC is not numerically precise at short maturities and very low strikes because there might be different implied volatilities that can yield the same option prices and therefore treated equally by the optimization method (see Footnote 10).

[Insert Figure 13 here]

However, there's another dimension that is very relevant in trading purposes: computing time. Hagan's solution is by far the fastest method taking much less than a second to compute. In contrast, it takes approximately 2.5 hours to compute the CMC-simulated volatility surface (Table 3).

[Insert Table 3 here]

6 Extension: The Fractional SABR

In previous chapters, we present the SABR model and discuss about their strengths and weakness. In this regard, the biggest weakness of the SABR model is that it can not reproduce the short time ATM volatility skew observed in market data. This weakness is share with all the diffusion stochastic volatility models.
As a result, a new kind stochastic volatility models were propose to reproduce this market characteristic [see Alós et al. (2007), Fukasawa (2011)]. This new kind of models were call rough volatility models [Gatheral et al. (2018)], because the resulting volatility paths (σ_t, $\forall\ t > t_0$) are rougher than volatility paths obtained from any diffusion stochastic volatility model. The reason of this is that rough volatility models uncertainty is model using a fractional brownian motion (W_t^H) while the others a Standard Brownian motion (W_t).

[Insert Figure 14 here]

As an example, we are going to present the log-normal Fractional SABR model (FSABR model) following Akahori et al. (2017) and Alós et al. (2018):

$$\begin{aligned} dF_t &= F_t\sigma_t(\sqrt{1-\rho^2}dZ_t + \rho dW_{2,t}) \\ \sigma_t &= \sigma_0 e^{\alpha W_{2,t}^H} \end{aligned} \tag{19}$$

Where, F_t is a forward price, σ_t is the volatility process, Z_t and $W_{2,t}$ are two independent standard brownian motions, and $W_{2,t}^H$ is a fractional brownian motion with Hurst exponent $H \in [0,1]$ generated by $W_{2,t}$, which is define as:

$$W_{2,t}^H = \int_0^t K_H(t,s)dW_{2,s}$$

where K_H is the Molchan-Golosov kernel. Also, the auto-covariance function of a fractional Brownian motion is denoted by $\Sigma(t,s)$ and defined as:

$$\Sigma(t,s) = E(W_{2,t}^H W_{2,s}^H) = \frac{1}{2}\big(t^{2H} + s^{2H} - |t-s|^{2H}\big)$$

We have the same parameters as in (6) plus the Hurst exponent $H \in [0,1]$. As proof in Alós et al. (2007), using $H < \frac{1}{2}$ reproduce the market skew. In this regard, Gatheral et al. (2018) show that we can inferred $H \approx 0.1$ from High Frequency time series data. This findings indicate that volatility should be model as a non-markovian rough process. The issue with this approach is that computing implied volatilities is still too complicated for practical purposes, because computing a fractional brownian motion is too time consuming. The reason of this is that the fractional brownian motion is obtained applying Cholesky descomposition to the autocovariance matrix, a high dimension matrix.

7 Conclusions

- The Hagan's formula allows us to reproduce the market volatility skew and surface almost perfectly when we calibrate for each maturity. In contrast, when we use a single set of parameters, Hagan's surface undershoots the observed surface for low maturities.

- The calibrated parameters and the implied volatility obtained with them are coherent with both market data as well as theoretical results since the SABR model we calibrated is able to reproduce the right price dynamics (Figure 8): as the price of the forward increases (decreases) the volatility smile shifts to higher (lower) values for each strike price.

- On the other hand, we can use Conditional Monte Carlo (CMC) Method as a general framework to compute option prices. However, this method has the following weakness:

  1) time consuming: compute the whole volatility surface demand several minutes

  2) high mean square error: relative to the one obtained from Hagan's formula

  3) overshoot short time skew: relative to market data, because of optimization instabilities

- The Hagan's implied volatility surface computed with a single parameter set and the CMC implied volatility surface fail to capture the behavior of the observed volatility for low strike prices and short times to maturity.

- Regarding computing time, Hagan's solution is by far the fastest method taking much less than the CMC method. The first method takes less than a second while the second one takes approximately 2.5 hours to compute the implied volatility surface. This is very relevant for trading purposes.

- Recent research works propose to model the volatility process with fractional Brownian motions. As proved in Alós et al. (2007), using $H < \frac{1}{2}$ reproduces the market skew. This is an indicator that volatility should be modeled as a rough process. However, computing implied volatilities using this approach is still too complicated for practical purposes. In consequence, we should continue with this line of research.

References

Akahori, J., Song, X., & Wang, T.-H. (2017, 02). Probability density of lognormal fractional sabr model.

Alós, E., Chatterjee, R., Tudor, S., & Wang, T.-H. (2018, 01). Target volatility option pricing in lognormal fractional sabr model. *Quantitative Finance*. doi: 10.1080/14697688.2019.1574021

Alós, E., Leon, J., & Vives, J. (2007). On the short-time behavior of the implied volatility for jump-diffusion models with stochastic volatility. *Finance and Stochastics*, *11*, 571–589.

Bergomi, L. (2016). *Stochastic Volatility Modeling*. Chapman Hall/ CRC Financial Mathematics Series.

Brent, R. P. (1971). An algorithm with guaranteed convergence for finding a zero of a function. *The Computer Journal*, *14*, 422-425.

Dekker, T. J. (1969). Finding a zero by means of successive linear interpolation. In B. Dejon & P. Henrici (Eds.), *Constructive aspects of the fundamental theorem of algebra* (p. 37-48). Wiley-Interscience, New York.

Derman, E., & Kani, I. (1994). Riding on a smile. *Risk*(7), 32-39.

Dupire, B. (1994). Pricing with a Smile. *Risk*(January), 18-20.

Fouque, J.-P., Papanicolaou, G., & R Sircar, K. (2000). *Derivatives in financial markets with stochastic volatility*.

Fukasawa, M. (2011). Asymptotic analysis for stochastic volatility: martingale expansion. *Finance and Stochastics*, *4*(15), 635–654.

Gatheral, J. (2006). *The Volatility Surface: A Practitioner's Guide*. Wiley.

Gatheral, J., Jaisson, T., & Rosenbaum, M. (2018). Volatility Is Rough. *Quantitative Finance*, *18*(6), 933–949.

Hagan, P., Kumar, D., Lesniewski, A., & E. Woodward, D. (2002). Managing Smile Risk. *Wilmott Magazine*, *1*(1), 84–108.

Heston, S. (1993). A closed-form solution for options with stochastic volatility, with application to bond and currency options. *Review of Financial Studies*(6), 327-343.

Hull, J. (2018). *Options, futures, and other derivatives, 10th ed.* Pearson.

Hull, J., & White, A. (1987, 06). The pricing of options on assets with stochastic volatilities. *The Journal of Finance*, *42*(2), 281-300.

Lee, R. (2005). Implied volatility: Statics, dynamics and probabilistic interpretation. In H. G. J. H. R. Baeza-Yates J. Glaz & J. Palacios (Eds.), *Recent advances in applied probability* (p. 241-268). Springer Link, Boston MA.

Lesniewski, A. (2014). Option Smile and the SABR Model of Stochastic Volatility. MIT. Retrieved from `http://lesniewski.us/papers/presentations/MIT_March2014.pdf`

Majmin, L. (2005). *Local and Stochastic Volatility Models: An Investigation into the Pricing of Exotic Equity Options* (Unpublished doctoral dissertation). University of the Witwatersrand.

Skantzos, N. (2016). Stability of the SABR model. Deloitte, Financial Services.

Willard, G. (1997, 11). Calculating prices and sensitivities for path-independent derivative securities in multifactor models. *Journal of Derivatives*, *5*. doi: 10.3905/jod.1997.407982

A Figures

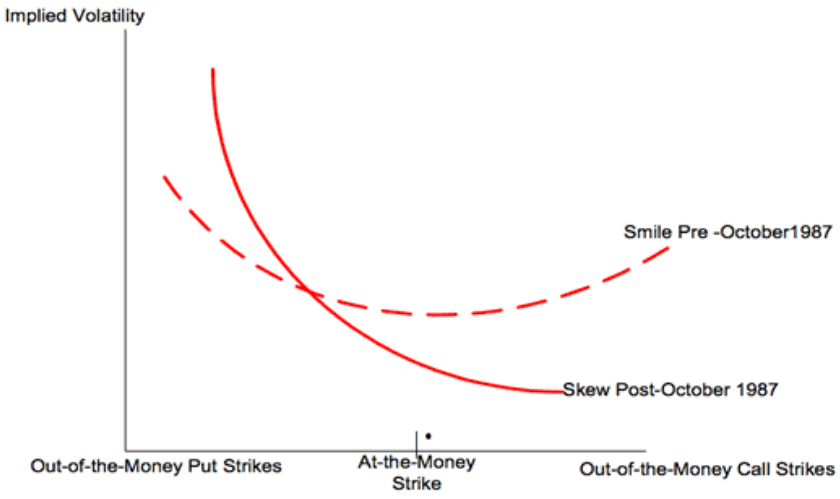


Figure 1: Typical Volatility Smile and Skew

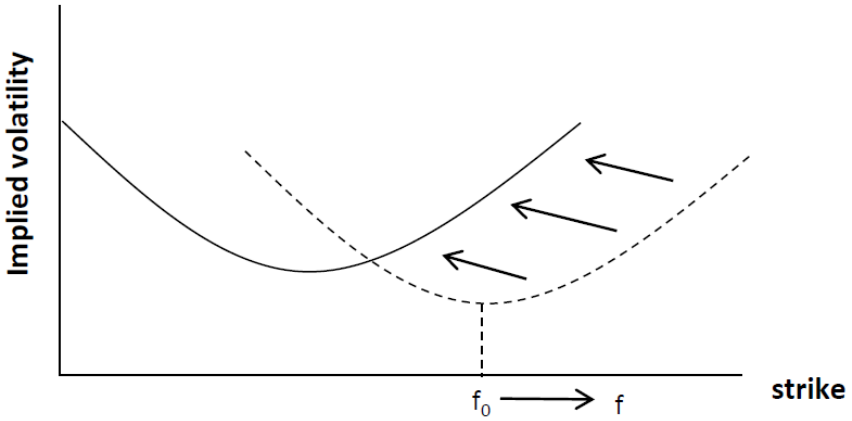


Figure 2: Price Dynamics of the LV model

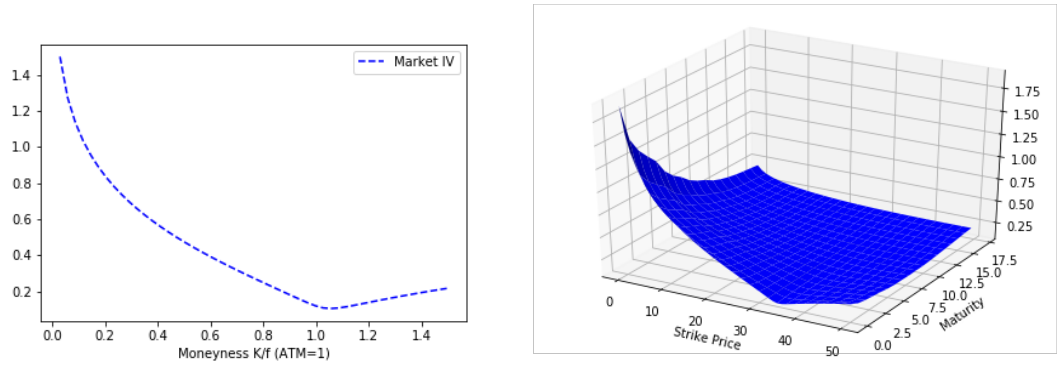
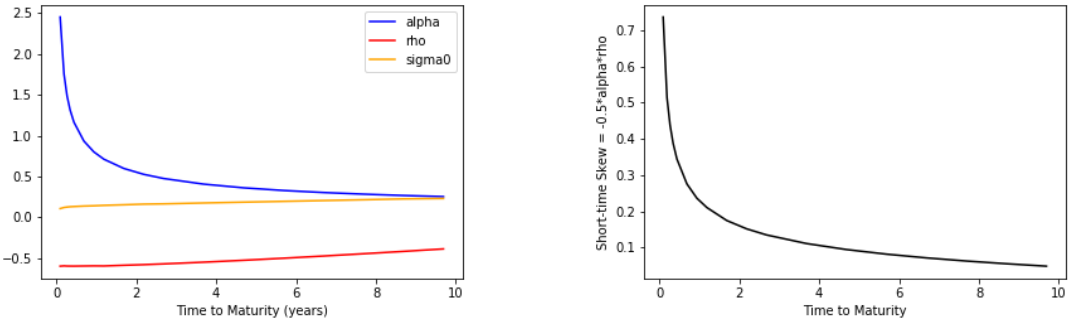


Figure 3: Volatility Skew (T=0.1 years) and Volatility Surface from market prices of European Call Options on the Euro Overnight Index Average - EONIA (April 11, 2019). Courtesy of David Garcia-Lorite



(a) Calibrated α, ρ, σ_0

(b) Short-time Skew and Time to Maturity

Figure 4: Calibration results

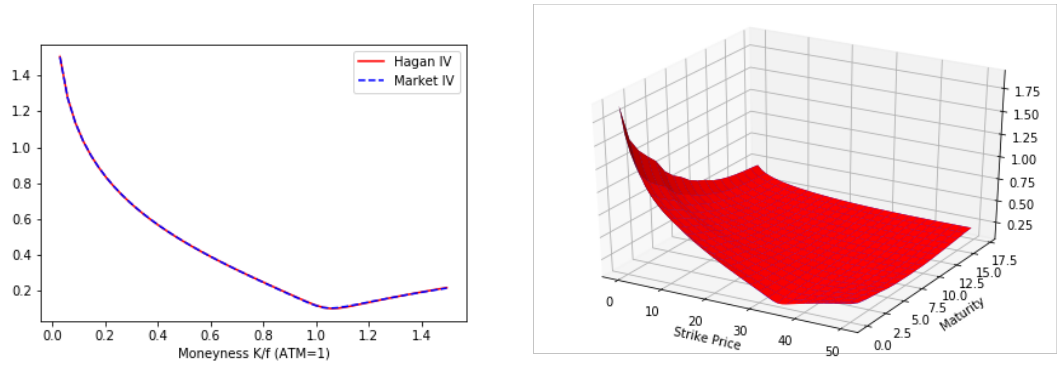


Figure 5: Volatility Skew ($\tau \approx 0.1$ years) and Volatility Surface obtained with Hagan's formula applying a set of parameters (α, ρ, σ_0) for each maturity. Both compared to the observed implied volatility surface.

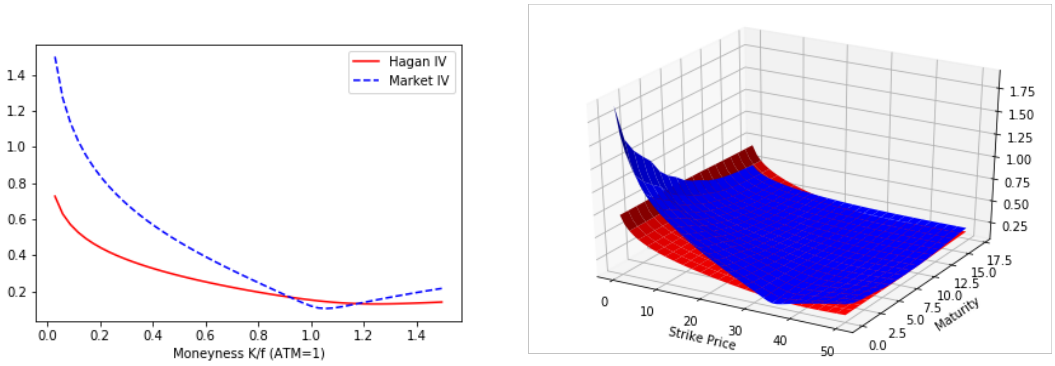


Figure 6: Volatility Skew ($\tau \approx 0.1$ years) and Volatility Surface obtained with Hagan's formula applying a single set of parameters (α, ρ, σ_0). Both compared to the observed implied volatility surface.

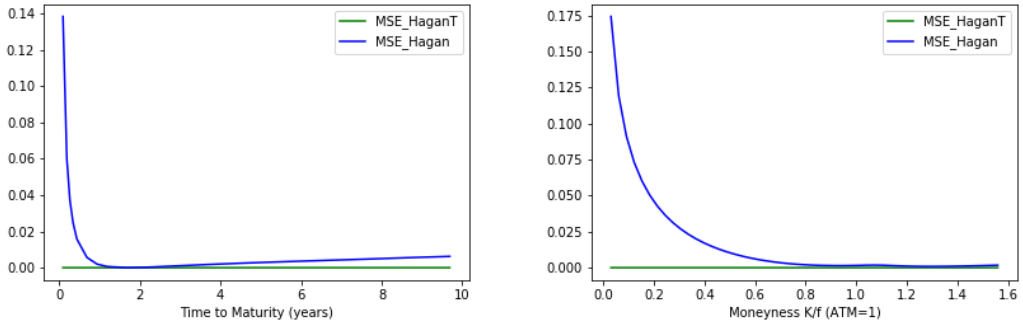


Figure 7: Mean Squared Errors w.r.t Observed data obtained by using Hagan's formula for Implied Volatility

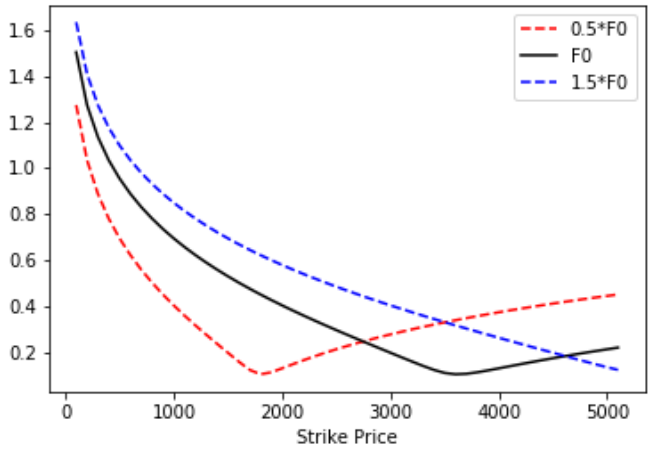


Figure 8: Price Dynamics of the Volatility Smile with the SABR model

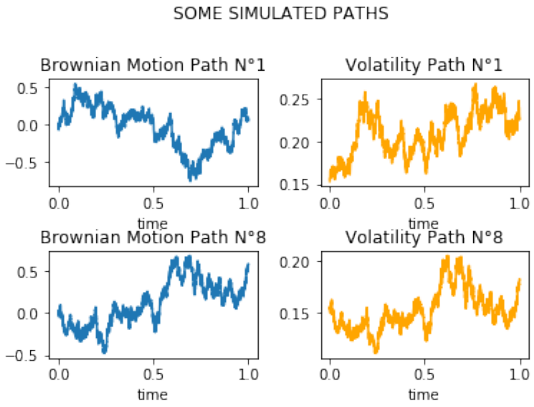


Figure 9: We generate these paths assuming: $\alpha \approx 0.5964$, $\sigma_0 \approx 0.1534$, $T = 1$ and $n = 10000$

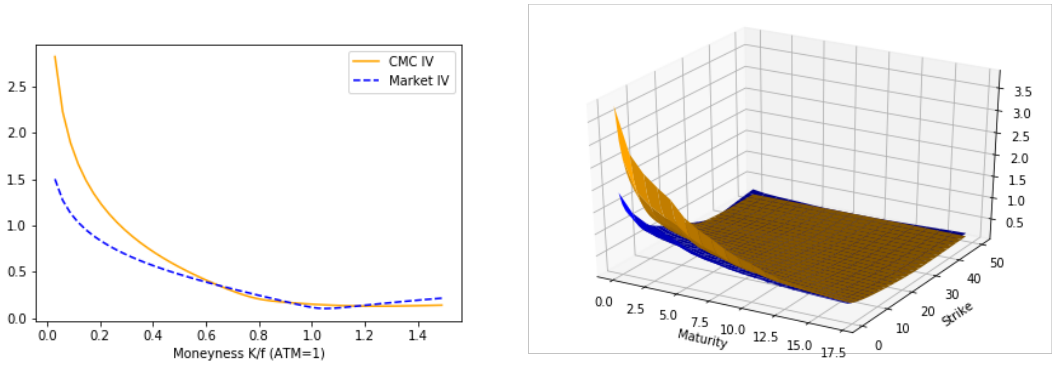


Figure 10: Volatility Skew ($\tau \approx 0.1$ years) and Volatility Surface obtained with CMC simulation applying a single set of parameters (α, ρ, σ_0). Both compared to the observed implied volatility surface.

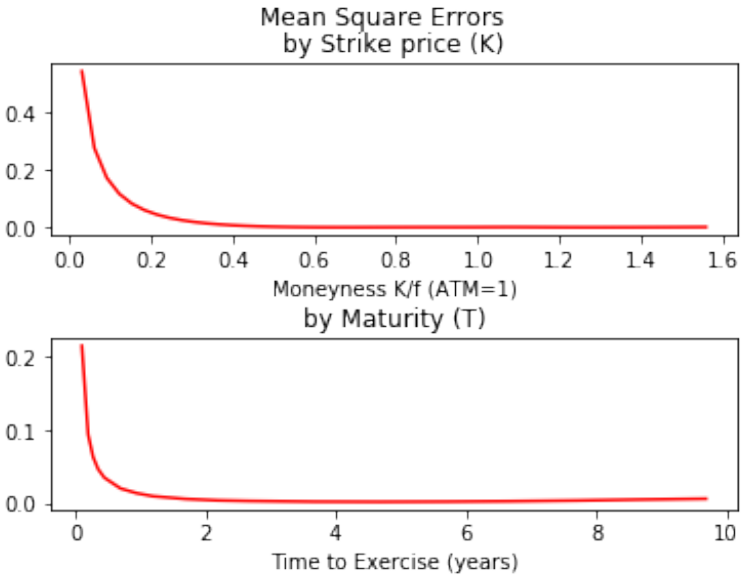


Figure 11: Mean Square Errors across strike prices and maturities

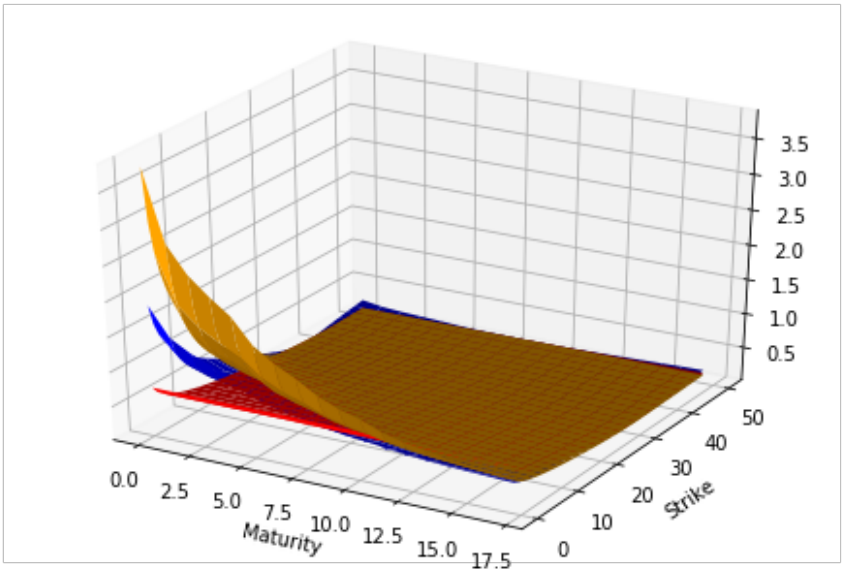


Figure 12: Implied Volatility Surface: Observed data (blue), Hagan's Solution (red) and CMC simulation (orange)

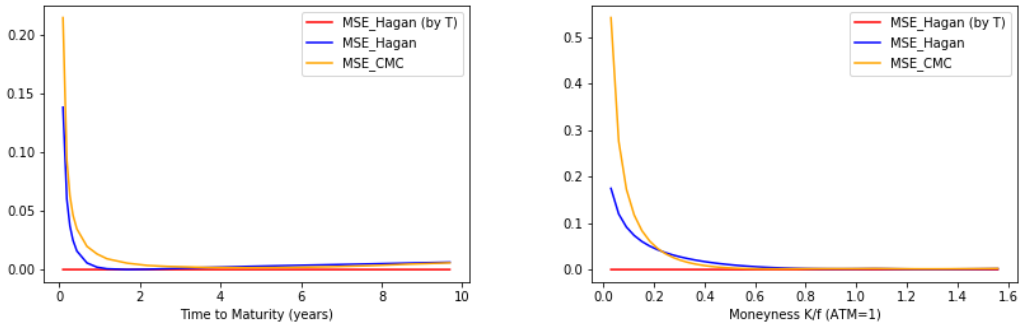


Figure 13: Mean Squared Errors w.r.t Observed Data: Hagan's formula vs CMC Simulation

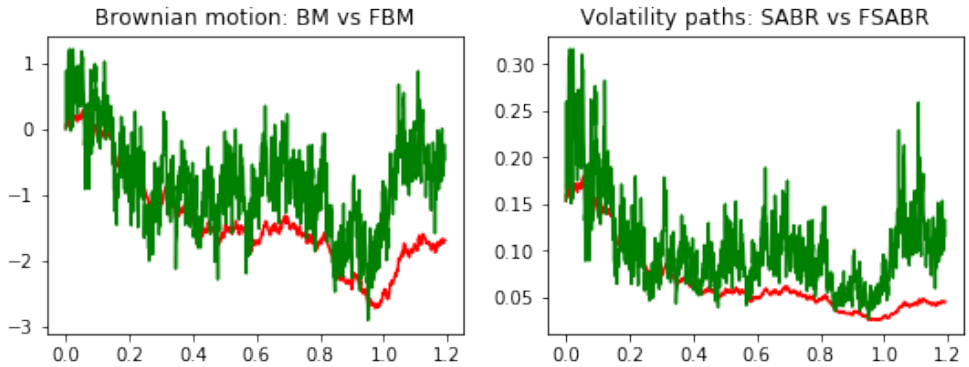


Figure 14: Computation were made with $n = 1000$, $T = 1.14$, $H = 0.13$, $\alpha \approx 0.5964$ and $\sigma_0 = 0.1534$. Standard Brownian motion depicted in red and rough Brownian motion in green.

B Tables

N	Function	boundaries	Bisection	Newton-Raphson	Brent-Dekker
1	$x^2 - 4$	0..3	15 : 0.003	2 : 0.0003	0.0003
2	$x^2 - 4$	0..100	15 : 0.003	2 : 0.0003	0.006
3	$x^2 - 4$	1..3	1 : 0.00009	3 : 0.0003	0.0003
4	3x-cos(x)-1	0..100	19 : 0.005	1 : 0.0005	0.003
5	1/x-10	0.00001..40	19 : 0.003	2 : 0.0003	0.002

Table 1: Numerical methods test.

Simulations	Average	Time in seconds
100	-0.0314649	1.27
1000	-0.0105249	2.21
10000	-0.0008755	19.38
100000	0.0001295	159.42

Table 2: Average of the Stochastic Integral for different simulations

Method	MSE (percent)	Computing Time (seconds)
Hagan (by T)	4.34E-05	0.07812
Hagan	1.743	0.07811
CMC	2.926	8908.33

Table 3: Summary of Performance Indicators

C Python Code

Coding used for our results in this work is in the following github repository:
https://github.com/Rodgrandez/Master-Thesis-code.-