



# **UNIVERSITA' DEGLI STUDI DI PADOVA**

**DIPARTIMENTO DI SCIENZE ECONOMICHE ED AZIENDALI "M.FANNO"**

**CORSO DI LAUREA MAGISTRALE / SPECIALISTICA IN ECONOMIA  
E FINANZA**

**TESI DI LAUREA**

**INTEREST RATE MODELS:**

**Simulation, Estimation and Comparison of some affine term structure models and  
the SABR model**

**RELATORE:**

**CH.MO PROF. Nunzio Cappuccio**

**LAUREANDO/A: Daniele Sambataro**

**MATRICOLA N. 1034641**

**ANNO ACCADEMICO 2014 – 2015**



Il candidato dichiara che il presente lavoro è originale e non è già stato sottoposto, in tutto o in parte, per il conseguimento di un titolo accademico in altre Università italiane o straniere.

Il candidato dichiara altresì che tutti i materiali utilizzati durante la preparazione dell'elaborato sono stati indicati nel testo e nella sezione "Riferimenti bibliografici" e che le eventuali citazioni testuali sono individuabili attraverso l'esplicito richiamo alla pubblicazione originale.

Firma dello studente

---





# *Contents*

## *First Chapter: Affine Term Structure Models*

Introduction	9
1.1 The Term structure equation	12
1.2 Short Interest Rate Model	16
1.2.1 The Vasicek Model	16
1.2.2 The Cox, Ingersoll and Ross model (CIR model)	20
1.3 Forward rate	22
1.3.1 Vasicek forward rate simulation	24
1.3.2 CIR forward rate simulation	25
Matlab Code	28

## *Second Chapter: The SABR model*

Introduction	33
2.1 SABR simulation introduction	35
2.1.1 The constant elasticity of variance process (CEV process)	36
2.2 Absorption probability	38
2.3 SABR conditional distribution	39
2.4 The integrated variance approximation	42
2.4.1 Chen et. al approximation	42
2.4.2 Kennedy's approximation	43
2.4.3 Andersen's approximation	44
2.5 Simulation of the forward path	44
2.5.1 Chen simulation scheme	45
2.5.2 Moment matching approximation	45
2.5.3 Direct inversion scheme	47
Matlab Code	48

## ***Third Chapter: Estimation of the ATSM parameters and pricing of Caplet option***

Introduction	53
3.1 Calibration procedure	55
3.2 Parameter Estimation	58
3.2.1 OLS estimation	58
3.2.1.1 Vasicek OLS parameter estimation	60
3.2.1.2 CIR OLS parameter estimation	61
3.2.2 Maximum likelihood estimation	62
3.2.2.1 Vasicek MLE	63
3.2.2.2 CIR MLE	65
3.3 Pricing	66
3.3.1 SABR Caplet price	67
3.3.2 Vasicek and CIR Caplet prices	68
3.3.3 Pricing result and conclusions	70
Matlab code	72
Bibliography	83
Acknowledgements	89

## **Abstract**

In questo lavoro vengono confrontati diversi modelli sui tassi di interesse. Tale confronto ha come obiettivo principale quello di valutare se i tassi di interesse forward generati dal modello SABR (Stochastic alfa, beta, rho) sono compatibili (in termini di valori e prezzi prodotti) con i tassi di interesse a breve generati da Modelli Affini, quali il modello di Vasicek ed il modello di Cox, Ingersoll e Ross. Per far ciò, vengono innanzitutto generate delle simulazioni dei tassi forward utilizzando il modello SABR con dei parametri plausibili e tratteremo tali tassi come dati reali. Poiché i dati generati sono dei tassi forward, ed i modelli affini descrivono la dinamica del tasso a breve, si procede con la fase di calibrazione, la quale ci consente di identificare il modello Vasicek e il modello CIR che generano i tassi forward più “vicini” possibili a quelli generati col SABR model. Utilizzando i modelli calibrati, si procede alle simulazioni del tasso a breve e, sulla base di queste, si stimano i parametri dei modelli affini utilizzando la stima OLS e quella di Massima Verosimiglianza. Dopo la fase di stima, si simulano i tassi di interesse forward utilizzando i modelli affini (stimati) e si procede al confronto di tali tassi con quelli generati inizialmente con il modello SABR. Si procede poi col calcolo dei prezzi delle opzioni caplet e al confronto degli stessi tra i diversi modelli. I risultati ottenuti mostrano che i tassi forward generati con i modelli affini seguono traiettorie che si discostano in maniera rilevante rispetto ai tassi generati con il modello SABR. Come conseguenza, anche i prezzi delle opzioni ottenuti usando tali modelli risultano avere differenze significative.



# First Chapter

## Affine Term Structure Models

### Introduction

In this chapter we expose the most used affine term structure models (ATSM) for the short term interest rate. Before exposing these models, we have to point out the differences between the interest rate market and the stock market. In the latter, in fact, we use to model the behavior of the asset price to estimate its price at a future date or the price of an option which have that asset as underlying. As argued in Bjork, the short rate of interest  $r$  is not the price of a traded asset, i.e. there is no asset on the market whose price process is given by  $r$ . Then, the short interest rate is not a stand-alone object, and in order to give a definition we need to refer to the zero coupon bond (ZCB) price.

A ZCB is an instrument which pays a certain amount, called *face value* ( for simplicity, we assume that the face value is equal to 1 unity of currency) at its maturity date  $T$ . We express the value at time  $t$  of a ZCB with maturity  $T > t$  by the notation  $P(t,T)$ . It should be clear the similarity between the investment in a zero coupon bond and the investment in a banking account  $X(t)$ , where we put money today to achieve an instantaneous return given by  $X(t)r(t)dt$ <sup>1</sup>.

The short term interest rate  $r(t)$  can be described as the rate applied on instantaneous borrowing (or lending). If we assume the absence of any kind of risk, lending money ( as investing in a bank account ) during the time interval  $[0,T]$ , is equivalent to a strategy in which at each time  $t$ , with  $0 \leq t \leq T$ , we buy a “just maturing” bonds, i.e bonds which will mature at “ $t + \Delta t$ ” with  $\Delta t \rightarrow 0$ . Let  $X(t)$  the value of a bank account at time  $t > 0$ , then at time  $t + \Delta t$  the value of the bank account  $X(t + \Delta t)$  will be:

$$X(t + \Delta t) = X(t) + X(t)r(t)\Delta t \quad 1.1$$

and applying the limit  $\Delta t \rightarrow 0$  we get:

---

<sup>1</sup> we assume that both ZCB and Banking account returns are risk free

$$X(t)r(t) = \lim_{\Delta t \rightarrow 0} \frac{X(t + \Delta t) - X(t)}{\Delta t} = \frac{d}{dt}X(t) \quad 1.2$$

Then the value of the bank account  $X(t)$  is the solution of the following Cauchy problem:

$$\begin{cases} \frac{d}{dt}X(t) = X(t)r(t) \\ X(0) = X_0 \end{cases} \quad 1.3$$

Solving this problem, we get the bank account process:

$$X(t) = X_0 e^{\int_0^t r(s) ds} \quad 1.4$$

By analogy, letting  $P(t,T) = X(t)$  and  $X_0 = P(0,T)$ , we get the price of a bond as function of  $r(t)$ :

$$P(t,T) = E \left[ e^{-\int_t^T r(s) ds} \right] \quad 1.5$$

If we assume that rates were constant, the price would be simply given by  $e^{-(T-t)r}$ . However, in the real world, the short term interest rate is stochastic and then we need a model which describe its dynamic in order to compute the value of a bond.

Another important relation is between the short term interest rate  $r(t)$  and the (continuously compounded) spot interest rate  $R(t,T)$ , defined as:

$$R(t,T) = -\frac{\log P(t,T)}{T-t} \quad 1.6$$

From which follows:

$$P(t,T) = e^{-(T-t)R(t,T)} \quad 1.7$$

Unlike  $r(t)$ , the spot interest rate  $R(t,T)$  can be directly calculated by the price of the ZCB observed in the market without any use of models which describe its dynamic. Now, we need to give a deterministic relation between the price of a ZCB and  $r(t)$  in order to derive the relation between  $r(t)$  and  $R(t,T)$ . To do this we will use the so-called *Affine term structure models*. For

these models, the spot price at time  $t$ , with  $0 \leq t < T$ , of a zero coupon bond (ZCB) with maturity  $T$  can be expressed as:

$$P(t, T) = \exp (A(t, T) - B(t, T)r(t)) \quad 1.8$$

Where  $A$  and  $B$  are deterministic function of  $t$  and  $T$ . The values of  $A$  and  $B$  depend on the affine model used to describe the dynamic of  $r(t)$ .

Then , given an affine model for the short interest rate, the relation between  $r(t)$  and  $R(t, T)$  is given by equating 1.7 and 1.8, and result in:

$$R(t, T) = -\frac{A(t, T)}{T - t} + \frac{B(t, T)}{T - t}r(t) \quad 1.9$$

This relation will be useful when one want to calibrate an affine model from the data or would simulate a path of  $R(t, T)$  using an affine model. In fact the only data we can observe in the market is  $R(t, T)$ , and it is crucial to understand the relation between the observed data and the object to be estimated or simulated. However, the continuously compounded notation (1.6) is used only in theoretical contexts. In the market practice, the bond price is expressed using the *simply compounded interest rate* defined as:

$$L(t, T) = \frac{1 - P(t, T)}{(T - t)P(t, T)} \quad 1.10$$

From which follows that the zero coupon bond price is given by:

$$P(t, T) = \frac{1}{1 + (T - t)L(t, T)} \quad 1.11$$

Rearranging the equation 1.11 and substituting the value of  $P(t, T)$  as in 1.8 we get the explicit relation between  $r(t)$  and  $L(t, T)$ :

$$L(t, T) = \frac{e^{-A(t, T) + B(t, T)r(t)} - 1}{T - t} \quad 1.12$$

Now, In order to link the theoretical quantity  $r(t)$  with the real world quantity  $R(t,T)$  (or  $L(t,T)$ ) we need to find the value of  $A(t,T)$  and  $B(t,T)$  of 1.8 for the model considered. The values of these functions must be such that the prices generated do not allow arbitrage opportunity. To meet this issue, we have to derive the value of  $A(t,T)$  and  $B(t,T)$  starting from the *term structure equation* for an arbitrage free bond market.

### 1.1 The term structure equation

The term structure equation is a Partial Differential Equation, which have as solution the family of prices  $\{P(t,T); 0 \leq t \leq T, T > 0\}$ . Now, let us see how to derive this equation in an arbitrage free bond market and how to combine this with an affine term structure model. Assuming that the short interest rate follows the following stochastic differential equation under the real measure  $P$ :

$$dr(t) = \mu(r,t)dt + \sigma(r,t)dW(t) \quad 1.13$$

where  $\mu(r,t)$  and  $\sigma(r,t)$  are two predictable function and  $dW(t)$  is a standard P-Brownian motion. For a fixed time  $T$ , the price of a zero coupon bond is only function of  $t$  and  $r$ , thus  $P(t,T) = f(t,r)$ . Assuming  $f(t,r)$  twice continuously differentiable, we can apply Ito's lemma on 1.13 obtaining:

$$\begin{aligned} df(t,r) &= f_t dt + f_r dr + \frac{1}{2} f_{rr} (dr)^2 \\ &= f_t dt + f_r (\mu(r,t)dt + \sigma(r,t)dW(t)) + \frac{1}{2} f_{rr} (\mu(r,t)dt + \sigma(r,t)dW(t))^2 \\ &= \left( f_t + f_r \mu(r,t) + \frac{1}{2} f_{rr} \sigma^2(r,t) \right) dt + f_r \sigma(r,t) dW(t) \quad 1.14 \end{aligned}$$

Where  $f_t$  and  $f_r$  are respectively the partial derivatives with respect to  $t$  and  $r$ . Finally, letting:

$$\begin{aligned} \tilde{\mu} &= \frac{f_t + f_r \mu(r,t) + \frac{1}{2} f_{rr} \sigma^2(r,t)}{f(t,r)} \quad 1.15 \\ \tilde{\sigma} &= \frac{f_r \sigma(r,t)}{f(t,r)} \end{aligned}$$

we can rewrite 1.14 as:

$$df(t,r) = f(t,r)\tilde{\mu}dt + f(t,r)\tilde{\sigma}dW(t) \quad 1.16$$



Now, we proceed by constructing a risk-free self-financing portfolio of bonds [7] having different times to maturity S and T (a similar procedure is used in the stock market to derive the Black and Scholes equation), and imposing the no arbitrage condition which require that this portfolio must have a rate of return equal to the short interest rate, we will get the term structure equation for the arbitrage free bond market. Let  $Q_S$  and  $Q_T$  the relative fraction amount of two bond  $f_S(t, r)$  and  $f_T(t, r)$ , and denoting by  $\Pi$  the value of the portfolio consisting in this two bonds, i.e  $\Pi = Q_S f_S(t, r) + Q_T f_T(t, r)$ , then the return of this portfolio is given by proportional return of the two bonds:

$$\frac{d\Pi}{\Pi} = Q_S \frac{df_S(t, r)}{f_S(t, r)} + Q_T \frac{df_T(t, r)}{f_T(t, r)} \quad 1.17$$

Substituting 1.16 into 1.17 we get:

$$\begin{aligned} \frac{d\Pi}{\Pi} = & Q_S \frac{\left( f_{S t} + f_{S r} \mu(r, t) + \frac{1}{2} f_{S r r} \sigma^2(r, t) \right) dt + f_{S r} \sigma(r, t) dW(t)}{f_S(t, r)} \\ & + Q_T \frac{\left( f_{T t} + f_{T r} \mu(r, t) + \frac{1}{2} f_{T r r} \sigma^2(r, t) \right) dt + f_{T r} \sigma(r, t) dW(t)}{f_T(t, r)} \end{aligned}$$

Which can be rewritten as:

$$\frac{d\Pi}{\Pi} = (Q_S \tilde{\mu}_S + Q_T \tilde{\mu}_T) dt + (Q_S \tilde{\sigma}_S + Q_T \tilde{\sigma}_T) dW(t) \quad 1.18$$

Where  $\tilde{\mu}$  and  $\tilde{\sigma}$  are as in 1.15. Now we want to eliminate the  $dW(t)$  term in order to get the risk-free portfolio, to do this we can only choose the quantities  $Q_S$  and  $Q_T$  in such way to have:

$$\begin{aligned} Q_S \tilde{\sigma}_S + Q_T \tilde{\sigma}_T &= 0 \\ \text{with} \quad Q_S + Q_T &= 1. \end{aligned}$$

The solution of this linear system of equations is given by:

$$Q_S = \frac{\tilde{\sigma}_T}{\tilde{\sigma}_T - \tilde{\sigma}_S},$$

$$Q_T = -\frac{\tilde{\sigma}_s}{\tilde{\sigma}_T - \tilde{\sigma}_s},$$

and substituting these solution into equation 1.18 we get the risk free portfolio dynamic:

$$\frac{d\Pi}{\Pi} = \left( \frac{\tilde{\sigma}_T}{\tilde{\sigma}_T - \tilde{\sigma}_s} \tilde{\mu}_s - \frac{\tilde{\sigma}_s}{\tilde{\sigma}_T - \tilde{\sigma}_s} \tilde{\mu}_T \right) dt \quad 1.19$$

Since the portfolio following this dynamic is risk-free, it must offer the same return of the bank account (risk-free investment). As argued in the previous section this return of the bank account is equal to the short term interest rate  $r(t)$ , and then must hold true:

$$r(t) = \frac{\tilde{\sigma}_T \tilde{\mu}_s - \tilde{\sigma}_s \tilde{\mu}_T}{\tilde{\sigma}_T - \tilde{\sigma}_s}$$

Or, written in a different way:

$$\frac{\tilde{\mu}_s - r(t)}{\tilde{\sigma}_s} = \frac{\tilde{\mu}_T - r(t)}{\tilde{\sigma}_T} \quad 1.20$$

The value of both-side hand value does not depend on the choice of the time to maturity. In fact, if the time to maturity of the left-hand side changes, the right-hand side remains unchanged and then even the left-hand side must remain the same, so we can define the process:

$$\lambda(t) = \frac{\tilde{\mu}_T - r(t)}{\tilde{\sigma}_T} \quad 1.21$$

which holds for all  $t$  and  $T$ . The process  $\lambda(t)$  is called *market price of risk*.

Now, inserting equation 1.15 into 1.21 and rearranging, specifying the boundary  $P^T(r, T) = 1$ , we obtain the so called “*term structure equation*” in an arbitrage free bond market:

$$\begin{cases} P_t^T + \{\mu(r, t) - \lambda(t)\sigma(r, t)\}P_r^T + \frac{1}{2}\sigma^2(r, t) P_{rr}^T - rP^T = 0 \\ P^T(r, T) = 1 \end{cases} \quad 1.22$$

Which represent the family of bond prices ( $P(t, T)$ ;  $0 \leq t \leq T$ ,  $T > 0$ ) which does not allow arbitrage opportunity in the bond market.

So, if we consider the following P-dynamic for the short term interest rate

$$dr(t) = \mu(r, t)dt + \sigma(r, t)dW(t)$$

Assuming that this model possesses an ATS so that:

$$P(t, T) = \exp( A(t, T) - B(t, T) r(t) )$$

If we introduce:

$$\begin{aligned} \hat{\mu} &= \mu(r, t) - \lambda\sigma(r, t) \quad (\text{drift term under Q measure}) \\ \hat{\sigma} &= \sigma(r, t) \end{aligned}$$

We can rewrite the term structure equation as

$$\begin{cases} P_t^T + \hat{\mu} P_r^T + \frac{1}{2} \hat{\sigma}^2 P_{rr}^T - rP^T = 0 \\ P^T(r, T) = 1 \end{cases} \quad 1.23$$

Now we can compute the partial derivatives  $P_t$ ,  $P_r$  and  $P_{rr}$  as:

$$P_t^T = (A_t(t, T) - B_t(t, T) r) P^T$$

$$P_r^T = -B(t, T) P^T$$

$$P_{rr}^T = B^2(t, T) P^T$$

And substituting into 1.23 we get:

$$A_t(t, T) - (1 + B_t(t, T))r - \hat{\mu} B(t, T) + \frac{1}{2} \hat{\sigma}^2 B^2(t, T) = 0 \quad 1.21$$

With boundary values

$$A(T, T) = 0$$

$$B(T, T) = 0$$

This equation give us the relation between  $A$ ,  $B$ ,  $\hat{\mu}$  and  $\hat{\sigma}^2$ , and then allow us to compute the price of a ZCB for an ATS model. In the next section, we use it in order to compute the zero-coupon bond price for the model studied.

## 1.2 Short Interest Rate Models

Based on the stochastic differential equation 1.10, in the literature were built several parametric models by giving to the functions  $\mu(r, t)$  and  $\sigma(r, t)$  a definite form. Most of the models developed imply dynamics for the short-term interest rate  $r(t)$  that can be included within the following SDE:

$$dr(t) = \alpha (\mu - r(t))dt + \sigma r^\gamma dW(t) \quad 1.25$$

Where  $\alpha$ ,  $\mu$ ,  $\sigma$ ,  $\gamma$  are constant parameters and  $dW(t)$  is a standard P-Brownian motion. By assigning different values to the parameters  $[\alpha, \mu, \sigma, \gamma]$ , we get different models. For  $\mu=0$ ,  $\gamma = 0$  we get the Merton(1973) model; for  $\gamma = 0$  we get the mean reverting process used by Vasicek model and for  $\gamma = 1/2$  we get the square root process which appears in the Cox, Ingersoll and Ross (CIR) model. The last two are the most popular model in the interest rate markets and are those we will consider in this work.

### 1.2.1 The Vasicek Model

In this model the short term interest rate  $r(t)$  is solution of the SDE:

$$dr(t) = \alpha(\mu - r(t))dt + \sigma dW^Q(t) \quad 1.26$$

where  $W(t)$  is a Q-standard Brownian motion<sup>2</sup> and  $\alpha, \mu, \sigma$  are strictly positive constants. The constant  $\mu$  represents the *risk-neutral* long-term mean risk-free rate,  $\alpha$  represents the rate at which  $r(t)$  reverts back to this long-term mean and  $\sigma$  represents the volatility of short-term interest rate. One of the main advantages of this model is to be a *mean reverting* process, but against a negative characteristic is to have positive probability that the interest rate can assume negative values.

---

<sup>2</sup> We assume  $\sigma > 0$

However, the feature that made the model of Vasicek one of the most used by traders is undoubtedly due to the simplicity of the model and on the fact of giving a closed form solution for pricing. One of the goals of this work consist in comparing the forward rate paths generate by this model with those generated by the SABR model. Then, we need a discrete version of 1.26 that allow us to simulate the short term interest rate. Phillips[50] showed that the exact discrete model corresponding to 1.26 is given by :

$$r(t + \Delta t) = r(t)e^{-\alpha\Delta t} + \mu(1 - e^{-\alpha\Delta t}) + \sigma\sqrt{(1 - e^{-\alpha\Delta t})/2\alpha}\epsilon_t \quad 1.27$$

where  $\epsilon_t \sim i. i. d N(0,1)$ . Looking at 1.27 it is easy to see that  $r(t)$  is normally distributed with conditional mean and variance respectively given by :

$$E[r(t + \Delta t|r(t))] = e^{-\alpha\Delta t}r(t) + (1 - e^{-\alpha\Delta t})\mu$$

$$VAR[r(t + \Delta t|r(t))] = \frac{\sigma^2}{2\alpha} (1 - e^{-2\alpha\Delta t})$$

The availability of a closed form expression for the transition density will be crucial when we will try to estimate parameters using MLE method. Now we want to connect the short-term interest rate given by this model with the price of the zero-coupon bond. We know that the price of a ZCB in an arbitrage free market must be solution of the term structure equation 1.24. For this model, we have:

$$\hat{\mu} = \alpha (\mu - r)$$

$$\hat{\sigma} = \sigma$$

And the equation 1.24 became:

$$A_t(t, T) - (1 + B_t(t, T))r(t) - \alpha (\mu - r(t))B(t, T) + \frac{1}{2}\sigma^2 B(t, T)^2 = 0$$

which, collecting the  $r(t)$  terms, result in:

$$A_t(t, T) - \left(1 + B_t(t, T) - \alpha B(t, T)\right)r(t) - \alpha\mu B(t, T) + \frac{1}{2}\sigma^2 B(t, T)^2 = 0 \quad 1.28$$

This equation must holds for every  $t$  and  $r$  , so we have that the coefficient of  $r$  must be equal to zero. Thus, we obtain the two ODE Cauchy problems:

$$1 + B_t(t, T) - \alpha B(t, T) = 0$$

$$\text{With boundary condition } B(T, T) = 0 \quad 1.29$$

and

$$A_t(t, T) - \alpha \mu B(t, T) + \frac{1}{2} \sigma^2 B(t, T)^2 = 0$$

$$\text{With boundary condition } A(T, T) = 0 \quad 1.30$$

Where  $A_t$  and  $B_t$  are the partial derivatives with respect to  $t$ .

For equation 1.29 it is easy to show that:

$$B(t, T) = \frac{1 - e^{-a(T-t)}}{a} \quad 1.31$$

To solve 1.27 we have to integrate, then we have:

$$A(t, T) = \frac{1}{2} \sigma^2 \int_t^T B^2(u, T) du - \alpha \mu \int_t^T B(u, T) du \quad 1.32$$

And Substituting 1.31 into 1.32 follows:

$$\begin{aligned} A(t, T) &= \frac{1}{2} \sigma^2 \int_t^T \left( \frac{1 - e^{-a(T-u)}}{a} \right)^2 du - \alpha \mu \int_t^T \left( \frac{1 - e^{-a(T-u)}}{a} \right) du \\ &= \frac{1}{2a^2} \sigma^2 \int_t^T (1 - e^{-a(T-u)} + e^{-2a(T-u)}) du - \mu \int_t^T 1 du + \mu \int_t^T e^{-a(T-u)} du \\ &= \frac{\sigma^2}{2a^2} (T-t) - \frac{\sigma^2}{2a^2} 2B(t, T) + \frac{\sigma^2}{2a^2} \left( \frac{1 - e^{-2a(T-t)}}{2a} \right) - \mu(T-t) + \mu B(t, T) \\ &= \left( \frac{\sigma^2}{2a^2} - \mu \right) ((T-t) - B(t, T)) - \frac{\sigma^2}{2a^2} B(t, T) + \frac{\sigma^2}{2a^2} \left( \frac{1 - e^{-2a(T-t)}}{2a} \right) \\ &= \left( \frac{\sigma^2}{2a^2} - \mu \right) ((T-t) - B(t, T)) - \frac{\sigma^2}{4a} B^2(t, T) \quad 1.33 \end{aligned}$$

Hence, the price of a bond in the Vasicek model is given by:

$$P(t, T) = \exp(A(t, T) - B(t, T)r(t)) \quad 1.34$$

with

$$A(t, T) = \left( \frac{\sigma^2}{2a^2} - \mu \right) ((T - t) - B(t, T)) - \frac{\sigma^2}{4a} B^2(t, T)$$

and

$$B(t, T) = \frac{1 - e^{-a(T-t)}}{a}$$

In order to get the P-dynamic for the continuously compounded interest rate  $R(t, T)$  we use 1.9, obtaining:

$$dR(t, T) = a \left( \frac{B_t(t, T)}{T - t} - \frac{A_t(t, T)}{T - t} - R(t, T) \right) dt + \frac{B_t(t, T)}{T - t} \sigma dW_t \quad 1.35$$

Which can be rewritten in the same form of  $r(t)$ :

$$dR(t, T) = a(\check{\mu} - R(t, T))dt + \check{\sigma}dW_t \quad 1.36$$

With:

$$\check{\mu} = \frac{B_t(t, T)}{T - t} - \frac{A_t(t, T)}{T - t}$$

$$\check{\sigma} = \frac{B_t(t, T)}{T - t} \sigma$$

Then, we can write the dynamic of  $R(t, T)$  in the same form of  $r(t)$  simply changing the parameters. This means that both  $r(t)$  and  $R(t, T)$  have the same distribution, and in this particular case are normal distributed. This aspect is very useful when we try to estimate the parameter from real data with a method like MLE, which use the marginal distribution to obtain the parameters. The link between  $L(t, T)$  and  $r(t)$  in the Vasicek model can be derived from equation 1.12, we get:

$$L(t, T) = \frac{e^{-A(t, T) + B(t, T)r(t)} - 1}{T - t} \quad 1.37$$

where  $A(t, T)$  and  $B(t, T)$  are as in 1.34.

### 1.2.2 The Cox, Ingersoll and Ross model (CIR model)

The stochastic differential equation of the model proposed by Cox, Ingersoll and Ross[16] is:

$$dr(t) = \alpha (\mu - r(t))dt + \sigma\sqrt{r(t)}dW^Q(t) \quad 1.38$$

$$\text{with } 2a\mu > \sigma^2$$

where  $\alpha$ ,  $\mu$  and  $\sigma$  are positive constant and  $W^Q(t)$  is a standard Brownian motion under the risk neutral measure Q. The condition  $2a\mu > \sigma^2$  is imposed to avoid that the short term interest rate reach the zero value. As for the Vasicek model, we need a discrete version of 1.38 to simulate  $r(t)$ . Phillips [50] show that the discrete version of 1.38 is given by:

$$r(t + \Delta t) = e^{-a\Delta t}r(t) + \mu(1 - e^{-a\Delta t}) + \sigma \int_t^{t+\Delta t} e^{-a(t+\Delta t-s)}\sqrt{r(s)} dW^Q(s) \quad 1.39$$

It is possible show that the conditional distribution of this process is, up to a scale factor, a non-central chi-squared distribution [19] :

$$p(r(t + \Delta t)|r(t)) \sim \frac{\sigma^2(1 - e^{-a\Delta t})}{4a} \chi_d^2 \left( \frac{4ae^{-a\Delta t}}{\sigma^2(1 - e^{-a\Delta t})} r(t) \right) \quad 1.40$$

Where  $\chi_d^2(\lambda)$  denotes the noncentral chi-squared distribution with  $d$  degrees of freedom equal to

$$d = \frac{4\mu a}{\sigma^2}$$

and non-centrality parameter  $\lambda$  given by

$$\lambda = \frac{4ae^{-a\Delta t}}{\sigma^2(1 - e^{-a\Delta t})} r(t)$$

Straightforward calculation give the expected value and variance of  $r(t)$  as [16]:

$$E(r(t + \Delta t)|r(t)) = r(t)e^{-a\Delta t} + \mu(1 - e^{-a\Delta t}),$$

$$\text{var}(r(t + \Delta t)|r(t)) = r(t) \left( \frac{\sigma^2}{a} \right) (e^{-a\Delta t} - e^{-2a\Delta t}) + \mu \left( \frac{\sigma^2}{2a} \right) (1 - e^{-a\Delta t})^2.$$



Another way to represent the distribution of  $r(t + \Delta t)|r(t)$ , which can be particularly convenient for the simulation, is given by Feller[19]. He show that the *transition density* of the square root model can also be expressed as:

$$p(r(t + \Delta t)|r(t)) = ce^{-u-v} \left(\frac{v}{u}\right)^{\frac{q}{2}} I_q(2\sqrt{uv}) \quad 1.41$$

Where:

$$c = \frac{2a}{\sigma^2(1 - e^{-a\Delta t})}, \quad u = cr(t)e^{-a\Delta t}, \quad v = cr(t + \Delta t), \quad q = \frac{2a\mu}{\sigma^2} - 1$$

And  $I_q(2\sqrt{uv})$  is the modified Bessel function of the first kind and order  $q$ .

Now, as made for the Vasicek model, we are interested in computing the value of the function  $A(t,T)$  and  $B(t,T)$  in order to determine the price of a zero coupon bond and the relation between  $r(t,T)$  and  $R(t,T)$ .

Then, by solving the term structure equation 1.24 with:

$$\hat{\mu} = \alpha (\mu - r)$$

$$\hat{\sigma} = \sigma\sqrt{r}$$

we get

$$A_t(t,T) - (1 + B_t(t,T))r - \alpha(\mu - r)B(t,T) + 1/2r\sigma^2B(t,T)^2 = 0.$$

Which can be rewritten as

$$A_t(t,T) - \alpha\mu B(t,T) - r \left( 1 + B_t(t,T) - a B(t,T) - \frac{1}{2\sigma^2 B(t,T)^2} \right) = 0. \quad 1.42$$

Now we can made the same observation as for the Vasicek model obtaining the follow two ODE Cauchy problems:

$$A_t(t,T) - \alpha\mu B(t,T) = 0$$

$$\text{With boundary condition } A(T,T) = 0 \quad 1.43$$

and

$$1 + B_t(t,T) - a B(t,T) - 1/2\sigma^2 B(t,T)^2 = 0$$

$$\text{With boundary condition } B(T,T) = 0 \quad 1.44$$

Here we skip the passages and give directly the solution for the term structure in the CIR model, which is given by

$$P(t, T) = C(t, T) \exp(-B(t, T)r(t)) \quad 1.45$$

where

$$C(t, T) = \left( \frac{2\gamma e^{(a+\gamma)((T-t)/2)}}{(a+\gamma)(e^{\gamma(T-t)} - 1) + 2\gamma} \right)^{2a\mu/\sigma^2}$$

$$B(t, T) = \left( \frac{2(e^{\gamma(T-t)} - 1)}{(a+\gamma)(e^{\gamma(T-t)} - 1) + 2\gamma} \right)$$

$$\gamma = \sqrt{a^2 + 2\sigma^2}$$

Now, as for the Vasicek model, we can get the stochastic differential equation for  $R(t, T)$ , that is

$$dR = \frac{B_t(t, T)}{(T-t)} \left( a \left( \mu - \frac{(T-t)R + \ln(C_t(t, T))}{B_t(t, T)} \right) dt + \sigma \sqrt{\frac{(T-t)R + \ln(C_t(t, T))}{B_t(t, T)}} dW_t \right)$$

Which cannot be written in the same form of 1.33, in fact in the CIR model  $r(t)$  follows a chi-squared distribution. As for the Vasicek model we need also to derive the relation between  $r(t)$  and  $L(t, T)$ , this is given by combining equation 1.12 with equation 1.45. The result is:

$$L(t, T) = \frac{1}{T-t} \left( \frac{e^{B(t, T)r(t)}}{C(t, T)} - 1 \right)$$

### 1.3 Forward rate

The main object of this work is the comparison between the SABR model with the ATS model. In particular, we want study the forward-rate path generated by these models, and figure out if any of these paths are compatible with those generated by the other models. To do this we need to implement a Monte Carlo simulation of forward interest rate with the ATS model, therefore we have to find the relation between the short rate of interest and the forward rate. The simply compounded forward interest rate  $f(t, S, T)$  is the rate decided at present time  $t$  and applied over a future period time  $[S, T]$ , with  $t \leq S \leq T$ . Let us determine the no-arbitrage value of this

rate using the bonds prices  $P(t, T)$  for various maturity  $T > t$ . Suppose that at time  $t$  our wealth is equal to zero and that we carry out the following trading strategy:


- 1- We borrow at time  $t$  the amount  $P(t, T)$  by shortselling one unit of bond with maturity  $T$ , then we must refund the quantity  $1\text{€}$  at time  $T$ .
- 2- We invest the amount  $P(t, T)$  over the period  $[t, S]$  by buying a (fractional) quantity  $P(t, T)/P(t, S)$  of a bond with maturity  $S$  priced  $P(t, S)$  at time  $t$ . This second operation will yield the amount

$$\frac{P(t, T)}{P(t, S)} \times 1\text{€}$$

at time  $S$ .

In order to avoid any arbitrage opportunity the simply compounded forward rate  $f(t, S, T)$  must be chosen so that :

$$\frac{P(t, T)}{P(t, S)} (1 + (T - S)f(t, S, T)) - 1\text{€} = 0 \quad 1.46$$



If this equality were not satisfy, there would be arbitrage opportunity, in fact:

a) If

$$\frac{P(t, T)}{P(t, S)} (1 + (T - S)f(t, S, T)) - 1\text{€} > 0$$

By starting from an initial wealth equal to zero, using the strategy described above it is possible to get a profit at time  $T$  without any risk.

b) If

$$\frac{P(t, T)}{P(t, S)} (1 + (T - S)f(t, S, T)) - 1\text{€} < 0$$

One can borrow the quantity  $P(t, S)$  shortselling the bond with maturity  $S$  and buy the quantity  $P(t, S)/P(t, T)$  of the bond with maturity  $T$ . At time  $t$  the value in  $S$  of the bond with maturity  $T$  is given by  $1 \times \exp -(T - S)f(t, S, T)$ , then in  $S$  I will receive the quantity:

$$\frac{P(t,S)}{P(t,T)} \frac{1}{1 + (T-S)f(t,S,T)}$$

Which is greater than 1 (because is the inverse of a value less than 1), and I will refund 1 at time S. In this way I have get a profit without any risk. Rearranging 1.46 we obtain the explicit formula for the no arbitrage continuously compounded forward interest rate, that is:

$$f(t,S,T) = \frac{1}{T-S} \left( \frac{P(t,S)}{P(t,T)} - 1 \right) \quad 1.47$$

Now, we can determine the bond price  $P(t,S)$  and  $P(t,T)$  in the ATS models and using 1.42 we can find the relation between  $r(t)$  and  $f(t,S,T)$  for each model.

### 1.3.1 Vasicek forward rate simulation

In this model, the forward rate is given by:

$$\begin{aligned} f(t,S,T) &= \frac{1}{T-S} \left( \frac{1 + (T-t)L(t,T)}{1 + (S-t)L(t,S)} - 1 \right) = \\ &= \frac{1}{T-S} \left( \frac{e^{-A(t,T)+B(t,T)r(t)}}{e^{-A(t,S)+B(t,S)r(t)}} - 1 \right) \\ &= \frac{1}{T-S} \left( \exp(A(t,S) - B(t,S)r(t) - A(t,T) + B(t,T)r(t)) - 1 \right) \\ &= \frac{1}{T-S} \left( \exp \left( A(t,S) - A(t,T) + (B(t,T) - B(t,S))r(t) \right) - 1 \right) \quad 1.48 \end{aligned}$$

And the substitution of the value  $A(t,S)$  and  $B(t,T)$  as given by 1.34 give as the relation between the short interest rate and the forward rate:

$$\begin{aligned}
f(t, S, T) = & \frac{1}{T-S} \exp \left( r(t) \left( \frac{1 - e^{-a(T-t)}}{a} - \frac{1 - e^{-a(S-t)}}{a} \right) \right. \\
& + \left( \frac{\sigma^2}{2a^2} - \mu \right) \left( (S-t) - \frac{1 - e^{-a(S-t)}}{a} \right) - \frac{\sigma^2}{4a} \left( \frac{1 - e^{-a(S-t)}}{a} \right)^2 \\
& \left. - \left( \frac{\sigma^2}{2a^2} - \mu \right) \left( (T-t) - \frac{1 - e^{-a(T-t)}}{a} \right) - \frac{\sigma^2}{4a} \left( \frac{1 - e^{-a(T-t)}}{a} \right)^2 - 1 \right) \quad 1.49
\end{aligned}$$

This formula might seem very complicated, but contrary to appearances, it is very simple to implement since all the object inside it are simply numbers.

Now we have all the tools needed for simulating the forward interest rate using the Vasicek model. We implement the simulation follow the following steps:

- 1) Simulate the path of  $r(t)$  using the exact discrete model given by 1.27;
- 2) For each  $t$  calculate the value of  $A(t, S), A(t, T), B(t, S), B(t, T)$ ;
- 3) For each  $t$  calculate the value of the ZCBs with maturity  $S$  and  $T$ ;
- 4) For each  $t$  calculate the forward rate using 1.49.

### 1.3.2 CIR forward rate simulation

Before talking about the simulation of the forward rate, we must deepen the topic of the simulation of the interest rate using the CIR model. Fortunately, many path simulation methods for the CIR process are been developed in the literature in the last years. The main reason of this is that the Heston model [28] has become in the recent years one of the most used stochastic volatility model in the stock market. In this model, the volatility of the asset follows a CIR process, then to sample path we can use the same methods used to simulate the path of the volatility in the Heston model. A classical approach to constructing simulation schemes for SDE 1.38 involves the application of the *Euler approximation*, which give the following process:

$$r(t + \Delta t) = r(t) + a(\mu - r(t))\Delta t + \sigma\sqrt{r(t)\Delta t}Z, \quad Z \sim N(0,1). \quad 1.50$$

The main problem of this approximation is due to the fact that certain realization  $z$  of the random variable  $Z \sim N(0,1)$  can lead to negative values of  $r$  when

$$z < -\frac{r(t) + a(\mu - r(t))\Delta t}{\sigma\sqrt{r(t)\Delta t}}.$$

To address this problem, many authors have proposed different simulation scheme. Van Haastrecht et. al [51] propose a non-central  $\chi^2$  inversion scheme, Glasserman et.al[21]use the gamma expansion scheme , Broadie and Kaya [11] propose an exact scheme, Andersen [6] use instead a quadratic exponential scheme. The last two are the most reliable simulation scheme and we will implement the Broadie and Kaya one for our path simulation. In their work the authors sample the simulated value of  $r(t)$  directly from the non-central chi-squared distribution. To do this they exploit the work of Johnson et al [37], which show that given a chi-squared random variable  $\chi_d^2$  with  $d > 1$  degree of freedom, the following equation is valid:

$$\chi_d^2(\lambda) = \chi_1^2(\lambda) + \chi_{d-1}^2 \quad 1.51$$

Along the line of Patnaik[42] we use the following approximation:

$$\chi_d^2(\lambda) \approx (Z - \sqrt{\lambda})^2 \quad 1.52$$

where  $Z$  is a standard normal random variable. Then, the algorithm we will follow to simulate  $r(t)$  is composed by the following steps:

- 1) Calculate  $d$  and  $\lambda$  and verify that  $d > 1$ ;
- 2) Draw a vector of standard normal random variable  $z$ ;
- 3) Draw a vector of chi-squared random variable  $\chi_{d-1}^2$  having  $d-1$  degree of freedom;
- 4) Sample  $r(t)$  from  $\chi_d^2(\lambda)$  using the relations:

$$\chi_d^2(\lambda) = (Z - \sqrt{\lambda})^2 + \chi_{d-1}^2 \quad 1.53$$

and,

$$p(r(t + \Delta t)|r(t)) = \frac{\sigma^2(1 - e^{-a\Delta t})}{4a} \chi_d^2(\lambda) \quad 1.54$$

With

$$d = \frac{4\mu a}{\sigma^2}; \quad \lambda = \frac{4ae^{-a\Delta t}}{\sigma^2(1 - e^{-a\Delta t})}r(t).$$

- 5) For each  $t$  calculate the value of  $C(t,T)$ ,  $C(t,S)$ ,  $B(t,T)$ ,  $B(t,S)$ ;
- 6) For each  $t$  calculate the values of ZCB with maturity  $S$  and  $T$ ;
- 7) For each  $t$  calculate the value of the continuously compounded forward rate from:

$$\begin{aligned} f(t,S,T) &= \frac{1}{T-S} \left( \frac{1 + (T-t)L(t,T)}{1 + (S-t)L(t,S)} - 1 \right) \\ &= \frac{1}{T-S} \left( \frac{\frac{e^{B(t,T)r(t)}}{C(t,T)}}{\frac{e^{B(t,S)r(t)}}{C(t,S)}} - 1 \right) \\ &= \frac{1}{T-S} \left( \frac{C(t,S) e^{B(t,T)r(t)}}{C(t,T) e^{B(t,S)r(t)}} - 1 \right) \\ &= \frac{1}{T-S} \left( \frac{C(t,S)}{C(t,T)} e^{r(t)(B(t,T)-B(t,S))} - 1 \right) \quad 1.55 \end{aligned}$$

**MATLAB code****Vasicek forward rate**

```

function pathFvas =
vasicekForwardSimulation(r0,k,mu,sigma,T_set,T_ex);

clc
clear all
close all

%r0=0.03;
%k=-0.18;
%mu=-0.080;
%sigma=0.04;
%T_set=2;
%T_ex=1;
%time_step=250;
%delta_t=T_ex/time_step;
%n_sim=500;

path_R = zeros(n_sim,time_step+1);
path_R(:,1) = r0;
zcb_set= zeros(n_sim,time_step*T_ex);
zcb_ex= zeros(n_sim,time_step*T_ex);
forward_rate=zeros(n_sim,time_step*T_ex);
B_Tex=zeros(n_sim,time_step*T_ex);
B_Tset=zeros(n_sim,time_step*T_ex);
A_Tex=zeros(n_sim,time_step*T_ex);
A_Tset=zeros(n_sim,time_step*T_ex);

for j=1:n_sim
    for i=1:time_step
        path_R(j,i+1)=path_R(j,i)*exp(-k*delta_t)+ mu*(1-
            exp(-k*delta_t))+sigma*sqrt((1-exp(-
                2*k*delta_t))/2*k)*randn();
    end
end

for j=1:n_sim
    for i=1:time_step*T_ex
        B_Tex(j,i)= (1-exp(-k*(time_step*T_ex-
            i+1)/time_step))/k;
        B_Tset(j,i)= (1-exp(-k*(time_step-
            i+1)/time_step))/k;
        A_Tex(j,i)= (((sigma^2)/(2*k^2))-
            mu)*((time_step*T_ex-i+1)/(time_step))-B_Tex(j,i))-
            ((sigma^2)/(4*k))*B_Tex(j,i)^2;
        A_Tset(j,i)= (((sigma^2)/(2*k^2))-mu)*((time_step-
            i+1)/(time_step))-B_Tset(j,i))-
            ((sigma^2)/4*k)*B_Tset(j,i)^2;
    end
end

```



```

_____SimplycompoundedForwardrate_____
zcb_ex(j,i)=1/(exp(-
    A_Tex(j,i)+path_R(j,i)*B_Tex(j,i)));
zcb_set(j,i)=1/(exp(-
    A_Tset(j,i)+path_R(j,i)*B_Tset(j,i)));
forward_rate(j,i)=(1/(T_set-
    T_ex))*(zcb_ex(j,i)/zcb_set(j,i)-1);

_____continuouslycompoundedfarwardrate_____
zcb_ex(j,i)=exp(A_Tex(j,i)-path_R(j,i)*B_Tex(j,i));
zcb_set(j,i)=exp(A_Tset(j,i)-
    path_R(j,i)*B_Tset(j,i));
forward_rate(j,i)=(log(zcb_ex(j,i))-
    log(zcb_set(j,i)))/(T_set-T_ex);
end
end

figure(1);
plot (forward_rate')
end

```

### CIR forward rate (Kaya scheme)

```

function pathFcir =
CirForwardSimulation(r0,k,mu,sigma,T_set,T_ex);
clc
close all
clear all

%r0=0.03;
%k=0.25;
%mu=0.0270;
%sigma=0.03;
%T_set=2;
%T_ex=1;
time_step=250;
delta_t=T_set/time_step;
n_sim=1000;
c=((sigma^2)*(1-exp(-k*delta_t)))/(4*k)
path_R = zeros(n_sim,time_step+1);
path_R(:,1) = r0;
gamma=sqrt((k^2)+2*(sigma^2))

Z=randn(n_sim,time_step);
lambda=zeros(n_sim,time_step);
c=((sigma^2)*(1-exp(-k*delta_t)))/(4*k)
d=(4*mu*k)/(sigma^2);

```

```

x=chi2rnd(d-1,[n_sim,time_step])

path_R = zeros(n_sim,time_step+1);
path_R(:,1) = r0;
zcb_set= zeros(n_sim,time_step*T_ex);
zcb_ex= zeros(n_sim,time_step*T_ex);
forward_rate=zeros(n_sim,time_step*T_ex);
B_Tex=zeros(n_sim,time_step*T_ex);
B_Tset=zeros(n_sim,time_step*T_ex);
C_Tex=zeros(n_sim,time_step*T_ex);
C_Tset=zeros(n_sim,time_step*T_ex);

    for j=1:n_sim
        for i=1:time_step
            lambda(j,i)=path_R(j,i)*(((4*k*exp(-
                k*delta_t)))/((1-exp(-k*delta_t))*sigma^2));

            path_R(j,i+1)=((Z(j,i)+sqrt(lambda(j,i)))^2+x(j,i))*c;
        end
    end

for j=1:n_sim
    for i=1:time_step*T_ex
        n(i)=(T_set-(i-1)/time_step);
        m(i)=(T_ex-(i-1)/time_step);

        B_Tex(j,i)= (2*(exp(gamma*m(i))-
            1))/((k+gamma)*((exp(gamma*m(i))-1))+2*gamma);

        C_Tex(j,i)=((2*gamma*exp((k+gamma)*m(i)/2))/((k+gamma)
            )*((exp(gamma*m(i))-
            1))+2*gamma))^((2*k*mu)/(sigma^2));

        B_Tset(j,i)= (2*(exp(gamma*n(i))-
            1))/((k+gamma)*((exp(gamma*n(i))-1))+2*gamma);

        C_Tset(j,i)=((2*gamma*exp((k+gamma)*n(i)/2))/((k+gamma)
            )*((exp(gamma*n(i))-
            1))+2*gamma))^((2*k*mu)/(sigma^2));

        forward_rate(j,i)=(1/(T_set-
            T_ex))*((C_Tex(j,i)/C_Tset(j,i))*ex
            p(path_R(j,i)*(B_Tset(j,i)-
            B_Tex(j,i)))-1);

%
%continuously compounded forward interest rate

%
%          zcb_ex(j,i)=C_Tex(j,i)*exp(-
path_R(j,i)*B_Tex(j,i));
%          zcb_set(j,i)=C_Tset(j,i)*exp(-
path_R(j,i)*B_Tset(j,i));

```

```

%           forward_rate(j,i)=(log(zcb_ex(j,i))-
log(zcb_set(j,i)))/(T_set-T_ex);
    end
end
figure(1);
plot (forward_rate')

end

```

**CIR interest rate using the Matlab function *ncx2rnd* to sampling directly from a non-central chi-squared distribution  $\chi_d^2(\lambda)$  (as made in [36])**

```

path_R = zeros(n_sim,time_step+1);
path_R(:,1) = r0;
B=zeros(n_sim,time_step);
A=4*k+mu/(sigma^2);
C=((sigma^2)*(1-exp(-k*delta_t)))/(4*k)

    for j=1:n_sim
        for i=1:time_step
            B(j,i)=path_R(j,i)*((4*k*exp(-k*delta_t)))/(((1-
exp(-k*delta_t))*sigma^2));
            path_R(j,i+1)=C*ncx2rnd(A,B(j,i));
        end
    end
end

```



## Second chapter

### THE SABR MODEL

#### INTRODUCTION

One of the main problem that trader face is the managing of the volatility smile and skew in the option market. It is well known that there is a one-to-one relation between the option price and the volatility of the underlying: the call and the put price are indeed increasing function of the volatility. Due to this, in many markets the price quotes are often expressed directly by the level of the volatility (as in cap/floor market or swaption market). It is also well known that the value of the implied volatility given by the option price vary across the different strikes producing the volatility smile. The models which assume constant the volatility, as Black's model[57], are clearly unable to manage the volatility risk, and then do not allow to make a good vega-hedging. It is sufficient to think at the case in which we hold two different option on the same asset with different strikes  $K_1$  and  $K_2$ , and then different implied volatilities  $\sigma_1$  and  $\sigma_2$ . If we would make vega-hedging on this portfolio using a constant volatility model we would not know which volatility we should put into the model. Another problem of using a constant volatility model is the pricing of the exotic option: if we would price a call option with strike  $K_1$  which has a barrier at a certain level  $K_2$ , what implied volatility we should put into the pricing formula? A third problem evidenced in Hagan et. al.[23] concerns the evolution of the implied volatility curve. Since the implied volatility depends on the strike  $K$ , it is likely to also depend on the current value  $S$  of the asset price: In this case there would be systematic changes in volatility as the asset price  $S$  of the underlying changes. Some of the vega risks of Black's model would actually be due to changes in the price of the underlying asset, and should be hedged more properly as delta risks.

To solve this problem are been developed the *local volatility model* [18] where the volatility is not a constant parameter, but it depend on the level of the underlying and on the time. However, this models are not consistent with the behavior of the market since, as argued by Wilmott[54]: "the dynamic behavior of smiles and skews predicted by local volatility models is exactly opposite the behavior observed in the marketplace: when the price of the underlying asset decreases, local volatility models predict that the smile shifts to higher prices; when the price increases, these models predict that the smile shifts to lower prices. In reality, asset prices and market smiles move in the same direction". To solve this inconsistency, Hagan et. al. developed

the SABR (stochastic  $\alpha, \beta, \rho$ ) model, which is a two factor model where the dynamic of the underlying (in our case a forward interest rate  $f(t, S, S + \tau)$ , which we indicate with  $F_t^S$ ) is given by a CEV process with a lognormal stochastic volatility :

$$\begin{aligned} dF_t^S &= \sigma_t^S (F_t^S)^{\beta^S} dW_{1,t}^S \\ d\sigma_t^S &= \alpha^S \sigma_t^S dW_{2,t}^S \end{aligned} \quad 2.1$$

Where  $W_{1,t}^S$  and  $W_{2,t}^S$  are two correlated Brownian motion with  $E[dW_{1,t}^S dW_{2,t}^S] = \rho^S dt$ .

The parameter  $\sigma_t^S$  is the volatility of the forward rate,  $\alpha^S$  is a positive constant representing the volatility of the volatility, and  $\beta^S$  is a constant parameter representing the elasticity of variance. All these parameters are specific to a particular forward rate, of expiry  $S$  and maturity  $S + \tau$ . In the following sections, we will omit the subscript  $S$  to make clearer the exposition.

### *Plain-vanilla European option*

Consider a European call option on the asset  $A$  with exercise date  $t_{ex}$ , settlement date  $t_{set}$  and strike  $K$ . If the holder exercise the option on  $t_{ex}$ , then on the settlement date  $t_{set}$  he receive the underlying asset  $A$  and pays the strike  $K$ . The exercise of this kind of option is equivalent to receive in  $t_{ex}$  a forward contract  $F(t)$  on the asset  $A$  with maturity date  $t_{set}$ . Clearly, the option will be exercised only if the value of the forward contract is positive, and the value of the option at time zero is:

$$V_{call}(0) = P(0, t_{set}) \{E_0[F(t_{ex}) - K]^+\}$$

Where  $P(0, t_{set})$  is the value in  $t=0$  of zero coupon with maturity  $t_{set}$ . Using singular perturbation techniques, Hagan et. al obtained the option price implied by the SABR model, and from this, the associated implied volatility:

$$V_{call}(0) = P(0, t_{set}) \{fN(d_1) - KN(d_2)\} \quad 2.2$$

With  $f = F(0)$ , and

$$d_{1,2} = \frac{\log \frac{f}{K} \pm \frac{1}{2} \sigma_B^2 t_{ex}}{\sigma_B \sqrt{t_{ex}}}$$

This expression is the same of that given by the Black's formula for European option, with the difference that here the Black volatility  $\sigma_B$  is not constant, but given by:

$$\sigma_B(K, f) = \frac{\sigma_0^S}{(fK)^{(1-\beta)/2} \left( 1 + \frac{(1-\beta)^2}{24} \log^2(f/K) + \frac{(1-\beta)^4}{1920} \log^4(f/K) + \dots \right)} \cdot \left( \frac{z}{x(z)} \right) \cdot \left( 1 + \left( \frac{(1-\beta)^2}{24} \frac{(\sigma_0^T)^2}{(fK)^{(1-\beta)}} + \frac{1}{4} \frac{\rho\beta\alpha\sigma_0^T}{(fK)^{(1-\beta)/2}} + \frac{2-3\rho^2}{24} \alpha^2 \right) t_{ex} \right)$$

Where

$$z = \frac{\alpha}{\sigma_0^T} (fK)^{(1-\beta)/2} \log(f/K)$$

And

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

Implementing SABR model for vanilla option is very easy thanks to the fact that the price can be expressed by the Black's formula, which is the standard in the interest rate option market.

## 2.1 Sabr simulation introduction

When we want to simulate a path using this model, the real problem is not the simulation of the volatility process, which can easily be solved writing:

$$\sigma_t = \sigma_0 \exp \left( \alpha W_{2,t} - \frac{1}{2} \alpha^2 t \right) \quad 2.3$$

The real drawback is how to simulate the asset price given this behavior of the volatility. Many authors have proposed methods to simulate the SABR model; Islah [32] proposes a scheme based on relation of SABR to a squared Bessel process, Andersen et. al [5] propose a log-euler approximation scheme, Chen et al. [14] propose a scheme which mix moment-matched approximation and direct inversion (for small value of F). Following Chen et. al [14], we will implement a simulation algorithm by exploiting the relation between the SABR model and the squared Bessel process shown by Islah [32]. This relation will be shown in the subsequent section and is obtained by applying an invertible transformation to the asset price process, i.e  $X(F) = F^{1-\beta} / (1-\beta)$ . This transformed process, conditional on  $\sigma_\Delta$  and  $\int_0^\Delta \sigma_s^2 ds$ , results in a time-changed Bessel process of which we know the dimension and the starting point. Therefore, we

can sample from the Bessel process and applying an inversion scheme we can find our simulated values in the original coordinate.

In order to simplify the presentation, we first describe the relation between the CEV process and the Bessel process and then we will extend this relation to the SABR model.

### 2.1.1 Cev process

Let  $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$  be a filtered probability space generated by one-dimensional Brownian motion  $\{W_t\}$ . For all  $0 \leq t \leq T$ , the CEV process is the solution of the following SDE:

$$dS_t = \sigma_t S_t^\beta dW_t \quad 2.4$$

With  $0 < \beta \leq 1$ . We implement a series of transformation in order to connect the CEV process with a squared Bessel (process of which we know the transition density). Then, starting from the transition density of the squared Bessel process, using an inverse transformation, we can obtain the transition density and the cumulative distribution function of the CEV process.

After letting  $X_t = S_t^{1-\beta} / (1-\beta)$  for  $\beta \neq 1$ , we obtain by Ito lemma the following SDE for  $X_t$  [49]:

$$\begin{aligned} dX_t &= (1-\beta) \frac{S_t^{1-\beta}}{1-\beta} \sigma S_t^\beta dW_{1,t} - \frac{1}{2} \beta(1-\beta) \frac{S_t^{-1-\beta}}{1-\beta} \sigma^2 S_t^{2\beta} dt \\ &= \sigma dW_t - \frac{\beta \sigma^2}{(2-2\beta)X_t} dt \end{aligned} \quad 2.5$$

Now, applying a second transformation,  $Y_t = X_t^2$ , we get a time-changed squared Bessel process, which satisfies the following SDE:

$$dY_t = 2\sqrt{Y_t} \sigma dW_t + \delta \sigma^2 dt \quad 2.6$$

Where  $\delta := (1-2\beta)/(1-\beta)$  is the dimension of the process.

Let  $\nu(t) = \sigma^2 t$ , then  $Y_t = Z_{\nu(t)}$ , where  $\{Z\}$  is a squared Bessel process solution of the SDE[60]:

$$dZ_t = 2\sqrt{Z_t} dW_t + \delta dt \quad 2.7$$

With degree of freedom  $\delta$ .

It can be shown [4] that hold true the following results:



1 – For  $0 < \delta < 2$  we get  $\delta < 1/2$  and the SDE 2.4 has a unique solution only if a boundary condition is specified at  $S=0$ . When we specify an absorbing boundary condition, the density function does not integrate to 1 and the shortcoming represent the absorption probability. The absorbing boundary at  $S=0$  is required to ensure the arbitrage free condition. As argued in [49], for  $\delta > 0$  the price cannot became negative. Then, when price process reaches zero, if there were a reflecting boundary, we could buy this asset for zero and sell it when the barrier reflects the price process toward positive area, realizing a risk-free profit.

2 – For  $1 < \delta < 2$  we get  $1/2 < \delta < 1$ . A unique strong solution to SDE 2.4 exists, and boundary value zero is absorbing. The density function does not integrate to 1 and the shortcoming represent the absorption probability.

These two results cover the entire range of value that we will assign to parameter  $\delta$  in the path simulation. Indeed, for most financial application, parameter  $\delta$  ranges between 0 and 1.

Moreover, another fundamental result is that for  $\delta > 0$  and for  $0 < \delta < 2$ , but only when the boundary at zero is absorbing, the transition density  $q(t, x, y)$  for the squared Bessel process in eq. 2.7 is (proof in [9]):

$$q^\delta(t, x, y) = \frac{1}{2t} \left(\frac{y}{x}\right)^{\frac{\delta-2}{4}} \exp\left(-\frac{x+y}{2t}\right) I_{|\frac{\delta-2}{2}|} \left(\frac{\sqrt{xy}}{t}\right) \quad 2.8$$

With  $y \geq 0, t > 0$ .

Where  $I_a(x)$  denotes the Bessel function, defined by:

$$I_a(x) = \sum_{j=0}^{\infty} \frac{(x/2)^{2j+a}}{j! \Gamma(a+j+1)}$$

with  $\Gamma(x) = \int_0^{\infty} u^{x-1} e^{-u} du$  (Gamma function).

Based on 2.8, we can obtain the transition density for the CEV process. Inverting the transformation used above we get:

$$S_t = \left( (1 - \beta) \sqrt{|Z_{v(t)}|} \right)^{\frac{1}{1-\beta}}.$$

We can define a map

$$h: s \rightarrow \left( (1 - \beta) \sqrt{s} \right)^{\frac{1}{1-\beta}}, \quad s \geq 0,$$

with inverse

$$h^{-1}: y \rightarrow \frac{y^{2(1-\beta)}}{(1-\beta)^2}, \quad y \geq 0$$

So,  $S_t = h(Z_{v(t)})$  and  $Z_t = h^{-1}(S_0) = S_0^{2(1-\beta)} / (1-\beta)^2$ . Then  $Z_{v(t)}$  has density  $q(v(t), Z_0, y)$ , from which follows that the conditional transition density for the CEV process is given by:

$$p(S_t|S_0) = q^\delta(v(t), Z_0, h^{-1}(s)) \frac{dh^{-1}(s)}{ds}$$

By combining this density with the transition density of the squared Bessel process as in 2.8 we get:

$$p(S_t|S_0) = \frac{1}{v(t)} \left(\frac{S_t}{S_0}\right)^{\frac{1}{2}} \exp\left(-\frac{S_t^{2(1-\beta)} + S_0^{2(1-\beta)}}{2(1-\beta)^2 v(t)}\right) I_{\left|\frac{\delta-2}{2}\right|} \left(\frac{(S_0 S_t)^{1-\beta}}{v(t)(1-\beta)^2}\right) \frac{S_t^{1-2\beta}}{1-\beta} \quad 2.9$$

Where  $v(t) = \delta^2 t$  and  $\delta = (1 - 2\beta) / (1 - \beta)$ . By integrating 2.9 we find that the cumulative distribution function of the CEV process as in eq. 2.4 is given by the following formula:

$$\Pr(S_t \leq x|S_0) = 1 - \chi^2(a; b, c). \quad 2.10$$

With:

$$a = \frac{S_0^{2(1-\beta)}}{(1-\beta)^2 v(t)}, \quad b = \frac{1}{1-\beta}, \quad c = \frac{x^{2(1-\beta)}}{(1-\beta)^2 v(t)}$$

where  $\chi^2(a; b, c)$  is the non-central chi-square cumulative distribution function for the value “a”, non-centrality parameter “b” and degree of freedom “c”.

## 2.2 Absorption probability

As said in the previous section, the density does not integrate to unity when the boundary is absorbing. The difference between unity and the integral of the density function represent the absorption probability at 0. Then, using the equation 2.10, we can obtain the formula for the absorption probability:

For  $0 < \beta < 1$ , the probability of  $S_t$  given the initial condition  $S_0$  is [14]:

$$\Pr(S_t = 0 | S_0) = 1 - \lambda^2(a; b, 0)$$

$$= 1 - \lambda^2\left(\frac{S_0^{2(1-\beta)}}{(1-\beta)^2 v(t)}; \frac{1}{1-\beta}\right) \quad 2.11$$

In the next section we extend the results concerning the CEV model 2.4 in the context of the SABR model 2.1

### 2.3 SABR conditional distribution

We start by the simplest case to treat, that is the one where  $\beta = 1$ . We rewrite the eq.1 considering the following Cholesky decomposition for the two Brownian motion:

$$dW_{1,t} = \rho dW_{2,t} + \sqrt{1-\rho^2} dU_t \quad 2.12$$

Wher  $U_t$  is a standard Brownian motion independent of  $W_{2,t}$ . Then 2.1 begin:

$$\begin{aligned} dF_t &= \sigma_t F_t (\rho dW_{2,t} + \sqrt{1-\rho^2} dU_t) \\ d\sigma_t &= \alpha \sigma_t dW_{2,t} \end{aligned} \quad 2.13$$

Solving the SDE of F and integrating the SDE of  $\sigma$  we get:

$$F_t = F_0 \exp\left\{-\frac{1}{2} \int_0^t \sigma_s^2 ds + \rho \int_0^t \sigma_s dW_{2,s} + \sqrt{1-\rho^2} \int_0^t \sigma_s dU_s\right\} \quad 2.14$$

$$\int_0^t \sigma_s dW_{2,s} = \frac{1}{\alpha} (\sigma_t - \sigma_0) \quad 2.15$$

Substituting 2.15 into 2.14, we obtain

$$F_t = F_0 \exp\left\{-\frac{1}{2} \int_0^t \sigma_s^2 ds + \frac{\rho}{\alpha} (\sigma_t - \sigma_0) + \sqrt{1-\rho^2} \int_0^t \sigma_s dU_s\right\} \quad 2.15$$

Since  $U_t$  and  $W_{2,t}$  (and then  $U_t$  and  $\sigma_t$ ) are independent, by Ito isometry, the integral  $\int_0^t \sigma_s dU_s$  has a normal distribution with mean zero and variance  $\int_0^t \sigma_s^2 ds$ , and follows that the conditional distribution of  $\ln S_t$ , given  $\sigma_0, \sigma_t, \int_0^t \sigma_s^2 ds$  and  $S_0$  is log-normal and satisfies

$$\ln F_t \sim N \left( \ln F_0 - \frac{1}{2} \int_0^t \sigma_s^2 ds + \frac{\rho}{\alpha} (\sigma_t - \sigma_0) + (1 - \rho^2) \int_0^t \sigma_s^2 ds \right) \quad 2.16$$

Then we can easily sample F from this distribution.

For the other case, i.e  $\beta \neq 1$ , we proceed by applying to F the same transformation used for the CEV process, i.e  $X(F) = F^{1-\beta} / (1-\beta)$ . However, here we have to consider that the volatility process is stochastic, and the realization of  $F_t$  depends on the realization of  $\sigma_t$ .

We consider the transformation  $X(F)$ , after the Ito lemma application we get:

$$X_t = X_0 + \int_0^t \sigma_s dW_{1,s} - \int_0^t \frac{\beta \sigma_s^2}{(2-2\beta)X_s} ds \quad 2.17$$

and substituting the Cholesky decomposition 2.12 and the integral 2.15 into 2.17 we obtain:

$$X_t = X_0 + \frac{\rho}{\alpha} (\sigma_t - \sigma_0) + \sqrt{1-\rho^2} \int_0^t \sigma_s dU_s - \int_0^t \frac{\beta \sigma_s^2}{(2-2\beta)X_s} ds \quad 2.18$$

If we would condition on  $\sigma_t$ , we could introduce the shifted process  $\tilde{X}_t$ , with initial condition:

$$\tilde{X}_0 = X_0 + \frac{\rho}{\alpha} (\sigma_t - \sigma_0)$$

And by 2.18 we get.

$$d\tilde{X}_t = \sqrt{1-\rho^2} \sigma_t dU_t - \frac{\beta \sigma_t^2}{(2-2\beta)X_t} dt \quad 2.19$$

Now, we define the variable transformation  $Y_t = \tilde{X}_t^2$ , which, after applying Ito lemma, gives

$$\begin{aligned} dY_t &= 2\tilde{X}_t d\tilde{X}_t + d\tilde{X}_t^2 \\ &= 2\sqrt{Y_t} \sqrt{1-\rho^2} \sigma_t dU_t + \frac{1-2\beta-\rho^2(1-\beta)}{(1-\beta)(1-\rho^2)} (1-\rho^2) \sigma_t^2 dt \end{aligned} \quad 2.20$$

Applying Ito isometry can be easily shown that the integral of  $\sigma_t dU_t$  is a Normal distribution with mean zero and variance  $\int_0^t \sigma_s^2 ds$ . Then, when we apply the time change transformation

(which means a condition on integrated variance)  $v(t) = (1 - \rho^2) \int_0^t \sigma_s^2 ds$ , the process resulting is a Brownian motion distributed as  $\sqrt{1 - \rho^2} \int_0^t \sigma_s dU_s$  [56], i.e:

$$U_{v(t)} = \sqrt{1 - \rho^2} \int_0^t \sigma_s dU_s$$

Substituting this time changed Brownian motion into 2.20 we get

$$dY_{v(t)} = 2\sqrt{Y_t} dU_{v(t)} + \frac{1 - 2\beta - \rho^2(1 - \beta)}{(1 - \beta)(1 - \rho^2)} dv(t) \quad 2.21$$

From which follows that, starting from the process for F as in 2.1, given the transformation  $Y(F) = F^{2-2\beta} / (1-\rho)^2$ ,<sup>3</sup> and applying a time-change transformation  $v(t) = (1 - \rho^2) \int_0^t \sigma_s^2 ds$ , the resulting process  $Y_{v(t)}$ , conditional on the terminal volatility  $\sigma_t$  and integrated variance  $\int_0^t \sigma_s^2 ds$ , is a squared Bessel process solving the SDE:

$$Y_{v(t)} = 2\sqrt{Y_t} dU_{v(t)} + \delta dv(t) \quad 2.22$$

With initial condition  $Y_0 = \left( \frac{F_0^{1-\beta}}{1-\beta} + \frac{\rho}{\alpha} (\sigma_t - \sigma_0) \right)^2$  and dimension  $\delta = \frac{1-2\beta-\rho^2(1-\beta)}{(1-\beta)(1-\rho^2)}$

and, proceeding as we did in the previous section to get the the cumulative distribution function of the CEV process, we obtain that for some  $F_0 > 0$ , the conditional cumulative distribution of  $F_t$  with an absorbing boundary at  $F_t = 0$  given  $\sigma_t$  and  $\int_0^t \sigma_s^2 ds$  is:

$$\Pr\left(F_t \leq K \mid F_0 > 0, \sigma_t, \int_0^t \sigma_s^2 ds\right) = 1 - \chi^2(a, b, c) \quad 2.23$$

Where

$$a = \frac{1}{v(t)} \left( \frac{F_0^{1-\beta}}{1-\beta} + \frac{\rho}{\alpha} (\sigma_t - \sigma_0) \right)^2, \quad b = 2 - \frac{1 - 2\beta - \rho^2(1 - \beta)}{(1 - \beta)(1 - \rho^2)}, \quad c = \frac{K^{2(1-\beta)}}{(1 - \beta)^2 v(t)},$$

$$v(t) = (1 - \rho^2) \int_0^t \sigma_s^2 ds$$

<sup>3</sup> the second transformation  $Y = \bar{X}^2$  is not applied on  $X$ , but on  $\bar{X}$ , and we have  $\bar{X} = X$  only when  $\rho = 0$ ; nevertheless ishah 2009 show how, if  $t$  is small, the result is applicable with a very small approximation error even when  $\rho \neq 0$  )

and  $\chi^2(a, b, c)$  represent the chi – square cumulative distribution function.

To verify the last result is sufficient to substitute the parameters  $\delta$  and  $Y_0$  given in equation 2.22 into the distribution function presented in equation 2.10.

## 2.4 The integrated variance approximation

Several methods have been proposed in the literature to approximate the integrated variance. In the following subsection we describe the most used approximation techniques.

### 2.4.1 Chen et. al approximation

Chen et al. propose a moments matching approximation method. This method is based on the fitting of the first two conditional moments of  $\int_0^t \sigma_s^2 ds$  in a lognormal distribution. The conditional mean and variance of  $\int_0^t \sigma_s^2 ds$  are approximated using a small disturbance expansion. They reformulate the volatility process by introducing a small parameter  $\epsilon$  in the diffusion coefficient so that:

$$\sigma_t^{(\epsilon)} = \sigma_0 + \epsilon \tilde{\alpha} \int_0^t \sigma_s^{(\epsilon)} dW_{2,s} \quad 2.24$$

Where  $\tilde{\alpha} = \epsilon \tilde{\alpha}$ .

After the small disturbance expansion around  $\epsilon = 0$ , they get the following approximation for the conditional mean “m” and variance “v”:

$$m = \sigma_0^2 t \left( 1 + \alpha W_{2,t} + \frac{1}{3} \alpha^2 \left( 2W_{2,t}^2 - \frac{t}{2} \right) + \frac{1}{3} \alpha^3 (W_{2,t}^3 - W_{2,t}t) \right. \\ \left. + \frac{1}{5} \alpha^4 \left( \frac{2}{3} W_{2,t}^4 - \frac{3}{2} W_{2,t}^2 t + 2t^2 \right) \right) \quad 2.25$$

$$v = \frac{1}{3} \sigma_0^3 \alpha^2 t^3 \quad 2.26$$

In order to avoid that the integrated variance takes negative value (which would be meaningless), it is convenient to choose log-normal distribution to reproduce the conditional distribution of the integrated variance. Then, the corresponding mean  $\mu$  and variance  $\sigma^2$  of the lognormal distribution are obtained by setting:

$$\mu = \ln(m) - \frac{1}{2} \ln \left( 1 + \frac{v}{m^2} \right) \quad 2.27$$

$$\sigma^2 = \ln\left(1 + \frac{v}{m^2}\right). \quad 2.28$$

And the integrated variance is given by the approximation:

$$\int_0^t \sigma_s^2 ds \approx \exp(\sigma N^{-1}(U) + \mu) \quad 2.29$$

Where U is a uniform random variable.

#### 2.4.2 Kennedy's approximation

Another way based on the moment matching is to calculate analytically the moments of the integrated variance. As shown by Kennedy[35], the mean  $\check{\mu}$  and the variance  $\check{\sigma}^2$  of the integrated variance are given by the following analytical expressions:

$$\check{\mu} = \frac{\sigma_0^2 \sqrt{t}}{2\alpha} \frac{\left[ \Phi\left(\frac{\ln\left(\frac{\sigma_s}{\sigma_0}\right)}{\alpha\sqrt{t}} + \alpha\sqrt{t}\right) - \Phi\left(\frac{\ln\left(\frac{\sigma_s}{\sigma_0}\right)}{\alpha\sqrt{t}} - \alpha\sqrt{t}\right) \right]}{\varphi\left(\frac{\ln\left(\frac{\sigma_s}{\sigma_0}\right)}{\alpha\sqrt{t}} + \alpha\sqrt{t}\right)}$$

$$\check{\sigma}^2 = -\frac{\sigma_0^4 \sqrt{t}}{4\alpha^3} \left(1 + e^{2\ln\left(\frac{\sigma_s}{\sigma_0}\right)}\right) \frac{\left[ \Phi\left(\frac{\ln\left(\frac{\sigma_s}{\sigma_0}\right)}{\alpha\sqrt{t}} + \alpha\sqrt{t}\right) - \Phi\left(\frac{\ln\left(\frac{\sigma_s}{\sigma_0}\right)}{\alpha\sqrt{t}} - \alpha\sqrt{t}\right) \right]}{\varphi\left(\frac{\ln\left(\frac{\sigma_s}{\sigma_0}\right)}{\alpha\sqrt{t}} + \alpha\sqrt{t}\right)}$$

$$+ \frac{\sigma_0^4 \sqrt{t}}{4\alpha^3} \frac{\left[ \Phi\left(\frac{\ln\left(\frac{\sigma_s}{\sigma_0}\right)}{\alpha\sqrt{t}} + 2\alpha\sqrt{t}\right) - \Phi\left(\frac{\ln\left(\frac{\sigma_s}{\sigma_0}\right)}{\alpha\sqrt{t}} - 2\alpha\sqrt{t}\right) \right]}{\varphi\left(\frac{\ln\left(\frac{\sigma_s}{\sigma_0}\right)}{\alpha\sqrt{t}} + 2\alpha\sqrt{t}\right)}$$



### 2.4.3 Andersen's approximation

Andersen [6] proposes to approximate the integrated variance by applying a simple trapezoidal rule:

$$\int_0^t \sigma_s^2 ds \approx t(\gamma_1 \sigma_0^2 + \gamma_2 \sigma_t^2) \quad 2.30$$

Where  $\gamma_1$  and  $\gamma_2$  are two constant parameter which can be chosen in several ways. The simplest is given by choosing  $\gamma_1 = 1$ ,  $\gamma_2 = 0$  (which gives an Euler-like approximation). However, the method used by Andersen is based on the central discretization, given by setting  $\gamma_1 = \gamma_2 = \frac{1}{2}$ . An extension of this method could consist in the simulation of intermediate volatilities in the interval  $(0,t)$ . The idea is to split the interval  $(0,t)$  in  $n$  sub-steps, and associating to any intermediate volatility the same weight  $w = 1/(1+n)$  we get the following approximation

$$\int_0^t \sigma_s^2 ds \approx t \sum_{i=0}^n w \sigma_i^2 \quad 2.31$$

## 2.5 Simulation of the forward rate path

It is possible to distinguish two different methodology for SABR model simulation. The most used and studied is the moment matching approach, which consist in the fitting of the first two moment of process distribution with a distribution function chosen by the operator. The first example of the application of moment matching applied in the contest of a stochastic volatility model is the QE (*quadratic exponential*) scheme introduced by Andersen[6] to simulate an underlying in the Heston model. Haven[59] extend the QE scheme to the SABR model implementing an algorithm very similar to Andersen one, solving the problem given by the fact that the QE scheme is based on a squared Bessel process with reflecting boundary at zero. The other methodology use a scheme developed by Marakow and Glew [47], which is used for the exact simulation of the squared Bessel process. Other methods may be used as the log-Euler discretization, but these do not take into account the absorbing or reflecting condition at zero. Chen et. al propose an algorithm which mix the use of the moment matching techniques under certain condition ( $F \gg 0$  and then “absorption probability”  $\rightarrow 0$ ) and a direct inversion scheme for small value of  $F$ . this last scheme is that we will use to simulate the forward rate.



### 2.5.1 Chen simulation scheme

The simulation scheme proposed by Chen use the property of the SABR seen in the previous section. In particular, we have seen that the asset price dynamic of the asset in the SABR model is a CEV process, and this can be linked with a transformed squared Bessel process. It is argued that, when the level of the asset price is high, the probability of hitting zero is very small and could be approximate to zero, i.e  $\Pr(\inf\{t|S_t = 0\} < \infty) \approx 0$ . Then, by using the conditional cumulative distribution of  $F_t$  given in equation 2.23 and the properties of non-central chi-square distribution, it is possible to make the following approximation:

$$\Pr\left(F_t \leq K \mid F_0 > 0, \sigma_t, \int_0^t \sigma_s^2 ds\right) = 1 - \chi^2(a, b, c) = \chi^2(c, 2 - b, a) + \Pr(\inf\{t|F_t = 0\} < \infty) \approx \chi^2(c, 2 - b, a)$$

Where  $a, b, c$  are as in 2.23.

It is possible to show (see [37]) that the non-central chi-square distribution approaches to a power function applied to a Gaussian variable when the non-centrality parameter is high. The preferable transformation of the Gaussian function is the cubic one (used in [6]), but such scheme does not guarantee positive value of  $F$ . Then, when the level of  $F$  is sufficiently high, we can match the first two moment of the distribution  $\chi^2(c, 2 - b, a)$  to a quadratic Gaussian function (as made in the QE scheme).

For small value of  $F$ , Chen et. al propose to implement a Newton-type root finding method to invert directly the distribution given by equation 2.23. Specifically, first it is determined the value of  $c^*$  which solves the equation  $1 - \chi^2(a, b, c^*) - U = 0$  (where  $U$  is an extraction from an uniform random distribution), and then an inversion scheme is applied to sample  $F$  from the value of  $c^*$ . In the next subsection we analyze more specifically the moment matching approximation and the direct inversion scheme.

### 2.5.2 Moment matching approximation

The quadratic representation of the Gaussian process  $Y_t$  used in Andersen[6] (and also in [14]) takes the form:

$$Y_t = d(e + Z)^2 \quad 2.32$$

Where  $Z \sim N(0,1)$  and  $d$  and  $e$  are two constant to be determined by the moment matching so that the mean and the variance of the quadratic Gaussian process  $Y_t$  corresponds to those of the non-central chi-square distribution given in 2.23. These constant parameters depend on  $t$  and on the value of  $Y$  at zero.

The mean and the variance of a non-central chi-square distribution  $\chi^2(x, k, \lambda)$  are defined respectively by  $m := k + \lambda$  and  $s^2 = 2(k + 2\lambda)$ . We want that:

$$E[Y] = m$$

$$\text{Var}[Y] = s^2$$

And since:

$$E[Y] = d(1 + e^2)$$

$$\text{Var}[Y] = 2d^2(1 + 2e^2)$$

Follows that:

$$m = d(1 + e^2); \quad s^2 = 2d^2(1 + 2e^2) \quad 2.33$$

Substituting  $d = \frac{m}{(1+e^2)}$  into the second equation, we get:

$$e^2 = 2\varphi^{-1} - 1 + \sqrt{2\varphi^{-1} - 1} \sqrt{2\varphi^{-1} - 1} \quad 0$$

$$\text{where } \varphi = \frac{s^2}{m^2} \quad 2.34$$

Which can be solved only if  $\varphi \leq 2$ . Then for High value of  $\varphi$  (corresponding to low value of  $F$ ), the moment matching fail.

In our case, we have from equation 2.23:

$$\lambda = \frac{1}{v(t)} \left( \frac{F_0^{1-\beta}}{1-\beta} + \frac{\rho}{\alpha} (\sigma_t - \sigma_0) \right)^2$$

$$k = 2 - \frac{1-2\beta-\rho^2(1-\beta)}{(1-\beta)(1-\rho^2)}$$

From  $\lambda$  and  $k$  we can derive the value of  $d$  and  $e$ , hence we can sample  $F_t$  by approximate the process  $F_t^{2(1-\beta)} / (1-\beta)^2 v(t)$  to the quadratic normal process  $Y$ :

$$\frac{F_t^{2(1-\beta)}}{(1-\beta)^2 v(t)} = d(e + Z)^2$$

From which

$$F_t = ((1 - \beta)^2 v(t) d(e + Z)^2)^{\frac{1}{2(1-\beta)}} \quad 2.36$$

### 2.5.3 Direct inversion scheme

When the moment matching scheme fail, i.e when  $F$  is low or  $\varphi > 2$ , we have to estimate  $F$  using another method. Chen et. al propose to estimate  $F$  inverting the equation which gives the non-centrality parameter  $c$  in equation 2.23. The main problem is that when  $F$  is small the density function does not integrate to one, and then we have to compute  $c$  considering this. In particular, we can determine the value  $c^*$  which solves the equation  $H(a,b,c^*) := 1 - \int_0^{\infty} \lambda^2(a,b,c) - U = 0$  and find  $F$  by inverting the equation for  $c$  in equation 2.23, which result in:

$$F_t = (c^*(1 - \beta)^2 v(t))^{\frac{1}{2-2\beta}} \quad 2.37$$

We can use several method to determine  $c^*$ , such as interpolation or a Newton's method. In Chen et. al is used the last one, in their algorithm they get  $c_0 = a$  and control if this value solves the equation  $H(a,b,c^*) := 1 - \int_0^{\infty} \lambda^2(a,b,c^*) - U = 0$  with a certain degree of tolerance  $\epsilon$ . If the tolerance is not achieved, i.e  $H(a,b,c^*) := 1 - \int_0^{\infty} \lambda^2(a,b,c^*) - U > \epsilon$ , the Newton's method is applied until the prescribed tolerance is reached:

$$c_{n+1} = c_n - \frac{H(a,b,c_n)}{q(a,b,c_n)} \quad 2.38$$

Where  $q(a,b,c_n)$  is the transition density for the squared Bessel process with an absorbing boundary at zero given by equation 2.8. The implementation of this method require some attention when is implemented with a programming language because  $c^*$  can became negative and this would be meaningless.

**SABR PATH SIMULATION USING CHEN ET. AL SCHEME[36]**

```

F0=0.05252;
T=1;
sigma_0=0.03;    %ANNUAL
alfa=0.60;      %ANNUAL
beta=0.5;
rho=-0.25;
time_step=250;
n_sim=500;

tollerance=0.001;
maxiter=200;

Delta = 1/time_step;
NSteps = T/Delta;
pathF = zeros(n_sim,NSteps+1);
pathF(:,1) = F0;

F_Delta = repmat(F0,n_sim,1) ;

for i = 1:NSteps

    Z = sqrt(Delta)*randn(n_sim,1);
    sigma_Delta = sigma_0.*exp(alfa*(Z-0.5*alfa*Delta));

    m = sigma_0.^2*Delta.*(1 + alfa*(Z + alfa*((2*Z.^2-
        0.5*Delta)/3 + alfa*((Z.^3-Z*Delta)/3 +
        alfa*(2/3*Z.^4-1.5*Z.^2*Delta+2*Delta^2)/5))));
    v = sigma_0.^4*alfa^2*Delta^3/3;

    % moment-matched log-normal distribution
    sigma2 = log(1+v./m.^2);
    mu = log(m) - 0.5*sigma2;

    %inverse transformation
    U = randn(n_sim,1);
    A_Delta = exp(sqrt(sigma2).*U + mu);

    v_Delta = (1-rho^2)*A_Delta;

    if beta == 1                % case: beta = 1
        dW = rho*Z + sqrt(1-
            rho^2)*sqrt(Delta)*randn(n_sim,1); % cholesky

        F_Delta = F_Delta.*(1 + sigma_0.*dW);
    else
        % Direct Inversion Scheme for Conditional CEV
        Process
        if alfa >0

```

```

        a = (F_Delta.^(1-beta)/(1-beta) +
            rho*(sigma_Delta-
                sigma_0)/alfa).^2./v_Delta;
    else
        a = (F_Delta.^(1-beta)/(1-beta)).^2./v_Delta;
    end
    b = 2 - (1-2*beta-rho^2*(1-beta))/(1-beta)/(1-
        rho^2);

    U1 = rand(n_sim,1);

    P_absorb = 1 - chi2cdf(a,b);

    % distinguish between the following cases

    % (1)
    I1 = F_Delta == 0;
    F_Delta(I1) = 0;
    % (2)
    I2 = F_Delta ~= 0 & U1 <= P_absorb;
    F_Delta(I2) = 0;
    % (3)
    I3 = F_Delta ~= 0 & U1 > P_absorb;

    if sum(I3) > 0

        % parameters for the moment-matched quadratic
        % gaussian approximation,
        k = 2-b;
        lambda = a;
        m = k+lambda;
        s2 = 2*(k+2*lambda);
        Psi = s2./m.^2;

        Psi_thres = 2;

        % moment matched quadratic gaussian
        approximation
        I4 = ((Psi > 0 & Psi <= Psi_thres) & m >= 0) &
        I3 == 1;
        if sum(I4) > 0
            e2 = 2./Psi-1+sqrt(2./Psi).*sqrt(2./Psi-
            1);
            d = m./(1+e2);

            F_Delta(I4) = ((1-
            beta)^2*v_Delta(I4).*d(I4).*(sqrt(e2(I4))+randn(sum(I4),1)).^2
            ).^(1/(2*(1-beta))));
            end

        % Step 7 otherwise

```

```

I5 = ((Psi > Psi_thres | ( m < 0 & Psi > 0 &
    Psi <= Psi_thres)) & I3 == 1);

if sum(I5) > 0
    % Newton method to determine the root c*
    % of the
    % function : H(a,b,c) = 1-Chi^2(a,b,c)-U =
    %              0
    % c* is a random number from a squared
    % Bessel
    % distribution with absorbing boundary at
    % zero

H = @(c,IVec)(1-ncx2cdf(a(IVec),b,c)-
    U1(IVec));
qbar = @(c,IVec)
    (0.5*(c./a(IVec)).^(0.25*(b-2)) .*
    exp(-0.5*(a(IVec)+c)) .*
    besseli(abs(0.5*(b-
    2)),sqrt(a(IVec).*c)));

nVec = (1:length(a))';
indexVec = nVec(I5);

IndOfInterest=H(0,indexVec)<0;
iVec=indexVec(IndOfInterest);

c_star=a(iVec);
c_old=c_star; %initial counter
k=0;
i6=abs(H(c_star,iVec))>tollerance;
while k < maxiter && sum(i6) > 0
    k=k+1;
    step=1;
    c_star(i6)=c_old(i6)-
        step*H(c_old(i6),iVec(i6))./
        qbar(c_old(i6),iVec(i6));

while sum(c_star<0)>0 && step>tollerance
    i7 = c_star < 0;
    step=0.5*step;
    c_star(i7)=c_old(i7)-
        step*H(c_old(i7),iVec(i7))./qbar(c_o
        ld(i7),iVec(i7));
end
c_old(i6)=c_star(i6);
i6=abs(H(c_star,iVec))>tollerance;
end

F_Delta(iVec)=(c_star.*(1-
beta)^2.*v_Delta(iVec)).^(1/(2-2*beta));

```

```
                                F_Delta(indexVec(~IndOfInterest)) = 0;
                                end
                            end
                        end

                        % update initial volatility
                        sigma_0 = sigma_Delta;
                        pathF(:,i+1) = F_Delta;
                        figure(3);
                        plot (pathF')
                        SABRsimulation= pathF;
                    end
                end
            end
        end
    end
end
```





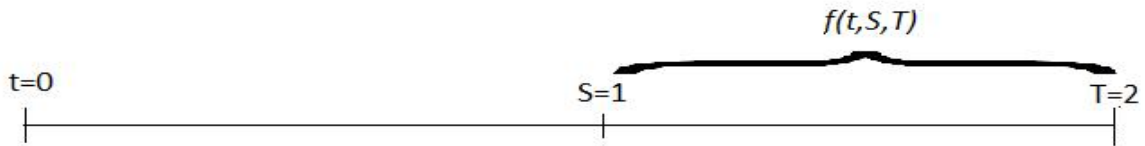
## **Third Chapter**

# **Estimation of the ATSM parameters and pricing of Caplet option**

### INTRODUCTION

The main purpose of this work is to try to answer questions like: are the forward rate paths generated by a SABR model compatible with the interest rate paths generated by an affine term structure model? Or, how diverge the option prices given by these models? In general, when traders want to make a comparison between different models in order to find the best one to use, they choose to do this starting from real data. Therefore, they estimate the parameters of the models from the historical time series observed in the market, hence they make the simulations of the paths and the pricing of the option using the estimated models, and finally they compare the resulting paths and option prices with ones observed in the market. This procedure is the market practice, and it is the best of what the operators can do. However, when we work with real data we do not know the true data generating process, the only information we have is the (unique) realization of this process. Then, the goodness of a model is measured with respect to a very limited set of information about the true model. In this work, we do not rely on the correspondence of each model with real data, but we compare each model directly with the others. In particular, we assume that our true data generating process is the SABR model and, on the basis of the data generated by this, we estimate the parameter of the Vasicek and CIR model. After the estimation, we make the simulation path and the pricing of some options and we compare them with those given by the SABR model. If the difference of the options price is low and the paths generated are similar, we can conclude that using the SABR model or the Vasicek and CIR models does not involve large differences. Then, we could use the simplest ATS models. Otherwise, we can conclude that the SABR model own some features that other models are not able to reproduce and then we should use this. However, we must emphasize that we are making only an econometric analysis, which does not include any consideration about the already established superiority of the SABR model in the management of the hedging strategy and of the volatility smile in the options market.

The first step of the analysis is to assign a value to the SABR model parameters. We will assume that these parameters refers to the dynamic of a specific forward rate  $f(t, S, T)$ , i.e the forward rate of expiry  $S = 1$  year and maturity  $T=2$  years.



In order to choose reasonable parameters, we have taken these from a SABR model estimated on the swaption quotes<sup>4</sup>, getting the following values:

$f(0, S, T)$	0.05252
$\alpha$	0.5
$\frac{\sigma_0}{\rho}$	0.029
$\rho$	- 0.25
$\frac{\sigma_0}{\rho}$	0.61

Tab 3.1 SABR parameters

Then, we put this set of parameter values into the SABR model and, using the method described in chapter 2, simulate  $n (=500)$  paths of the forward rate for 250 time steps (TS) and with  $t$  equal to  $1/250$ <sup>5</sup>. Now, we have several paths generated by model we consider as “true” model and we want to see if the Vasicek and CIR model are able to replicate the behavior of these paths. To do this, we first need to estimate the set of parameters  $\tilde{\theta} = (\tilde{\alpha}, \tilde{\mu}, \tilde{\sigma})$  for both affine term structure models. In order to estimate the set of parameters  $\tilde{\theta}$  we need the discretization version of stochastic differential equation of the model considered. From equations 1.24 and 1.44 of chapter 1 we have the following discrete equations:

- Discrete version of the Vasicek model:

$$r(t + \Delta t) = r(t)e^{-a\Delta t} + \mu(1 - e^{-a\Delta t}) + \sigma\sqrt{(1 - e^{-a\Delta t})/2a}\epsilon_t \quad 3.1$$

<sup>4</sup> These data are taken from R.Rebonato et. al , The SABR/ Libor Market model pag. 29

<sup>5</sup> This setup simulates a daily forward rate path

where  $\epsilon_t \sim i.i.d N(0,1)$ .

- Discrete version of the CIR model:

$$r(t + \Delta t) = r(t) + a(\mu - r(t))\Delta t + \sigma\sqrt{r(t)}Z_t, \quad 3.2$$

where  $Z_t \sim i.i.d N(0, \Delta t)$ .

If the data generated by the SABR model were interest rates, we could directly estimate such models on these data. However, since it is not so, we first have to implement a calibration procedure which “convert”, as well as possible, the SABR forward rate in short term interest rater<sup>\*</sup>( $t$ ).

### 3.1 Calibration procedure

The In order to find the set of parameters  $\theta^* = (\alpha^*, \mu^*, \sigma^*)$  we use the one-to-one relation between the forward rate  $f^*(t, S, T)$  and  $r^*(t)$  given in equation 1.55 for the CIR model and in equation 1.49 for the the Vasicek model<sup>6</sup>,mathematically we have:

$$f^*(t, S, T) = r^*(t)g(\theta^*, t, S, T) \quad 3.3$$

Where  $g$  is a deterministic function of  $\theta^*, t, S$  and  $T$ .

The calibration procedure we will implement consist of the following steps:

1 – Assign a range of values to the parameters to be calibrated; i.e

$$0 < \alpha^* < 0.3$$

$$0 < \mu^* < 0.1$$

$$0 < \sigma^* < 0.1$$

And, within this range, we select each parameter by varying its value using a step equal to 0.001, so that we get 0.1/0.001 values of  $\mu^*$  and  $\sigma^*$  and 0.3/0.001 values of  $\alpha^*$

2 – Put any possible combination of  $\alpha^*, \mu^*$  and  $\sigma^*$  <sup>7</sup>into 3.1 and 3.2. For each combination, generate  $m(=100)$  paths of  $r^*(t)$  starting from  $r^*(0)$ , which is calculated by 3.2.1 putting  $f(0, S, T)$  in the left side . Then, use 3.1 to generate the forward rates  $f^*(t, S, T)$ .

<sup>6</sup> Since analogous considerations can be made for the CIR model, to describe the calibration procedure we refer only to the Vasicek model.

<sup>7</sup> In our simulation we have  $100 \times 100 \times 300 = 3.000.000$  of combinations

3 – To decide what combination of parameters is the best, compare each  $f^*(t, S, T)$  with each  $f(t, S, T)$  and choose the combination of parameters that minimize the mean of *root mean square deviation*:

$$\sqrt{\frac{1}{TS} \sum_{t=1}^{TS} (f(t, S, T) - f^*(t, S, T))^2} \quad 3.4$$

Where TS is the number of observation for each path (in our case, TS=250).

The Matlab implementation of this procedure give us the following result:

Parameter	Value	Mean of RMSerrors <sup>8</sup>	CV(RMSE) <sup>9</sup>
$\bar{a}^*$	0.220	0.0042	8.1 %
$\bar{\mu}^*$	0.041		
$\bar{\sigma}^*$	0.038		

Tab. 3.2 Vasicek calibrated parameters

Parameter	Value	Mean of RMSerrors	CV(RMSE)
$\bar{a}^{ctr}$	0.262	0.0033	6.36 %
$\bar{\mu}^{ctr}$	0.035		
$\bar{\sigma}^{ctr}$	0.029		

Tab 3.3 CIR calibrated parameters

The value of CV(RMSE) show that the calibrated CIR model generate paths that, on average, are closer (to SABR paths) then ones generated by the Vasicek model. A view of the figures 3.4 – 3.6 show that the “density” of 5000 forward rate at time t=1 generated by Vasicek and CIR model is not so far from the SABR ones.

<sup>8</sup> Of forward rate given using estimated model

<sup>9</sup> Coefficient of variation of the RMSE calculated as the ratio  $\text{RMSerror} / \sqrt{\text{mean}(f(t, S, T))}$ , where the denominator represents the mean of the forward value  $f(t, S, T)$

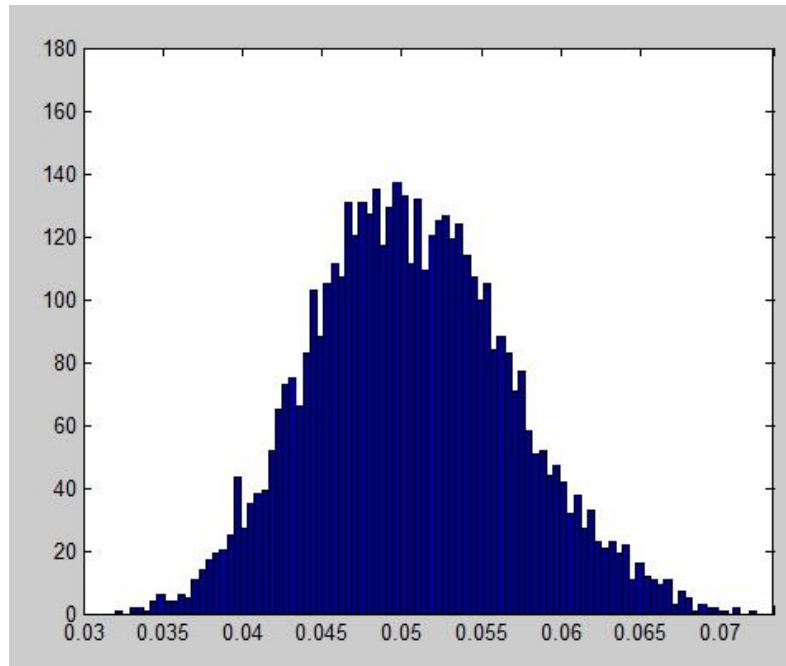


Fig.3.4 forward rates density generated by the SABR model with parameters as in tab 3.1

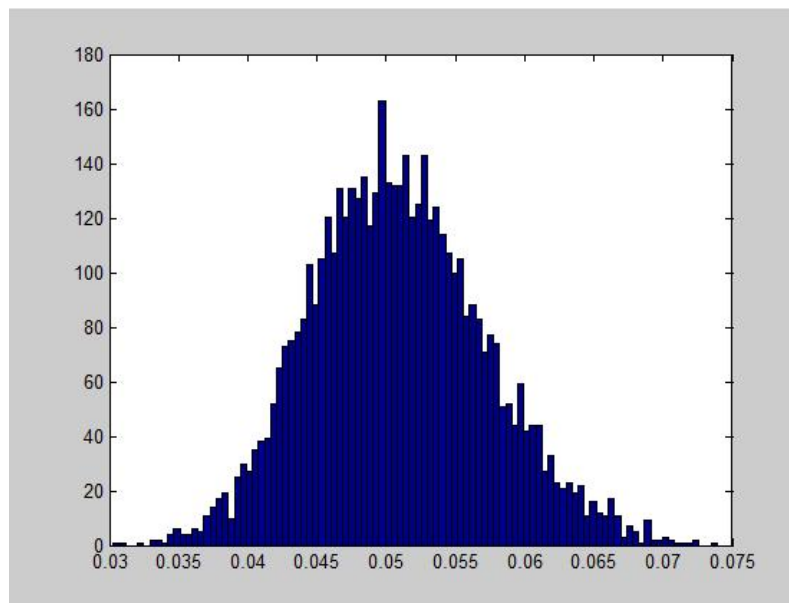


Fig. 3.5 forward rate at time  $S=1$  generated by calibrated CIR

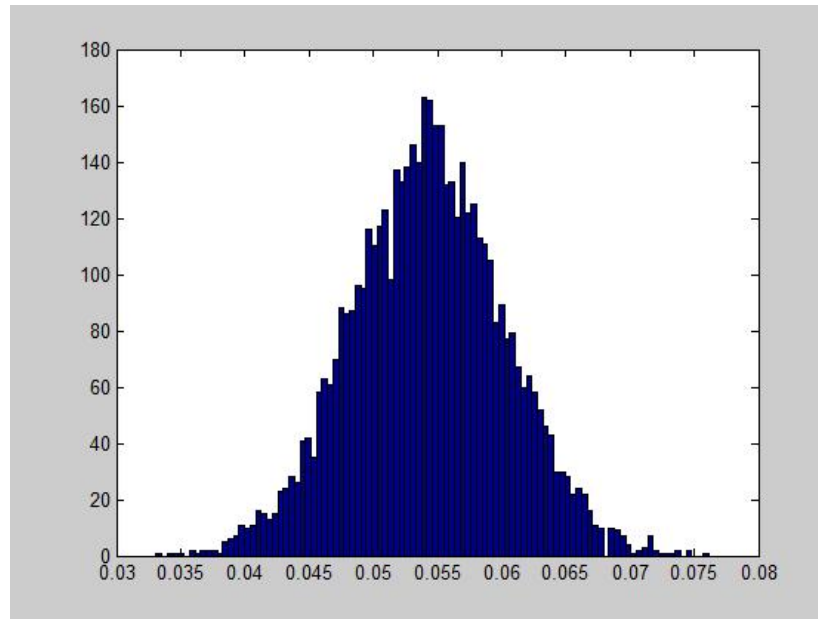


Fig. 3.6 histogram of forward rate at time S=1 generated by calibrated Vasicek

The value of the CV(RMSE) for the CIR model indicates that, on average, the forward rates  $f^*(t, S, T)$  generated by this model diverge by 6.36% (of the value  $f(t, S, T)$ ) respect to the same  $f(t, S, T)$ . Then, looking at the value of CV(RMSE), the CIR model seems to fit better than the Vasicek model the paths generated by SABR model.

Now, we consider the interest rates  $r^*(t)$  generated by the calibrated model as the “true” interest rates resulting from the SABR model and apply the estimation procedure.

### 3.2 Parameter Estimation

In the literature, we can find several estimation methods for both the Vasicek and CIR model. In the following paragraph we will describe and implement two of the most used: the OLS estimation and the Maximum likelihood estimation.

#### 3.2.1 OLS estimation

In order to apply the OLS to the studied model we have to consider the discrete version given in equations 3.1 and 3.2. We assume that the relationship between two consecutive observations  $r(t)$  and  $r(t + 1)$ , with  $0 \leq t < T$ , is given by the following regression with an i.i.d random term  $\omega_t$ :

$$r(t + 1) = \varphi + br(t) + \omega_t \quad 3.5$$

The OLS estimation consist of minimizing the variance of residual, or equivalently the *sum of the squared residual (SSR)*, which results in:

$$SSR = \sum_{t=0}^{T-1} \omega_t^2 = \sum_{t=0}^{T-1} (r(t + 1) - (\varphi + br(t)))^2 \quad 3.6$$

Where T represent the number of observations.

Hence, the OLS estimators  $\hat{\varphi}$  and  $\hat{b}$  can be found by solving the following optimization problem:

$$\min \sum_{t=0}^{T-1} (r(t + 1) - (\varphi + br(t)))^2 \quad 3.7$$

By differentiating 3.7 with respect to  $a$  and with respect to  $b$  and setting these derivatives equal to 0, we find the estimators  $\hat{\varphi}$  and  $\hat{b}$ , which are equal to[39]:

$$\begin{aligned} \hat{b} &= \frac{Tr_{xy} - r_x r_y}{Tr_{xx} - r_x^2} \\ \hat{\varphi} &= \frac{r_y - \hat{b}r_x}{T} \\ st. deviation(\omega) &= \sqrt{\frac{Tr_{yy} - r_y^2 - \hat{b}(Tr_{xy} - r_x r_y)}{T(T - 2)}} \end{aligned}$$

Where

$$r_x = \sum_{t=0}^{T-1} r(t), \quad r_y = \sum_{t=0}^{T-1} r(t + 1), \quad r_{xx} = \sum_{t=0}^{T-1} r(t)^2,$$

$$r_{xy} = \sum_{t=0}^{T-1} r(t)r(t + 1), \quad r_{yy} = \sum_{t=0}^{T-1} r(t + 1)^2,$$

In Matlab it is possible to find  $\hat{\varphi}$  and  $\hat{b}$  by using the function *fminsearch*.

### 3.2.1.1 - Vasicek OLS parameter estimation

The discrete Vasicek equation 3.1 can be expressed by the linear relation 3.5 setting the parameters as follow:

$$b = e^{-a\Delta t}$$

$$\varphi = \mu(1 - e^{-a\Delta t})$$

$$st. deviation(\omega) = \sigma\sqrt{(1 - e^{-a\Delta t})/2a}$$

Rewriting these equation gives:

$$a = -\frac{\ln b}{\Delta t} \quad 3.8$$

$$\mu = \frac{\varphi}{1 - b} \quad 3.9$$

$$\sigma = st. deviation(\omega) \sqrt{\frac{-2 \ln b}{\Delta t(1 - b^2)}} \quad 3.10$$

Then, to estimate the parameter of the Vasicek model it is sufficient to find  $\hat{\varphi}$  and  $\hat{b}$  solving the equation 3.7 and apply the equation 3.8-3.9. To get  $\hat{\sigma}$  we calculate the standard deviation of residuals and substitute this into 3.10. The result are shown in tab.3.4

Calibrated Parameter	Value	Estimated parameter	Value	Mean of RMSerrors <sup>10</sup>	CV(RMSE)
$\hat{a}^*$	0.220	$\hat{a}$	0.1550	0.0047	9.06 %
$\hat{\mu}^*$	0.041	$\hat{\mu}$	0.0665		
$\hat{\sigma}^*$	0.038	$\hat{\sigma}$	0.0233		

Tab. 3.4 Vasicek calibrated and estimated parameters using OLS

<sup>10</sup> Of forward rate given using estimated model



### 3.2.1.2 CIR OLS parameter estimation

For the CIR process it is not possible to give an explicit relation between the parameters of 3.5 and the parameters of 3.2 because the diffusion term is affected by  $\sqrt{r(t)}$ . Then, we have to perform OLS directly on the equation 3.2 by isolating  $\sigma Z_t$  :

$$\sigma Z_t = \frac{r(t + \Delta t) - r(t)}{\sqrt{r(t)}} - \frac{a\mu\Delta t}{\sqrt{r(t)}} + a\sqrt{r(t)}\Delta t$$

And the objective function corresponding to 3.7 is:

$$\min \sum_{t=0}^{T-1} \left( \frac{r(t+1) - r(t)}{\sqrt{r(t)}} - \frac{a\mu\Delta t}{\sqrt{r(t)}} + a\sqrt{r(t)}\Delta t \right)^2 \quad 3.11$$

which is solved by:

$$\hat{a}_{cir} = \frac{T^2 - 2T + 1 + \sum_{t=0}^{T-1} r(t+1) \frac{1}{r(t)} - \sum_{t=0}^{T-1} r(t) \sum_{t=0}^{T-1} \frac{1}{r(t)} - (T-1) \sum_{t=0}^{T-1} \frac{r(t+1)}{r(t)}}{\left( T^2 - 2T + 1 - \sum_{t=0}^{T-1} r(t) \sum_{t=0}^{T-1} \frac{1}{r(t)} \right) \Delta t}$$

$$\hat{\mu}_{cir} = \frac{(T-1) \sum_{t=0}^{T-1} r(t+1) - \sum_{t=0}^{T-1} \frac{r(t+1)}{r(t)} \sum_{t=0}^{T-1} r(t)}{T^2 - 2T + 1 + \sum_{t=0}^{T-1} r(t+1) \frac{1}{r(t)} - \sum_{t=0}^{T-1} r(t) \sum_{t=0}^{T-1} \frac{1}{r(t)} - (T-1) \sum_{t=0}^{T-1} \frac{r(t+1)}{r(t)}}$$

And  $\hat{\sigma}_{cir}$  is found as a standard deviation of residuals. As for the Vasicek model, it is possible to solve 3.11 using *fminsearch* function in Matlab. The result of the estimation obtained by using this function are shown in tab 3.5.

Calibrated Parameter	Value	Estimated parameter	Value	Mean of RMSerrors <sup>11</sup>	CV(RMSE)
$\frac{\partial \alpha_{CIR}}{\partial \alpha}$ meter	0.262	$\frac{\partial \alpha_{CIR}}{\partial \alpha}$ meter	0.1771	0.0038	7.32 %
$\frac{\partial \beta_{CIR}}{\partial \beta}$ meter	0.035	$\frac{\partial \beta_{CIR}}{\partial \beta}$ meter	0.0469		
$\frac{\partial \gamma_{CIR}}{\partial \gamma}$ meter	0.029	$\frac{\partial \gamma_{CIR}}{\partial \gamma}$ meter	0.0260		

Tab. 3.5 CIR calibrated and estimated parameters using OLS

### 3.2.2 Maximum likelihood estimation

Given an observation  $r$  on a random variable, we define *likelihood* the value of the density function associated to the random variable considered at  $r$ . The *likelihood* of a sample  $\mathbf{r} = (r_1, r_2, \dots, r_n)$  is the product of the values of the density function at each  $r_i$ . We can write:

$$L(\theta|\mathbf{r}) = \prod_i^n f(r_i|\theta). \quad 3.12$$

The value of the density function at each point depend on the parameters  $\theta$ , then, if the parameters change, even the value of the likelihood will change. The *Maximum likelihood estimator* of  $\theta$  is the value  $\hat{\theta}$  of  $\theta$  that maximize the likelihood function:

$$\hat{\theta} = \max L(\theta|\mathbf{r})$$

In general, it is more convenient to work with the logarithm of the likelihood (*log-likelihood*) because we have to differentiate the likelihood function to find the value of  $\hat{\theta}$ , and it is much easier to differentiate the *log-likelihood* because is a sum of log-density:

$$\ln L(\theta|\mathbf{r}) = \sum_{i=1}^n \ln f(r_i, \theta). \quad 3.13$$

The logarithmic function is a monotonic increasing function and then the value  $\hat{\theta}$  which maximize the likelihood also maximize the log-likelihood.

<sup>11</sup> Of forward rate given using estimated model

### 3.2.2.1 Vasicek MLE

As show in chapter 1, in the Vasicek model  $r(t)$  is normally distributed with :

$$E[r(t + \Delta t)|r(t)] = e^{-\alpha\Delta t}r(t) + \mu(1 - e^{-\alpha\Delta t})$$

$$VAR[r(t + \Delta t)|r(t)] = \frac{\sigma^2}{2\alpha} (1 - e^{-2\alpha\Delta t})$$

and the conditional probability density of  $r(t + \Delta t)|r(t)$  is given by:

$$f(r(t + \Delta t)|r(t); a, \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2\left(\frac{1 - e^{-2\alpha\Delta t}}{2a}\right)}} \exp\left(-\frac{(r(t + \Delta t) - e^{-\alpha\Delta t}r(t) - \mu(1 - e^{-\alpha\Delta t}))^2}{2\sigma^2\left(\frac{1 - e^{-2\alpha\Delta t}}{2a}\right)}\right)$$

Hence, the log-likelihood function of a set of observation  $\mathbf{r} = (r_1, r_2, \dots, r_n)$  results as follows:

$$\begin{aligned} \ln L(a, \mu, \sigma|\mathbf{r}) &= \sum_{t=0}^{n-1} \ln f(r(t + \Delta t)|r(t); a, \mu, \sigma) \\ &= -\frac{n}{2} \ln(2\pi) - n \ln\left(\sigma^2 \left(\frac{1 - e^{-2\alpha\Delta t}}{2a}\right)\right) \\ &\quad - \sum_{t=0}^{n-1} (r(t + \Delta t) - e^{-\alpha\Delta t}r(t) - \mu(1 - e^{-\alpha\Delta t}))^2 \quad 3.14 \end{aligned}$$

The maximum of this function can be found differentiating with respect to  $a, \mu$  and  $\sigma$  and setting this partial derivative equal to zero [1]:

$$\begin{aligned} \frac{\partial \ln L(a, \mu, \sigma|\mathbf{r})}{\partial a} &= -\frac{\Delta t e^{-\alpha\Delta t}}{\sigma^2 \left(\frac{1 - e^{-2\alpha\Delta t}}{2a}\right)} \sum_{t=1}^{n-1} [(r(t - \Delta t) - \mu)(r(t) - \mu) - e^{-\alpha\Delta t}(r(t) - \mu)^2] = 0 \\ a &= -\frac{1}{t} \ln \frac{\sum_{t=1}^{n-1} [(r(t - \Delta t) - \mu)(r(t) - \mu)]}{\sum_{t=1}^{n-1} [(r(t) - \mu)^2]} \quad 3.15 \end{aligned}$$

$$\frac{\partial \ln L(a, \mu, \sigma | \mathbf{r})}{\partial \mu} = \frac{1}{\sigma^2 \left( \frac{1 - e^{-2\alpha\Delta t}}{2a} \right)} \sum_{t=1}^n r(t + \Delta t) - e^{-\alpha\Delta t} r(t) - \mu(1 - e^{-\alpha\Delta t}) = 0$$

$$\mu = \frac{\sum_{t=1}^n (r(t + \Delta t) - e^{-\alpha\Delta t} r(t))}{n(1 - e^{-\alpha\Delta t})} \quad 3.16$$

$$\frac{\partial \ln L(a, \mu, \sigma | \mathbf{r})}{\partial \sigma} = \frac{n}{\sqrt{\sigma^2 \left( \frac{1 - e^{-2\alpha\Delta t}}{2a} \right)}} - \frac{n}{\left( \sigma^2 \left( \frac{1 - e^{-2\alpha\Delta t}}{2a} \right) \right)^{\frac{3}{2}}} \sum_{t=0}^{n-1} [(r(t + \Delta t) - \mu) - e^{-\alpha\Delta t} (r(t) - \mu)]^2$$

$$\sigma^2 = \frac{1}{n \left( \frac{1 - e^{-2\alpha\Delta t}}{2a} \right)} \sum_{t=0}^{n-1} [(r(t + \Delta t) - \mu) - e^{-\alpha\Delta t} (r(t) - \mu)]^2 \quad 3.17$$

And after straightforward calculation we get:

$$\hat{\mu} = \frac{r_y r_{xx} - r_x r_{xy}}{n(r_{xx} - r_{xy}) - (r_x^2 - r_x r_y)}$$

$$\hat{a} = -\frac{1}{\Delta t} \ln \frac{r_{xy} - \mu(r_x + r_y) + n\mu^2}{r_{xx} - 2\mu r_x - n\mu^2}$$

$$\hat{\sigma}^2 = \frac{2a}{(1 - e^{-2\alpha\Delta t})n} (r_{yy} - 2e^{-\alpha\Delta t} r_{xy} + e^{-2\alpha\Delta t} r_{xx} - 2\mu(1 - e^{-\alpha\Delta t})(r_y - e^{-\alpha\Delta t} r_x) + n\mu^2(1 - e^{-\alpha\Delta t})^2)$$

Where:

$$r_x = \sum_{t=0}^{T-1} r(t), \quad r_y = \sum_{t=0}^{T-1} r(t+1), \quad r_{xx} = \sum_{t=0}^{T-1} r(t)^2,$$

$$r_{xy} = \sum_{t=0}^{T-1} r(t)r(t+1), \quad r_{yy} = \sum_{t=0}^{T-1} r(t+1)^2.$$

The Matlab implementation of the MLE estimation give us the following result:

Calibrated Parameter	Value	Estimated parameter (MLE)	Value	Mean of RMSerrors <sup>12</sup>	CV(RMSE)
$\bar{a}^*$	0.220	$\hat{a}$	0.2044	0.0045	8.68 %
$\bar{\mu}^*$	0.041	$\hat{\mu}$	0.0565		
$\bar{\sigma}^*$	0.038	$\hat{\sigma}$	0.0411		

Tab. 3.4 Vasicek calibrated and estimated parameters (MLE)

### 3.2.2.2 CIR Maximum Likelihood estimation

As shown in chapter 1, the distribution function of  $r(t)$  follows, up to a scale factor, a Non-central chi-squared distribution. Then we have:

$$p(r(t + \Delta t)|r(t)) \sim \frac{\sigma^2(1 - e^{-a\Delta t})}{4a} \chi_d^2 \left( \frac{4ae^{-a\Delta t}}{\sigma^2(1 - e^{-a\Delta t})} r(t) \right) \quad 3.18$$

Where  $\chi_d^2(\lambda)$  denotes the non-central chi-squared distribution with  $d$  degrees of freedom equal to:

$$d = \frac{4\mu a}{\sigma^2}$$

and non-centrality parameter  $\lambda$  given by:

$$\lambda = \frac{4ae^{-a\Delta t}}{\sigma^2(1 - e^{-a\Delta t})} r(t)$$

Given this distribution of the  $r(t)$  we cannot find (as made for Vasicek) an explicit formulas for the values of the parameters to estimate. Then, in the algorithm developed in Matlab we need to calculate first the log-likelihood function using 3.13 and in order to maximize this we will

<sup>12</sup> Of forward rate given using estimated model

minimize the value of the log-likelihood function (using *fminsearch*)<sup>13</sup>. The result of the estimation (see Matlab code in appendix C). The estimation is made for any simulation of  $r^*(t)$ , and the resulting parameters are calculated as mean of estimated parameter of each simulation. The result is shown in the following table:

Calibrated Parameter	Value	Estimated parameter (MLE)	Value	Mean of RMSerror <sup>14</sup>	CV(RMS E)
$\sigma_{CIR}$	0.262	$\sigma_{CIR}$	0.2703	0.0034	7.32 %
$\alpha_{CIR}$	0.035	$\alpha_{CIR}$	0.0334		
$\beta_{CIR}$	0.029	$\beta_{CIR}$	0.0292		

Tab. 3.5 CIR calibrated and estimated parameters (MLE)

### 3.3 Pricing

Another comparison that can be made between the different models analyzed regards the prices of the options that they generate. Of course, to make this kind of comparison we have to choose the options having an explicit form solution that allows an easy pricing for all the model used. Relying on this issue, we chose to calculate the prices of the *caplet*. A *caplet* is an European call option having as underlying the interest rate  $L(S,T)$  prevailing at time  $S$ , which pay at the maturity date  $T$  the payoff<sup>15</sup>:

$$\max(0, L(S, T) - K) * (T - S) * N \quad 3.19$$

where  $N$  is the notional amount on which the interest rate is calculated (in the following paragraph this amount is assumed equal to 1 for simplicity).

Such kind of option is mainly used to cover a short position on  $L(S,T)$  against a possible increase in the level of rates. For example, when I have a position that requires payment of the

<sup>13</sup> The use of *fminsearch* require to assign an initial value to the parameters to be found, follows [1] we set these initial value equal to values estimated using OLS estimation.

<sup>14</sup> Of forward rate given using estimated model

<sup>15</sup> The difference (T-S) is called *tenor*

floating rate  $L(S,T)$  at the date  $T$  and I want to fix (equal to  $K$ ) the maximum amount I will have to pay, in fact if I buy a caplet:

$$\text{TOT PAYOFF: CAPLET PAYOFF} - \text{SHORT POSITION} = (L(S,T) - K)^+ - L(S,T)$$

If  $L(S,T) > K$  the total payoff is  $-K$ , if  $L(S,T) < K$  the total payoff is  $-L(S,T)$  which is less than  $K$ , then the maximum amount I will have to pay is  $K$ .

The payoff in 3.19 can be written in term of forward rate  $L(S,S,T) = L(S,T)$  that is a  $P(0,T)$ -martingale<sup>16</sup>, then:

$$C(0, R(t, S, T), K) = P(0, T) E_t^{P(0, T)} (\max(0, L(S, S, T) - K)) (T - S)$$

And assuming  $L(t, S, T)$  lognormally distributed with deterministic volatility, i.e.:

$$dL(t, S, T) = \sigma_B L(t, S, T) dW^{P(t, T)} \quad 3.20$$

it is possible to apply the Black formula<sup>17</sup>:

$$C(0, L(0, S, T), K) = (T - S) P(0, T) [L(0, S, T) * N(d_1) - K * N(d_2)] \quad 3.21$$

$$d_1 = \frac{\ln\left(\frac{L(0, S, T)}{K}\right) + \frac{1}{2} \sigma_B^2 (S - 0)}{\sigma_B \sqrt{S - 0}}$$

$$d_2 = d_1 - \sigma_B \sqrt{S - 0}$$

Where  $\sigma_B$  is determined by the market, or in our case by the model used.

### 3.3.1 SABR Caplet price

We can now use the result of the second chapter to apply the SABR formula for European option as in equation 2.2. Hence, to get the caplet price using SABR model is sufficient to put into the Black formula 3.21 the following value of volatility:

---

<sup>16</sup> See [10]

<sup>17</sup> See [57]

$$\begin{aligned} \sigma_B(K, L(0, S, T)) &= \\ &= \frac{\sigma_0^T}{(L(0, S, T) * K)^{(1-\beta)/2} \left( 1 + \frac{(1-\beta)^2}{24} \log^2(L(0, S, T)/K) + \frac{(1-\beta)^4}{1920} \log^4(L(0, S, T)/K) + \dots \right)} \\ &\cdot \left( \frac{z}{x(z)} \right) \cdot \left( 1 + \left( \frac{(1-\beta)^2}{24} \frac{(\sigma_0^T)^2}{(L(0, S, T) * K)^{(1-\beta)} + \frac{1}{4} \frac{\rho\beta\alpha\sigma_0^T}{(L(0, S, T) * K)^{(1-\beta)/2}} \right. \right. \\ &\quad \left. \left. + \frac{2-3\rho^2}{24} v^2 \right) S \right) \end{aligned}$$

Where

$$z = \frac{\alpha}{\sigma_0^T} (L(0, S, T) * K)^{(1-\beta)/2} \log(L(0, S, T)/K)$$

And

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

Note that  $P(0, T)$  in 3.21 appears only as “discounting term”, and since we are not in the context of “real market” we cannot derive this from real data. We “solve” this problem by calculating the value of  $P(0, T)$  using the ATSm model. Then, for the SABR caplet price that has to be compared with the Vasicek caplet price we calculate  $P(0, T)$  using this last model, and we do the same with the CIR model. It is clear that this approach is far from what is done in the market, but in the “artificial world” in which we are operating it is a necessary forcing that allows us to compare these models.

### 3.3.2 Vasicek and CIR Caplet prices

To price a caplet with an ATS model we have to rearrange the payoff. Following the book of Cesari [58] we can make the following consideration: the payoff at time  $T$   $(L(S, T) - K)^+ * (T - S)$  is equivalent to the payoff at time  $S$   $(L(S, T) - K)^+ * (T - S) * P(S, T)$ . In fact, I can invest the payoff at time  $S$  in the bond  $P(S, T)$ , obtaining  $(L(S, T) - K)^+ * (T - S) * 1$  at time  $T$ . Now we manipulate the form of the payoff in  $S$  in the following way assuming a simply compound interest rate:

$$(L(S, T) - K)^+ * (T - S) * P(S, T) =$$



$$\begin{aligned}
&= (T - S) * P(S, T) * \left( \frac{1}{T - S} \left( \frac{1}{P(S, T)} - 1 \right) - K \right)^+ \\
&= (1 - (1 + (T - S)K) * P(S, T))^+ \\
&= (1 + (T - S)K) * \left( \frac{1}{1 + (T - S)K} - P(S, T) \right)^+ \\
&= (1 - (1 + (T - S)K) * P(S, T))^+
\end{aligned}$$

From which follows that a caplet can be interpreted as a European Put option with maturity  $S$  on the  $(1 + (T - S)K)$  T-ZCB and strike equal to 1. As shown in [10] the price at time zero of a Put option written on a zero coupon bond  $P(0, T)$ , having strike  $X$  and maturing at time  $S$  is given by:

In the Vasicek model:

$$ZBP^{VAS}(t, S, T, X) = XP(t, S)N(-h + \sigma_p) - Nom * P(t, T)N(-h)$$

with

$$\begin{aligned}
h &= \frac{1}{\sigma_p} \ln \left( \frac{Nom * P(t, T)}{X * P(0, S)} \right) + \frac{\sigma_p}{2} \\
\sigma_p &= \frac{\sigma}{a} (1 - e^{-a(T-S)}) \sqrt{\frac{1 - e^{-2a(T-t)}}{2a}}
\end{aligned}$$

Where  $Nom$  is the Nominal value of the ZCB.

In the CIR model:

$$ZBP^{CIR}(t, S, T, X) = P(t, T)\chi^2(x_1; v, w_1) - XP(t, S)\chi^2(x_2; v, w_2) - P(t, T) + XP(t, S)$$

where

$$\begin{aligned}
x_1 &= 2\bar{r}(\rho + \delta + B(S, T)) \\
x_2 &= 2\bar{r}(\rho + \delta) \\
v &= \frac{4a\mu}{\sigma^2} \\
w_1 &= \frac{2\rho^2 r(t) e^{(S-t)\sqrt{a^2 + 2\sigma^2}}}{\rho + \delta + B(S, T)} \\
w_2 &= \frac{2\rho^2 r(t) e^{(S-t)\sqrt{a^2 + 2\sigma^2}}}{\rho + \delta}
\end{aligned}$$

With  $\rho, \delta$  and  $\bar{r}$  defined as

$$\rho = \frac{2\sqrt{a^2 + 2\sigma^2}}{\sigma^2(e^{(S-t)\sqrt{a^2+2\sigma^2}} - 1)}$$

$$\delta = \frac{a + \sqrt{a^2 + 2\sigma^2}}{\sigma^2}$$

$$\bar{r} = \frac{\ln\left(\frac{C(S,T)}{X}\right)}{B(S,T)}$$

And  $\chi^2(x_1; \nu, w_1)$  represent the cumulative distribution function of the non-central chi-squared distribution with  $\nu$  degree of freedom and non-centrality parameter  $w$ . Both Vasicek and CIR caplet price are calculate also using Monte Carlo technique by simulating<sup>18</sup> the zero coupon bond price and calculating the discounted value (multiplying by  $P(0,S)$ ) of the mean of the payoffs.

### 3.3.3 Pricing result and conclusions

Table 3.6 contains the pricing result of a *caplet* with maturity  $S=1y$ , tenor  $=1y$ , strike  $= 0.06$  and written on a notional equal to 1:

SABR		Vasicek (calibrated)	CIR (calibrated)	Vasicek (estimated)	CIR (estimated)
0.0056	Monte Carlo	0.0051	0.0053	0.050	0.053
	Black formula	0.0051	0.0052	0.050	0.051

Table 3.6 Comparison of Caplet prices among SABR, Vasicek and CIR model

The result shown big difference in the caplet price among the models. The relative error made using the calibrated CIR model (equal to 5.3% using Monte Carlo and 7.1% using Black formula) is smaller than that committed using the calibrated Vasicek model (equal to 8.9% using both Monte Carlo and Black formula). The relative error is bigger if we look at the estimated model. Considering the estimation errors, the differences in prices are in line with the

<sup>18</sup> The caplet price is very sensitive to the number of simulations. Choosing  $n=10000$ , the resulting price of one simulation can differ up to 50% from the price of another simulation. For this reason we have set  $n=1000000$

expectation. In fact, we can deduce that the main cause of the price differences is due to the failure of the affine term structure model in replicating the forward rate behavior of the SABR model, and this results in different values of  $L(S, T)$  among the models, from which follows a difference value of  $P(S, T)$  and then of the caplet prices. Furthermore, looking at the market, the caplets are not traded singularly; they are the building blocks of more complex derivative contracts, the CAPs. This contract is essentially a sum of caplets having different maturity and the same tenor. We have not found precise data about the size of the CAPs market, also because these contracts are negotiated in the OTC market. However, a report of the Bank International Settlement<sup>19</sup> show that for single-currency interest rate derivatives at end-December 2014, the notional amount of outstanding contracts totaled \$505 trillion, which represented 80% of the global OTC derivatives market. Then, given the magnitude of the interest rate option market and the exposure of financial institutions on interest rate derivatives, it is clear the importance of using the most accurate models possible for pricing and hedging these options. As already said, the SABR model is probably the best model for pricing and hedging interest rate option, but when in a portfolio there are other positions, like for example on a certain quantity of bonds, these cannot be priced using Vasicek or CIR model.

---

<sup>19</sup> See [http://www.bis.org/publ/otc\\_hy1504.pdf](http://www.bis.org/publ/otc_hy1504.pdf)

## Matlab Code

### Vasicek calibration

```
clc
clear all
close all

load ('Sabr_path')
data=Sabr_path(1,:);

a=[0.001:0.001:0.001]; % alpha
b=[0.001:0.001:0.1]; % mu
c=[0.001:0.001:0.20];%sigma
o=length(data(:,1));

time_step=length(data(1,:));
nSim=100;

al=length(a);
bl=length(b);
cl=length(c);
T_ex=1; T_set=2;

tots=al*bl*cl;

path_R = zeros(nSim,time_step*T_set/T_ex);
r0=0.03;
mean_R=zeros(1,time_step);
path_R(:,1) = r0;

delta_t=1/time_step;
r=zeros(nSim,time_step);

zcb_set= zeros(nSim,time_step);
zcb_ex= zeros(nSim,time_step);
forward_rate=zeros(1,time_step);
B_Tex=zeros(1,time_step);
B_Tset=zeros(1,time_step);
A_Tex=zeros(1,time_step);
A_Tset=zeros(1,time_step);

error=zeros(nSim,time_step);
```

```

mean_error=zeros(1,1);
mean_error_vector=zeros(tots,1);
count=0;
%calibration referred to a single SABR simulation
for i=1:a1
    for j=1:b1
        for w=1:c1

            for h=1:nSim
                for q=1:(time_step*(T_set/T_ex)-1)
                    %per ogni tripletta ci calcoliamo delle
simulazioni di r
                    path_R(h,q+1)=path_R(h,q)*exp(-
a(i)*delta_t)+ b(j)*(1-exp(-a(i)*delta_t))+c(w)*sqrt((1-exp(-
2*a(i)*delta_t))/2*a(i))*randn();

                    end
                end

%
                for s=1:nSim
                    for m=1:time_step
                        x(i)=(T_set-(i-1)/time_step);
                        y(i)=(T_ex-(i-1)/time_step);
                        B_Tex(m)= (1-exp(-a(i)*y(m)))/a(i);
                        B_Tset(m)= (1-exp(-a(i)*x(m)))/a(i);

                        A_Tex(m)= (((c(w)^2)/(2*a(i)^2))-b(j))*((y(m))-
B_Tex(m))-((c(w)^2)/(4*a(i)))*B_Tex(m)^2;
                        A_Tset(m)=(((c(w)^2)/(2*a(i)^2))-b(j))*((x(m))-
B_Tset(m))-((c(w)^2)/(4*a(i)))*B_Tset(m)^2;

                        zcb_ex(s,m)=exp(A_Tex(m)-path_R(s,m)*B_Tex(m));
                        zcb_set(s,m)=exp(A_Tset(m)-
path_R(s,m)*B_Tset(m));
                        forward_rate(s,m)=((log(zcb_ex(s,m))-
log(zcb_set(s,m))))/(T_set-T_ex);

                    end
                end

                for ss=1:nSim %fisso la riga vasicek
                    for dd=1:time_step
                        error(ss,dd)=(data(dd)-
forward_rate(ss,dd))^2; %per ogni riga sabr calcolo la root
mean square deviation rispetto alle simulazioni vasicek
                    end
                end
            end
        end
    end
end

```

```

mean_error=sqrt((1/(time_step*nSim))*sum(sum(error(:,:))));
        count=count+1
        mean_error_vector(count,1)=mean_error;
[val,idx]=min(mean_error_vector) ;
    end
end
end

```

## Ols vasicek

```

clc
clear all
close all

load Sabr_path
data=Sabr_path;
n=length(data(1,:));
nSim=length(data(:,1))
delta_t=1/(n-1)

for i=1:nSim

rx(i)=sum(data(i,1:n-1));
ry(i)=sum(data(i,2:n));
rxx(i)=sum(data(i,1:n-1).^2) %k
rxy(i)=sum(data(i,1:n-1).*data(i,2:n)); %k
ryy(i)=sum(data(i,2:n).^2);

b(i)=(n*rxy(i)-rx(i)*ry(i))/(n*rxx(i)-(rx(i)^2));
phi(i)=(ry(i)-b(i)*rx(i))/n;
s.devW(i)=sqrt((n*ryy(i)-(ry(i)^2)-b(i)*(n*rxy(i)-
rx(i)*ry(i)))/(n*(n-2)));

a(i)=-log(b(i))/delta_t
mu(i)=phi(i)/(1-b(i));
sigma(i)=s.devW(i)*sqrt((-2*log(b(i)))/(delta_t*(1-b(i)^2)));

end
aa=mean(a);
mm=mean(mu);
ss=mean(sigma);

```

**MLE Vasicek**

```

clc
clear all
close all

load path_R
data=path_R
n=length(data(1,:));
nSim=length(data(:,1))
delta_t=1/(n-1)

for i=1:nSim

rx(i)=sum(data(i,1:n-1));
ry(i)=sum(data(i,2:n));
rxx(i)=sum(data(i,1:n-1).^2) %k
rxy(i)=sum(data(i,1:n-1).*data(i,2:n)); %k
ryy(i)=sum(data(i,2:n).^2);

mu(i)= ((ry(i)*rxx(i)-rx(i)*rxy(i))/(n*(rxx(i)-rxy(i))-
((rx(i)^2)-rx(i)*ry(i))));
a(i)= -(1/delta_t)*log((rxy(i)-mu(i)*rx(i)-
mu(i)*ry(i)+n*mu(i)^2)/(rxx(i)-2*mu(i)*rx(i)+n*mu(i)^2));

sigma2(i)=((2*a(i))/(1-exp(-a(i)*2*delta_t)))*(1/n)*(rxx(i)-
2*(exp(-a(i)*delta_t))*rxy(i)+(exp(-2*a(i)*delta_t))*rxx(i)-
2*mu(i)*(1-exp(-a(i)*delta_t))*(ry(i)-rx(i)*exp(-
a(i)*delta_t))+n*(mu(i)^2)*(1-exp(-a(i)*delta_t))^2);

end
aa=mean(a)
mm=mean(mu)
ss=mean(sigma2)

```

**CIR OLS and ML estimation**

```

clc
clear all
close all

load PathCalibratedCIR

CIR.Data = PathCalibratedCIR;
Results = CIRestimation(CIR)

```

---

```

function lnL = loglike(Parameters, CIR)
    Data = CIR.Data;
    Data1 = Data(2:end);
    Data2 = Data(1:end-1);
    TimeStep = 1/250;
    a = Parameters(1);
    mu = Parameters(2);
    sigma = Parameters(3);

    df=(4*mu*a)/sigma^2
    nc=Data1*(4*a*exp(-a*TimeStep))/((sigma^2)*(1-exp(-
a*TimeStep)))
    tdf =(((sigma^2)*(1-exp(-a*TimeStep)))/4*a)*
ncx2pdf(Data2, df, nc); %
                                                    transition
density function
    lnL = sum(-log(tdf));
end

```

```

function Results = CIRestimation(CIR)
    % OLS estimation
    Nobs = length(CIR.Data);
    r = CIR.Data(1:end-1);
    dr = diff(CIR.Data);
    dr = dr./r.^0.5;
    regressors = [CIR.TimeStep./r.^0.5, CIR.TimeStep*r.^0.5];
    drift = regressors\dr;
    res = regressors*drift - dr;
    a = -drift(2);
    mu = -drift(1)/drift(2);
    sigma = sqrt(var(res, 1)/CIR.TimeStep); % standard
deviation of residual
    InitialParameters = [a mu sigma];

    % MLE with initial parameter given by OLS

    [Parameters] = fminsearch(@(Parameters)
loglike(Parameters, CIR), InitialParameters);
    Results.Parameters = Parameters;
end

```



**SABR Caplet price**

```

clc
clear all
close all

K=0.060
CapData=[K 13]; %13 indica che la day count basis è calcolata
considerando 1y=252days
FwdRates=0.05252;
ZeroPrice=89;%P(0,T)
Settle=0;
StartDate=250;
EndDate=500;
sigma0=0.029;
rho=-0.25;
beta=0.5;
v=0.61;
z=(v/sigma0)*((0.5*K)^((1-beta)/2))*log(0.5/K);
x=log((z-rho+((1-2*rho*z+(z^2))^(1/2)))/(1-rho));
Sigma=(sigma0/(((0.05*K)^((1-beta)/2))*(1+(((1-
beta)^2)*(log(0.5/K))^2)/24)+(((1-
beta)^4)*(log(0.5/K))^4)/1920)))*(z/x)*(1+(((1-
beta)^2)/24)*((sigma0^2)/(0.5*K)^(1-
beta)))+(1/4)*(rho*beta*v*sigma0)/(0.5*K)^(1-
beta)/2)+((v^2)*(2-3*rho^2)/24))*1

CapPrices = bkcaplet(CapData, FwdRates, ZeroPrice, Settle,
StartDate, EndDate, Sigma)/100 %because the default notional
is 100

```

**Vasicek Caplet price (Black-76)**

```

clc
clear all
close all

a=0.4244;
mu=0.0465;
sigma=0.0311;
strike_cap=0.060
quantity=1+strike_cap
PT=0.8905;
PS=0.9458;
K=1
L=1%face value
sigmap=(sigma/a)*(1-exp(-a))*sqrt((1-exp(-2*a))/2*a);
h=(1/sigmap)*log((L*PT*1.06)/(K*PS))+(sigmap/2);
price=(K*PS*normcdf(-h+sigmap)-L*PT*normcdf(-h)*quantity)

```

**CIR caplet price (Black-76)**

```

r0=0.0551;
k=0.2703;
mu=0.0434;
sigma=0.0272;
PS=0.954;
PT=0.896;
CST=0.9946;
BST=0.8852;
quantity=1.06;
K=1
rho=(2*sqrt((k^2)+(2*sigma^2)))/((sigma^2)*(exp(sqrt((k^2)+(2*
sigma^2)))-1));
delta=(k+sqrt((k^2)+(2*sigma^2)))/(sigma^2);
r1=(log(CST/K)/BST);
x1=2*r1*(rho+delta+BST);
x2=2*r1*(rho+delta);
v=4*k*mu/(sigma^2);
w1=(2*(rho^2)*r0*exp(sqrt((k^2)+(2*sigma^2))))/(rho+delta+BST)
;
w2=(2*(rho^2)*r0*exp(sqrt((k^2)+(2*sigma^2))))/(rho+delta);

price=((quantity*PT*ncx2cdf(x1,v,w1)-K*PS*ncx2cdf(x2,v,w2))-
quantity*PT+K*PS)

```

**CIR caplet price (Monte Carlo)**

```

function pathFcir =
CirForwardSimulation(r0,k,mu,sigma,strike,T_set,T_ex);
clc
close all
clear all
r0=0.0551;
k=0.2703;
mu=0.0434;
sigma=0.0272;
T_set=2;
T_ex=1;
time_step=250;
delta_t=T_set/time_step;
n_sim=1000000;
c=((sigma^2)*(1-exp(-k*delta_t)))/(4*k) ;
path_R = zeros(n_sim,time_step+1);
path_R(:,1) = r0;
gamma=sqrt((k^2)+2*(sigma^2));
strike=0.06;
Z=randn(n_sim,time_step);

```

```

lambda=zeros(n_sim,time_step);
d=(4*mu*k)/(sigma^2);
x=chi2rnd(d-1,[n_sim,time_step]);

path_R = zeros(n_sim,time_step);
path_R(:,1) = r0;
zcb_set= zeros(n_sim,time_step*T_ex);
zcb_ex= zeros(n_sim,time_step*T_ex);
B_Tex=zeros(n_sim,time_step*T_ex);
B_Tset=zeros(n_sim,time_step*T_ex);
C_Tex=zeros(n_sim,time_step*T_ex);
C_Tset=zeros(n_sim,time_step*T_ex);

    for j=1:n_sim
        for i=1:time_step-1
            lambda(j,i)=path_R(j,i)*(((4*k*exp(-
k*delta_t)))/((1-exp(-k*delta_t))*sigma^2));

path_R(j,i+1)=(((Z(j,i)+sqrt(lambda(j,i)))^2)+x(j,i))*c;
            end
        end

        for j=1:n_sim
            for i=1:time_step*T_ex
                n(i)=(T_set-(i-1)/time_step);
                m(i)=(T_ex-(i-1)/time_step);

                B_Tex(j,i)= (2*(exp(gamma*m(i))-
1))/((k+gamma)*((exp(gamma*m(i))-1))+2*gamma);

C_Tex(j,i)=((2*gamma*exp((k+gamma)*m(i)/2))/((k+gamma)*((exp(g
amma*m(i))-1))+2*gamma))^(2*k*mu)/(sigma^2));

                B_Tset(j,i)= (2*(exp(gamma*n(i))-
1))/((k+gamma)*((exp(gamma*n(i))-1))+2*gamma);

C_Tset(j,i)=((2*gamma*exp((k+gamma)*n(i)/2))/((k+gamma)*((exp(
gamma*n(i))-1))+2*gamma))^(2*k*mu)/(sigma^2));

                zcb_ex(1,1)=C_Tex(1,1)*exp(-
path_R(1,1)*B_Tex(1,1));
                zcb_set(1,1)=C_Tset(1,1)*exp(-
path_R(1,1)*B_Tset(1,1));
            end
        end

        for g=1:n_sim

```

```

        payoff(g)=max(0,(1-1.05*zcb_set(g,time_step)));
    end
    mean_payoff=mean(payoff(:))
    price_capletCIR=1.06*zcb_ex(1,1)*mean_payoff

end

```

### Vasicek caplet Price

```

function pathFvas =
vasicekForwardSimulation(r0,k,mu,sigma,strike,T_set,T_ex);

close all
clear all
clc

r0=0.0579;
k=0.175;
mu=0.0465;
sigma=0.02911;
T_set=2;
T_ex=1;
time_step=250;
delta_t=T_set/time_step;
n_sim=100000;
strike=0.06
path_R = zeros(n_sim,time_step);
path_R(:,1) = r0;
zcb_set= zeros(n_sim,time_step*T_ex);
zcb_ex= zeros(n_sim,time_step*T_ex);
forward_rate=zeros(n_sim,time_step*T_ex);
B_Tex=zeros(n_sim,time_step*T_ex);
B_Tset=zeros(n_sim,time_step*T_ex);
A_Tex=zeros(n_sim,time_step*T_ex);
A_Tset=zeros(n_sim,time_step*T_ex);

    for j=1:n_sim
        for i=1:time_step*T_ex-1
            path_R(j,i+1)=path_R(j,i)*exp(-k*delta_t)+ mu*(1-
exp(-k*delta_t))+sigma*sqrt((1-exp(-
2*k*delta_t))/2*k)*randn();
        end
    end

    for j=1:n_sim
        for i=1:time_step*T_ex

```

```

        x(i)=(T_set-(i-1)/time_step);
        y(i)=(T_ex-(i-1)/time_step);
        B_Tex(j,i)= (1-exp(-k*y(i)))/k;
        B_Tset(j,i)= (1-exp(-k*x(i)))/k;
        A_Tex(j,i)= (((sigma^2)/(2*k^2))-mu)*(y(i)-
B_Tex(j,i))-(((sigma^2)/(4*k))*B_Tex(j,i)^2);
        A_Tset(j,i)= (((sigma^2)/(2*k^2))-mu)*(x(i)-
B_Tset(j,i))-(((sigma^2)/4*k)*B_Tset(j,i)^2);
        zcb_ex(j,i)= 1/(exp(-
A_Tex(j,i)+path_R(j,i)*B_Tex(j,i)));
        zcb_set(j,i)=1/(exp(-
A_Tset(j,i)+path_R(j,i)*B_Tset(j,i)));

        end
    end

    for g=1:n_sim
        payoff(g)=max(0,(1-(1+strike)*zcb_set(g,time_step)));
    end
    mean_payoff=mean(payoff(:))
    price_capletVas=(1+strike)*zcb_ex(1,1)*mean_payoff

end

```



### Bibliography

- [1] Aït-Sahalia, Y., 2002, Maximum likelihood estimation of discretely sampled diffusion: A closed-form approximation approach. *Econometrica*, 70, 223–262.
  
- [2] Ait-Sahalia, Y. (1996a), Testing Continuous-Time Models of the Spot Interest Rate," *Review of Financial Studies*, 9, 385-426.
  
- [3] Alexander C. , Market Risk Analysis, Practical Financial Econometrics (Volume II), 2008
  
- [4] Andersen L. , Andreasen J. , Volatility skews and extensions of the libor market model, *Applied Mathematical Finance* 7:1 (2000) 1–32.
  
- [5] Andersen, L., Brotherton-Ratcliffe R., Extended libor market models with stochastic volatility, *Journal of Computational Finance* 9:1 (2005) 1–27
  
- [6] Andersen, L. Simple and Efficient Simulation of the Heston Stochastic Volatility Model. *Journal of Computational Finance*, 11:1–42, 2008
  
- [7] Björk, T. (2009) *Arbitrage Theory in Continuous Time*. 3rd ed., Oxford, Oxford University Press.
  
- [8] Black, F. (1976) The Pricing of Commodity Contracts. *Journal of Financial Economics*, 3, pp. 351-67.
  
- [9] Borodin S. , *Handbook of Brownian Motion*, 2nd ed., Birkhauser, 2002
  
- [10] Brigo, D., and Mercurio, F. (2006) *Interest Rate Models – Theory and Practice*. 2 ed., Heidelberg, Springer
  
- [11] Broadie, M. and Kaya, O. Exact simulation of stochastic volatility and other affine jump diffusion processes. *Operations Research*, 542, 2006.
  
- [12] Cerrato M, , *The Mathematics of Derivatives Securities with Applications in MATLAB*, 2012

- [13] Chan, K. C., G. A. Karolyi, F. A. Longstaffe, and A. B. Sanders 1992, An Empirical Comparison of Alternative Models of Short Term Interest Rates, *Journal of Finance*, 47, 1209–1227.
- [14] Chen, B., Oosterlee, C.W. and van der Weide, H. Efficient unbiased simulation scheme for the SABR stochastic volatility model. *Preprint*, 2010
- [15] Conley, T., L.P. Hansen, E.G.J. Luttmer and J.A. Scheinkman (1997), "Short-Term Interest Rates as Subordinated Diffusions," *Review of Financial Studies*, 10, 525-577.
- [16] Cox, J., Ingersoll, J., and S. Ross, 1985, A theory of the term structure of interest rates, *Econometrica*, 53, 385–407
- [17] Cox, John C. and Stephen A. Ross, 1976. "The valuation of options for alternative stochastic processes." *Journal of Financial Economics* 3, pages 145-166.
- [18] Dupire, B. Pricing with a Smile. *Risk*, 7:18–20, 1994.
- [19] Feller, W. Two Singular Diffusion Problems. *Annals of Mathematics*, Vol. 54, No. 1 (July,1951)
- [20] Fima C. Klebaner , *Introduction To Stochastic Calculus With Applications*, 2nd Edition, 2005
- [21] Glasserman, P. and Kim, K.-K. Gamma Expansion of the Heston Stochastic Volatility Model. *Preprint*, [ssrn.com](http://ssrn.com), 2008.
- [22] Glasserman, P, *Monte Carlo Methods in Financial Engineering*, 2003
- [23] Hagan, P.S., Kumar, D., Lesniewski, A.S. and Woodward, D.E. Managing Smile Risk. *Wilmott Magazine*, 1:84–108, 2002.
- [24] Hagan, P., Lesniewski, A. and Woodward, D. Probability Distribution in the SABR Model of Stochastic Volatility. *Working Paper*, 2005.



- [25] Hamilton – Time Series Analysis (1994)
- [26] Heath, D., Jarrow, R., Morton, A. (1992) Bond Pricing and the Term Structure of Interest Rates: A New Methodology. *Econometrics*, 60, pp. 77-105.
- [27] Henriksen P. N., Lognormal moment matching and pricing of basket options, statistical research report no 1, issn 0806–3842, January 2008
- [28] Heston, S.L. (1993), “A closed-form solution for options with stochastic volatility with applications to bond and currency options,” *Review of Financial Studies*, vol. 6, no. 2, pp. 327-343.
- [29] Honorè P. , Maximum Likelihood Estimation of Non-Linear Continuous-Time Term-Structure Models, 1997
- [30] Huu Tue Huynh, Van Son Lai and Issouf Soumar, Stochastic Simulation and Applications in Finance with MATLAB, 2008
- [31] Ioannis Karatzas, Steven E. Shreve, Brownian Motion and Stochastic Calculus, Springer Science & Business Media, 1991
- [32] Islah O., Solving SABR in Exact Form and unifying it with LIBOR Market Model, SSRN eLibrary.2009
- [33] Luc Bauwens , Christian M. Hafner, Sebastien Laurent, Handbook of Volatility Models and Their Applications, 2012
- [34] Wooldridge Jeffrey M. Introductory Econometrics: A Modern Approach, 2012
- [35] Joanne E. Kennedy, Subhankar Mitra, and Duy Pham, on the approximation of the SABR model: a probabilistic approach, 2012
- [36] Joerg Kienitz, Daniel Wetterau, Financial Modelling : Theory, Implementation and Practice with MATLAB Source, 2012

- [37] Johnson, N. L., S. Kotz, N. Balakrishnan. 1994. *Continuous Univariate Distributions*, Vol. 2, 2nd ed. John Wiley and Sons, New York.
- [38] Lo, A. W., 1988, Maximum Likelihood Estimation of Generalized Itô Processes with Discretely Sampled Data, *Econometric Theory*, 4, 231–247.
- [39] Lord, R., R. Koekkoek and D. van Dijk (2006), “A Comparison of biased simulation schemes for stochastic volatility models,” Working Paper, Tinbergen Institute.
- [40] Melino, A. (1994), “Estimation of Continuous-Time Models in Finance,” in C.A. Sims, ed., *Advances in Econometrics, Sixth World Congress, Volume II*, Cambridge University Press, Cambridge, 313-351.
- [41] Merton, Robert C., 1973. “Theory of rational option pricing.” *Bell Journal of Economics and Management Science* 4, pages 141-183.
- [42] Patnaik, P. (1949), “The non-central chi quadro distribution- and  $F$ -distributions and their applications,” *Biometrika*, 36, pp. 202-232
- [43] Phillips, P.C.B. (1972): The Structural Estimation of a Stochastic Differential Equation System. *Econometrica* 40, 1021-1041.
- [44] Phillips, P.C.B. and J. Yu, 2011, Corrigendum to “A Gaussian Approach for Continuous Time Models of the Short Term Interest Rate”, *Econometrics Journal*, forthcoming.
- [45] Phillips, P.C.B. and J. Yu, 2009a, Maximum Likelihood and Gaussian Estimation of Continuous Time Models in Finance. *Handbook of Financial Time Series*, 497-530.
- [46] Riccardo Rebonato, Kenneth McKay, Richard White *The SABR/LIBOR Market Model: Pricing, Calibration and Hedging for Complex Interest-Rate Derivatives*, 2010
- [47] Roman N. Makarov and Devin Glew, Exact Simulation of Bessel Diffusions, *Monte Carlo methods and application* 16(3), 283-306 (2010)

- [48] Sawyer, K.R. (1993), "Continuous-Time Financial Models: Statistical Applications of Stochastic Processes," in G.S. Maddala, C.R. Rao and H.D. Vinod, eds., *Handbook of Statistics*, Vol. 11, Elsevier Science Publishers, 739-764
- [49] Schroder M., Computing the constant elasticity of variance option pricing formula, *Journal of Finance* 1:44 (1989) 211 – 218.
- [50] Torben G. Andersen, Richard A. Davis, Jens-Peter Kreiß, Thomas Mikosch , *Handbook of Financial Time Series*
- [51] Van Haastrecht, A. and Pelsser, A. Efficient, almost exact simulation of the Heston stochastic volatility model. *Preprint, SSRN.com*, 2010.
- [52] Vasicek, O., 1977, An Equilibrium Characterization of the Term Structure, *Journal of Financial Economics*, 5, 177–186.
- [53] Wilmott P., *Paul Wilmott on Quantitative Finance 3 Volume Set (2nd Edition)*, 2006
- [54] Wilmott, P., *The Best of Wilmott 1: Incorporating the Quantitative Finance Review*, 2004
- [55] Wooldridge J. M. , *Introductory Econometrics: A Modern Approach*, 2012
- [56] Karatzas I., Shreve S., *Brownian Motion and Stochastic Calculus*, Springer 1991
- [57] F. Black, The pricing of commodity contracts, *Jour. Pol. Ec.*, 81 (1976), pp. 167-179.
- [58] C. Riccardo, *Introduzione alla finanza matematica, Derivati, prezzi e coperture*, Springer 2009
- [59] Haven Kyle D., *Efficient SABR Simulation Schemes Revisited*, 2015
- [60] L. Yuan, J. Kalbfleisch, On the Bessel distribution and related problems, *Annals of the Institute of Statistical Mathematics* 53:3 (2000) 438 – 477.



## **AKNOWLEDGMENTS**

Un ringraziamento alla mia famiglia e ai miei amici. Ognuno di loro è stato fondamentale e ha saputo supportarmi e sopportarmi giorno dopo giorno facendomi sentire sempre molto fortunato. Un ringraziamento particolare lo devo a mio padre, che mi è stato sempre vicino come amico e come genitore, soprattutto nei momenti difficili.

Sono molto grato al Prof. Cappuccio per la disponibilità avuta, per il tempo che mi ha dedicato e per i preziosi consigli che ha saputo darmi.

Infine grazie ad una persona speciale.

