He_Yizhe_02177261

by Yizhe He

Imperial College London

IMPERIAL COLLEGE LONDON

DEPARTMENT OF MATHEMATICS

CIR++ vs Shifted Squared Vasicek in Interest Modelling

Author: Yizhe He (CID: 02177261)

A thesis submitted for the degree of

MSc in Mathematics and Finance, 2021-2022

Declaration

The work contained in this thesis is my own work unless otherwise stated.

Acknowledgements

I would like to express my deepest gratitude to Prof. Damiano Brigo, my academic supervisor as well as my personal tutor, for allowing me the opportunity to do the research. His invaluable patience and detailed guidance helped me a lot throughout the study period.

My sincere thanks also go to the MSc Mathematics and Finance program, for the cutting-edge curriculum, superb faculty, and unparalleled career resources.

Last but not least, I am grateful to my family and my friends for their continuing companionship, encouragement, and support, which are indispensable and extremely meaningful to me.

Abstract

This paper looks at a shifted squared Vasicek model as an alternative to the CIR++ model for interest rates. The advantage of shifted squared Vasicek is that it works with a shifted chi-squared distribution as opposed to shifted non-central chi-squared. Also, the work attempts a calibration to either old or new market data consisting of a zero-coupon curve through the models' shift and of several caps through the model dynamic parameters, comparing it with the classic CIR++ model. The results show that both models can fit the zero-coupon curve perfectly and a set of caps with the same excellent efficiency and produce curve patterns similar to each other. It compares and analyzes the simulation speeds of these two models. By comparing and analyzing the simulation speeds of these two models, we conclude that the shifted squared Vasicek model is faster and superior.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Interest Rate and Basic Definitions

The idea of interest rates has become ingrained in our daily lives and has become a familiar concept that we are able to manage. Everyone anticipates that when they put some money into a bank, the money will increase at a specific pace over time. It is common knowledge and conventional sense that borrowing money requires remuneration. Therefore, obtaining a particular sum of cash tomorrow is not the same as receiving the identical sum of cash now.

Interest rates exist in a wide variety of formats. The rates that are utilized by the government and those that are used by interbank institutions may be distinguished from one another in several ways. The issuance of bonds by governments is a common factor considered for calculating the rates charged by governments. Instead, when we refer to "interbank rates", we are referring to the rates at which deposits are traded from one bank to another, as well as the rates at which swap transactions take place between banks.

The LIBOR (which means London Interbank Offered Rate) rate, which is fixed daily in London, is often regarded as the most significant of all the interbank rates that may be used as a point of reference when negotiating contracts. Although there are similar mechanisms for establishing interbank rates in other different places, when we want to mention any of the interbank rates, we just call them by that acronym "LIBOR".

The following is a condensed version of an introduction to some fundamental definitions of the realm of interest rates. [1, Page 84-90]

- **Bank account:** For $t \geq 0$, the worth of a bank account is expressed by the notation $B(t)$, which has the following definition:

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

where r_t means the instantaneous spot rate (briefly referred to as the short rate) driving the development of a bank account, which is a positive function of t. Thus, when investing one unit of money at time 0, what we will get at time t is the value:

$$B(t) = \exp\left(\int_0^t r_s ds\right)$$

- **Stochastic discount factor:** The term "discounting" refers to the process of applying the appropriate rates in order to bring the valuation of a future payoff that will occur at maturity T back to the present moment t. We are able to discount using the numeraire B because this serves as a standard for the value of risk-free money over a period of time. As a result, in order to calculate the discount for a payoff from its maturity T to a time t which is earlier than T, we multiply the payoff by

$$D(t,T) = \frac{B(t)}{B(T)} = \frac{\exp\left(\int_0^t r_s ds\right)}{\exp\left(\int_0^T r_s ds\right)} = \exp\left(-\int_t^T r_s ds\right)$$

- **Risk neutral Valuation:** To quickly review the risk neutral valuation paradigm that Harrison et al. (1983) introduced, which exemplifies the no-arbitrage theory: a stochastic payoff in the future V_T, which is constructed on an underlying, will be paid at a future time T, and must fulfill certain technical constraints, has a unique price at present t, which is the risk-neutral expectation:

$$E_t^B\left[\frac{B(t)}{B(T)} V_T\right] = E_t^Q\left[\exp\left(-\int_t^T r_s ds\right) V_T\right]$$

- **Zero-coupon bond:** In our daily life, we often think about how much we need to pay to the bank today to get a unit of cash in one year. To solve this problem, we want to know what is the price at t, and this is where a zero-coupon bond comes from. It is a binding agreement that ensures the payout of one unit of money at the specified time T. Obviously, $P(T,T) = 1$ holds for all maturity T. The value of the agreement at moment t, which is earlier than T, is shown as the symbol $P(t,T)$:

$$P(t,T) = E_t^Q\left[\frac{B(t)}{B(T)} 1\right] = E_t^Q[D(t,T)]$$

The factor $D(t,T)$ denotes an equal quantity of money, while $P(t,T)$ implies a price of a contract. Once we have the concept of $P(t,T)$, we can easily define a lot of different interest rates, such as linear rate $L(t,T)$, continuously compounded rate $R(t,T)$, and so on.

- **Zero coupon curve:** There are a few distinct models for interest rate curves to choose from. The zero coupon curve, also known as the term structure or yield curve, might be any one of the following curves $T \mapsto L(t,T), T \mapsto R(t,T)$ etc. One of the most popular ways to define it is as a plot, taken at time t, of the function that is described as follows:

$$T \mapsto \begin{cases} L(t,T) & \text{if } T \leq t+1 \text{ years} \\ Y(t,T) & \text{if } T > t+1 \text{ years} \end{cases}$$

The initial value $r_t \approx L(t, t+\epsilon)$

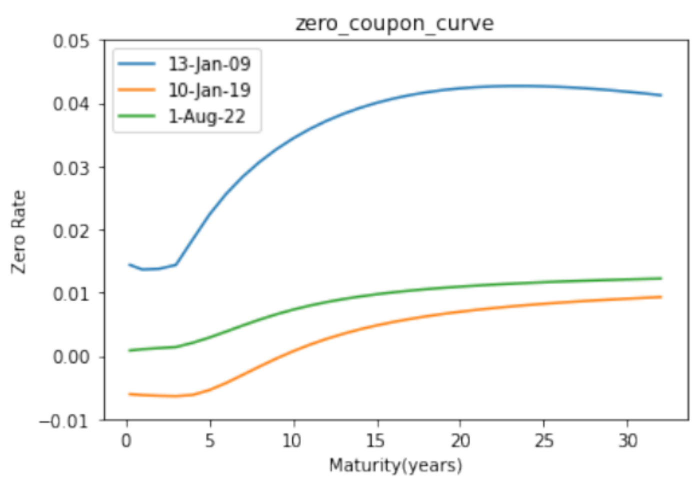

Figure 1.1: Zero-coupon curve of market EURO rates on three different dates

Figure 1.1 depicts some examples of this kind of curve. From the figure, we can see that the curves on January 13, 2009 and August 1, 2022 are not monotonic and there is an inverted shape at the beginning. A possible explanation is that investors are more willing to invest in two weeks instead of one year. The market there may have less liquidity and activity, so the rates are lower. Generally speaking, the curve of European rates is often monotonic, which is shown as the one on January 10, 2009.

Additionally, rates in recent years have been substantially lower than in previous years, and may even be negative, which would not have happened decades ago. The negative rates mean that rather than putting your money into a bond or in a bank, you should put it in the market. It is kind of like you need to pay the bank to take your money.

1.2 Interest-Rate Derivative Products

A financial instrument having a value that is tied to the changes in interest rates is referred to as an interest rate derivative. Contracts such as futures, options, and swaps may fall within this category. On the one hand, It is common practice for investment firms, banks, companies, and private investors to use these kinds of derivatives as hedging tools in order to protect themselves against fluctuations in market interest rates. On the other hand, the holders may utilize them to raise or modify their risk profile, as well as predict rate changes in the coming days.

The primary distinction between the two kinds of derivative products is based on whether or not it is dependent on the curve dynamics [2]. In the first group, which does not rely on dynamics, the primary products include FRA (the abbreviation of "forward rate agreement") and swaps. They do not require a model to be priced, so to price these products at a given time, we do not need to know how the rate curve will move. What we need is just the curve shape. While the second category, which is diametrically opposed to the previous one, is comprised largely of caps and swaptions. Next, I will proceed to provide an overview of these products.

- **FRA:** A forward rate agreement, also known as a FRA, is a kind of contract that contains three different time moments: the present time, which is denoted by t, the future expiry time, which is represented by $T(T > t)$, and the final maturity date, which is characterized by $S(S > T)$. In exchange for a payment for the period $T \mapsto S$ with rate $L(T, S)$, the agreement provides its owner with a payment for the same time with a set rate K at the agreement's maturity S. We represent the selected time measure between T and S using the notation $\tau(T, S)$. This measure of time is often referred to as the year fraction. $L(T, S)$ is a variable at the current moment, the random payoff $\tau L(T, S)$ is something we do not know now, but τK is known because K is a fixed rate. Thus, we can conclude that by FRA, we are agreeing to exchange linear rates between two future time instants for a fixed desired rate K which is known now. [1, Page 99-101]

  The price of the agreement at current moment t is

  $$\mathbf{FRA}(t, T, S, \tau(T, S), N, K) = N[P(t, S)\tau(T, S)K - P(t, T) + P(t, S)]$$

  where N represents the nominal amount of the agreement.

- **Swap:** The most fundamental and widespread kind of interest-rate derivative is referred to as swap. A swap involves two different parties: the one who pays money based on the fixed rate and receives a stream of money at a floating rate, called 'payer', while the other, referred to as 'receiver', receive a stream of money at K. The notional value for both money streams is identical. The two parties hope that by exchanging cash flows with one another, they can lessen the degree of uncertainty and the risk of loss that are caused by swings in the market rates.

  As what we learned in the lectures, the value of a payer swap can be computed as follows:

$$\begin{aligned}\mathbf{RFS}\left(t,\mathcal{T},\tau,N,K\right) &= \sum_{i=\alpha+1}^{\beta}\mathbf{FRA}\left(t,T_{i-1},T_i,\tau_i,N,K\right)\\ &= -NP\left(t,T_\alpha\right)+NP\left(t,T_\beta\right)+N\sum_{i=\alpha+1}^{\beta}\tau_i KP\left(t,T_i\right)\end{aligned}$$

  As for the value of the receiver, it's just analogous.

- **Caplet and Cap:** If a firm has a loan with a variable rate and does not want to switch to a loan with a fixed rate but still wants some protection, it may consider a cap. One way to think of a **cap** is as a kind of payer swap, in which an exchange payment is executed only if it is expected to result in a positive value. Correspondingly, a **floor** may be interpreted in the same manner as a receiver swap, which can be executed only if the value is positive.

  The value of a cap, after discounting, is [1, Page 109]

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

  The positive part is a non-linear transformation of the future rate. In addition, in order to calculate the expectation of a non-linear transformation, we need to know the whole distribution of future rates, and hence we need to model how this rate will move in the future.

  Caplet is the term that decomposes a cap contract. To put it another way, the payout of a caplet after discounting is [3, Page 17]

$$D\left(t,T_i\right)N\tau_i\left(L\left(T_{i-1},T_i\right)-K\right)^{+}$$

  There is a standard practise for the market would be to price a cap, at time

0, with the sum of Black's formulae as follows [4]:

$$\mathrm{Cap}^{\mathrm{Bl}}\left(0,\mathcal{T},\tau,N,K,\sigma_{\alpha,\beta}\right)=N\sum^{\beta}P\left(0,T_i\right)\tau_i\,\mathrm{Bl}\left(K,F\left(0,T_{i-1},T_i\right),\varsigma_i,1\right), \tag{1.2.1}$$

where

$$\begin{aligned}\mathrm{Bl}(K,F,\varsigma,\omega)&=F\omega\Phi\left(\omega d_1(K,F,\varsigma)\right)-K\omega\Phi\left(\omega d_2(K,F,\varsigma)\right)\\ d_1(k,F,\varsigma)&=\frac{1}{\varsigma}\ln\left(\frac{F}{k}\right)+\frac{\varsigma}{2}\\ d_2(k,F,\varsigma)&=\frac{1}{\varsigma}\ln\left(\frac{F}{k}\right)-\frac{\varsigma}{2}\\ \varsigma_i&=\sigma_{\alpha,\beta}\sqrt{T_{i-1}}\end{aligned}$$

and Φ represents the distribution function of standard normal, $\sigma_{\alpha,\beta}$ is the parameter volatility which we can get from the market data.

- **Swaption:** A (payer) swaption gives us the right, without obligation, to enter into a swap at a point in the future. By this definition, we know that the payoff of a swaption is [1, Page 129]

$$\begin{aligned}&ND\left(t,T_\alpha\right)C_{\alpha,\beta}\left(T_\alpha\right)\left(S_{\alpha,\beta}\left(T_\alpha\right)-K\right)^+\\ =&ND\left(t,T_\alpha\right)\left(\sum_{i=\alpha+1}^{\beta}P\left(T_\alpha,T_i\right)\tau_i\left(F\left(T_\alpha;T_{i-1},T_i\right)-K\right)\right)^+.\end{aligned}$$

If we put the '+' inside, it is a call option on the future forward rate, similar to a cap, which is a cap option on the future LIBOR rate. But now, as for a swaption, we do not have an option on every exchange of swap, we only have an option at the first date T_α. We can only execute it at T_α. What we want to do is to price all the future cash flows at T_α, and consider the total value of the swap. From this point of view, we can see that cap, in a way, protects you more because the product protects every single exchange in the whole process, while a swaption only protects the present value at T_α.

Another thing that is different from caps is that the payoff of this kind of product can not be decomposed. As a consequence of this, the most fundamental element that we need to understand in order to deal with it is the correlation between the rates that are included in the contract, which provides us with the joint action of the rates.

1.3 The Endogenous Term-Structure Model

Interest-rate derivative pricing has long been a significant academic study issue that has been tackled in a variety of ways over the last fifty years. There are many classical models which are focused on modelling the evolution of the short rate, including the Vasicek (1977), Dothan (1978), Cox-Ingersoll-Ross (1985) models, as well as the Exponential-Vasicek model. They all come with their own unique set of benefits and drawbacks, which will discussed in more detail.

Model	Dynamics
Vasicek	$x_t = k(\theta - x_t)\,dt + \sigma dW_t$
CIR	$dx_t = k(\theta - x_t)\,dt + \sigma\sqrt{x_t}dW_t, \quad 2k\theta > \sigma^2$
Dothan	$dx_t = ax_t dt + \sigma x_t dW_t, \quad x_t = x_0 e^{\left(a-\frac{1}{2}\sigma^2\right)t+\sigma W_t}$
Exponential − Vasicek	$x_t = \exp(z_t)\,, \quad dz_t = k(\theta - z_t)\,dt + \sigma dW_t$

Table 1.1: Summary of endogenous term-structure models.

In the Vasicek model [5], it is anticipated that the rate follows a Gaussian distribution and the joint distributions of lots of significant quantities are also normally distributed, which means that it is easy to calculate some analytical formulas. Also, the model is mean-reverting. When the time approaches infinity, the expectation of the rate stabilizes to a constant value with finite variance. It's important to note that Gaussian rates may take on negative values with a possibility. In the past, this was a disadvantage, particularly when rates were only positive in a few countries and regions, such as Japan. But as time goes by, interest rates have lately dropped below zero for several maturities, and as a result, we are pleased with the negative rates feature in this instance. Last but not least, because Gaussian distributions for the rates have tails that are excessively thin, they are incompatible with the inferred distributions that are generated by the market. Usually, financial variables are seen to have a fatter tail in the market. From this point of view, the Gaussian distribution may not be the ideal one.

To solve the problem that the short rate, in the Vasicek model, could be negative with a positive probability, Dothan (1978) and Rendleman and Bartter (1980) [6] presented a new model. They assumed that the short rate was log-normally distributed. However, the biggest problem is that this model is not mean-reverting. The expected value rate will become larger and larger or go to zero in time. Intuitively, this is quite different from market data, so it does not make sense from this perspective.

After some time, Cox, Ingersoll, and Ross (1985) (CIR) [7] came up with an alternative idea of a non-central chi-square distribution. In comparison to the Vasicek model, this one has two key distinctions. To begin, this model predicts a short rate that is unambiguously positive. Second, since the distribution of rates has wider tails than the distribution of Vasicek, CIR is often closer to the market-indicated distributions of rates than Vasicek. It is within walking distance to the square of Vasicek. However, when it comes to the quality of being analytically tractable, the fact that the rate is distinguished by a non-central chi-squared distribution appears to mean that we can still compute some formulas, but it is less tractable than Vasicek, in particular for multi-factor extensions where the addition of correlation between components is not possible.

In spite of the fact that the models described above each have their own set of advantages and disadvantages, they all have one important characteristic in common: they are all endogenous [8]. Relying on the model's parameters, $P(t,T) = E_t\left(e^{-\int_t^T r(s)ds}\right)$ could be calculated. In Vasicek and CIR, for instance, if we know $T \mapsto P$, with parameters k, θ, σ and $r(t)$, we are able to determine the total short-rate curve at the given moment t. Instead of being input to the model at the beginning moment ($t = 0$), the term structure of the short rate is an output, and its value is dependent on k, θ, σ, r_0 in the dynamics.

The calibration process is an extremely significant aspect of the operational aspects of a model. What we want to do is to price, hedge, and maybe manage the risk associated with an exotic instrument that is sophisticated and whose quotes are not readily accessed or liquid. In this situation, we need to use a model that is able to take into account the maximum number of liquid market data points that are currently accessible when these data are relevant to the product that is being evaluated. If we already have the real world curve $T \mapsto P^M(0,T)$ and we want to use the model we have to fit this curve, then what we should do is to set the parameters that can generate a curve that is as similar to the market curve as it is feasible to get. This step involves fitting. After the model has been fitted so that it corresponds as closely as possible to the data, it is next used to figure out the price of the complicated product. But on the other hand, the issue is that there are not nearly enough parameters. No matter what values are used for the parameters, it will never be possible to produce some forms, such as an inverted shape. As a result, exogenous yield curve models are typically taken into consideration in order to remedy this circumstance and calibrate the data of caplets, caps, and swaptions.[1, Page 164]

1.4 The Exogenous Term-Structure Model

To create exogenous models, one must first make the necessary modifications to the models shown above. The inclusion of "time-varying" parameters is the fundamental method that is used to change a time-homogeneous model (implies that the dynamics of rates only rely on constant coefficients, as we discussed before) into an exogenous model. For example, one would proceed as follows in the Vasicek case:[1, Page 167]

$$dr(t) = k[\theta - r(t)]dt + \sigma dW(t)$$

$$\longrightarrow dr(t) = k[\vartheta(t) - r(t)]dt + \sigma dW(t)$$

The long-term mean is transformed into a time-dependent function, and it is no longer a constant, and hence that we can put the initial curve from the market inside $\vartheta(t)$. In such a manner that, starting at time 0, the model precisely replicates the actual curve $T \mapsto L^M(0, T)$.

The Hull and White (1990) [9] extended Vasicek model, probable expansions of the Cox, Ingersoll, and Ross (1985) model, as well as the Black and Karasinski (1991) model [10] are all examples of well-known theories. Damiano Brigo and Fabio Mercurio (1998) [11] presented an outstanding technique that demonstrates how to extend a general time-homogeneous model in order to perfectly recreate the yield curve. The CIR++ model is one of the remarkable outcomes that they obtained. Markus Leippold and Liuren Wu (2003) [12] addressed the construction and estimation of general quadratic term structure models. This paper mostly expands upon the methodologies and findings that were provided by Damiano Brigo and Fabio Mercurio (1998).

1.5 Structure of the Thesis

The remaining sections of this thesis are organized as follows: methodology, implementation and results, Monte Carlo simulation, and conclusion.

In Chapter 2, the main structures of the CIR++ model and the Shifted Squared Vasicek model are introduced, including the derivation of analytical formulae for spot rate, options, caps, etc. Section 2.3 shows how to calibrate for real-world data based on exogenous short-rate models. The concepts and calculation methods of market cap volatility and model-implied cap volatility are briefly presented in Section 2.4.

In Chapter 3, we attempt a calibration to market data, comparing it with the classic CIR++ model. The quality of fitting the zero-coupon curve in the two models is first checked. Section 3.2 attempts a calibration to either old or new market data consisting of several caps through the model dynamic parameters.

Chapter 4 introduces the background and methods of Monte Carlo simulation. Through each method, we compare and analyze the simulation speed of these two models.

Finally, the results are summed up in Chapter 5. We also outline the project's future development and potential upgrades.

Chapter 2

Methodology

2.1 The CIR++ Model

The CIR++ model is proposed by Damiano Brigo and is an extension of the Cox-Ingersoll-Ross(1985) model. Here, the fundamental time-homogeneous model evolves in accordance with

$$dx_t^\alpha = k\left(\theta - x_t^\alpha\right)dt + \sigma\sqrt{x_t^\alpha}dW_t$$

where the vector $\alpha = (k, \theta, \sigma)$. From the CIR model, we know that the process x^α follows a non-central chi-square distribution and generates an affine term structure of interest rates. As a result, it is possible to derive analytical equations for the pricing of bond options, caps, floors, and swaptions.

Now we define the new short rate as

$$r_t = x_t^\alpha + \varphi(t;\alpha), \quad t \geq 0,$$

which is a sum of process x_t^α and a deterministic function $\varphi(t;\alpha)$. And $\varphi(t;\alpha)$ is dependent on the parameter vector and integrable on closed intervals. There is another parameter x_0 which satisfies:

$$\varphi(0;\alpha) = r_0 - x_0$$

And $\varphi(t;\alpha)$ can be designed to fit the initial yield curve perfectly. After adding a general shift to the original model, the short-rate dynamics are shown as [13]

$$\begin{aligned} dx(t) &= k(\theta - x(t))dt + \sigma\sqrt{x(t)}dW(t), \quad x(0) = x_0 \\ r(t) &= x(t) + \varphi(t) \end{aligned}$$

where $dW(t)$ are Brownian motions, the parameter vector now is $\alpha = (k, x_0, \theta, \sigma)$

and these parameters are positive constants, whose meanings are as follows:

- k: the speed at which values revert to their mean
- x_0: the initial value of x
- σ: the volatility
- θ: the mean reversion value throughout the long term

The Feller condition $2k\theta > \sigma^2$ should be satisfied, because it guarantees that x stays greater than zero. Now, taking into account the CIR++ model made up of such an extension, we compute the analytical mathematical formulae.

The instantaneous forward rates of the market and model at time t with maturity T are represented by $f^x(t,T;\alpha)$ and $f^M(t,T)$, which satisfy:

$$f^x(t,T;\alpha) = -\partial \ln P^x(t,T)/\partial T$$
$$f^M(t,T) = -\partial \ln P^M(t,T)/\partial T$$

where $P^x(t,T)$ and $P^M(t,T)$ are bond prices of the model and market.

If we assume the original yield curve of the discount terms fits exactly, then we have $\varphi(t;\alpha) = \varphi^{CIR}(t;\alpha)$ where [14]

$$\begin{aligned}\varphi^{CIR}(t;\alpha) &= f^M(0,t) - f^{CIR}(0,t;\alpha),\\ f^{CIR}(0,t;\alpha) &= \frac{2k\theta(\exp\{th\}-1)}{2h+(k+h)(\exp\{th\}-1)}\\ &\quad + \frac{4h^2 x_0 \exp\{th\}}{[2h+(k+h)(\exp\{th\}-1)]^2}\end{aligned} \tag{2.1.1}$$

with

$$h = \sqrt{k^2 + 2\sigma^2}$$

The Bond Price

In addition to this, the price of a zero-coupon bond whose maturity is T at time t is [3]:

$$P(t,T) = \bar{A}(t,T)e^{-B(t,T)r(t)}$$

where

$$\bar{A}(t,T) = \frac{P^M(0,T)A(0,t)\exp\{-B(0,t)x_0\}}{P^M(0,t)A(0,T)\exp\{-B(0,T)x_0\}}A(t,T)e^{B(t,T)\varphi^{CIR}(t;\alpha)}$$

so that

$$P(t,T) = \frac{P^M(0,T)A(0,t)\exp\{-B(0,t)x_0\}}{P^M(0,t)A(0,T)\exp\{-B(0,T)x_0\}}A(t,T)\exp\left\{-B(t,T)\left[r_t - \varphi^{CIR}(t;\alpha)\right]\right\}$$

with

$$A(t,T) = \left[2h\frac{\exp\{(k+h)(T-t)/2\}}{2h+(k+h)(\exp\{(T-t)h\}-1)}\right]^{2k\theta/\sigma^2},$$
$$B(t,T) = 2\frac{(\exp\{(T-t)h\}-1)}{2h+(k+h)(\exp\{(T-t)h\}-1)}.$$

The Spot Rate

Recall the concept of the continuously compounded spot rate, which means that the rate increases exponentially in time. If we invest an amount X at time t at this kind of rate with maturity T, the value of the investment will be $X\exp(R(t,T)(T-t))$ at the given time T. So that, [15]

$$P(t,T)\exp(R(t,T)(T-t)) = 1$$

$$R(t,T) = -\frac{1}{T-t}\ln P(t,T)$$

Notice,

$$r(t) = \lim_{T\to t^+} R(t,T) \approx R(t,t+\epsilon)$$

where ϵ is very small.

Therefore, at time t, the spot rate with maturity T is [3]

$$\begin{aligned} R(t,T) = {} & \frac{1}{T-t}\left[\ln\frac{P^M(0,t)A(0,T)\exp\{-B(0,T)x_0\}}{P^M(0,T)A(0,t)\exp\{-B(0,t)x_0\}}\right] \\ & -\frac{1}{T-t}\left[\ln(A(t,T)+B(t,T)\varphi^{CIR}(t;\alpha)-B(t,T)r_t\right] \end{aligned} \quad (2.1.2)$$

The European Call and Put Option

The value of a European call option at time t withi maturity T, with the strike price K, issued on a zero-coupon bond that would mature at time τ is calculated as the following [14]:

$$\begin{aligned} \mathbf{ZBC}(t,T,\tau,K) = {} & \frac{P^M(0,\tau)A(0,t)\exp\{-B(0,t)x_0\}}{P^M(0,t)A(0,\tau)\exp\{-B(0,\tau)x_0\}} \\ & \cdot\Psi^{CIR}\left(t,T,\tau,K\frac{P^M(0,T)A(0,\tau)\exp\{-B(0,\tau)x_0\}}{P^M(0,\tau)A(0,T)\exp\{-B(0,T)x_0\}}, r_t - \varphi^{CIR}(t;\alpha);\alpha\right), \end{aligned}$$

where $\Psi^{CIR}(t,T,\tau,X,x;\alpha)$ is the option price in CIR model, given $r(t)=x$.

$$\Psi^{CIR}(t,T,\tau,X,x;\alpha) = A(t,\tau)\exp\{-B(t,\tau)x\}\chi^2\left(\mu;\frac{4k\theta}{\sigma^2},\frac{2\rho^2 x\exp\{h(T-t)\}}{\rho+\psi+B(T,\tau)}\right) - XA(t,T)\exp\{-B(t,T)x\}\chi^2\left(2\bar{r}[\rho+\psi];\frac{4k\theta}{\sigma^2},\frac{2\rho^2 x\exp\{h(T-t)\}}{\rho+\psi}\right)$$

with

$$\rho = \rho(T-t) := \frac{2h}{\sigma^2(\exp[h(T-t)]-1)}$$
$$\psi = \frac{k+h}{\sigma^2}$$
$$\bar{r} = \bar{r}(\tau-T) := \frac{\ln(A(T,\tau)/X)}{B(T,\tau)}$$
$$\mu = 2\bar{r}[\rho+\psi+B(T,\tau)]$$

$\chi^2(\cdot;r,\rho)$ represents the CDF of non-central chi-squared distribution [16]. In this formula, r tends to mean the degrees of freedom, while ρ appears to imply the non-centrality parameter. The PDF of this kind of distribution is represented by the notation $p_{\chi^2(r,\rho)}$.

Through the further reduction of complexity, we are able to arrive at the following formulae [11]:

$$\begin{gathered}\mathbf{ZBC}(t,T,\tau,K) = \\ P(t,\tau)\chi^2\left(2\hat{r}[\rho+\psi+B(T,\tau)];\frac{4k\theta}{\sigma^2},\frac{2\rho^2\left[r(t)-\varphi^{CIR}(t;\alpha)\right]\exp\{h(T-t)\}}{\rho+\psi+B(T,\tau)}\right) \\ -KP(t,T)\chi^2\left(2\hat{r}[\rho+\psi];\frac{4k\theta}{\sigma^2},\frac{2\rho^2\left[r(t)-\varphi^{CIR}(t;\alpha)\right]\exp\{h(T-t)\}}{\rho+\psi}\right)\end{gathered} \tag{2.1.3}$$

with

$$\hat{r} = \frac{1}{B(T,\tau)}\left[\ln\frac{A(T,\tau)}{K} - \ln\frac{P^M(0,T)A(0,\tau)\exp\{-B(0,\tau)x_0\}}{P^M(0,\tau)A(0,T)\exp\{-B(0,T)x_0\}}\right]$$

Put–call parity allows for the straightforward calculation of the corresponding put option price, which will be discussed in more detail afterward. At the time t, their prices meet the following equation [3]:

$$\mathbf{ZBC}(t,T,\tau,K) + KP(t,T) = \mathbf{ZBP}(t,T,\tau,K) + P(t,\tau)$$

and hence we get

$$\mathbf{ZBP}(t,T,\tau,K) = \mathbf{ZBC}(t,T,\tau,K) - P(t,\tau) + KP(t,T) \tag{2.1.4}$$

To compute the formula of the put option price, what we need to do is just substitute (2.1.3) into (2.1.4).

Caplet and Cap

By equation (2.1.3), the price formulas of caplet, cap, swaption can all be obtained. Here, the derivation process of caps and caplets will be mainly described. By no-arbitrage theory, the $i-$th caplet's price is [3]

$$\begin{aligned}&\mathbf{CPL}\left(t, t_{i-1}, t_i, \tau_i, N, X\right)\\&\quad= E\left(e^{-\int_t^{t_i} r_s ds} N\tau_i \left(L\left(t_{i-1}, t_i\right) - X\right)^+ \mid \mathcal{F}_t\right)\\&\quad= NE\left(e^{-\int_t^{t_{i-1}} r_s ds} P\left(t_{i-1}, t_i\right)\tau_i \left(L\left(t_{i-1}, t_i\right) - X\right)^+ \mid \mathcal{F}_t\right)\end{aligned}$$

Recall that

$$L\left(t_{i-1}, t_i\right) = \frac{1 - P\left(t_{i-1}, t_i\right)}{(t_i - t_{i-1})P\left(t_{i-1}, t_i\right)}$$

hence

$$\begin{aligned}&\mathbf{CPL}\left(t, t_{i-1}, t_i, \tau_i, N, X\right)\\&\quad= NE\left(e^{-\int_t^{t_{i-1}} r_s ds} P\left(t_{i-1}, t_i\right)\left[\frac{1}{P\left(t_{i-1}, t_i\right)} - 1 - X\tau_i\right]^+ \mid \mathcal{F}_t\right)\\&\quad= NE\left(e^{-\int_t^{t_{i-1}} r_s ds}\left[1 - \left(1 + X\tau_i\right)P\left(t_{i-1}, t_i\right)\right]^+ \mid \mathcal{F}_t\right)\end{aligned}$$

so we obtain

$$\mathbf{CPL}\left(t, t_{i-1}, t_i, \tau_i, N, X\right) = \frac{1}{1 + X\tau_i}\mathbf{ZBP}\left(t, t_{i-1}, t_i, N\left(1 + X\tau_i\right)\right)$$

Following that is the derivation of a cap. To begin, the meanings of a few factors, particularly those relating to time, will be elaborated upon:

- $\mathcal{T}$: $\mathcal{T} = \{t_0, t_1, \ldots, t_n\}$ the set of the exchange times
- t_0: the first reset date
- τ_i: the year fraction between t_{i-1} and t_i, $i = 1, ..., n$
- t: the current time, $t < t_0$
- X: the strike rate
- N: the nominal value

And hence that the price of the cap at time t is [3]

$$\mathbf{CAP}(t, \mathcal{T}, N, X) = N\sum_{i=1}^{n}\left(1 + X\tau_i\right)\mathbf{ZBP}\left(t, t_{i-1}, t_i, \frac{1}{1 + X\tau_i}\right)$$

Analogously, the floor's value is written as

$$\mathbf{FLR}(t, \mathcal{T}, N, X) = N \sum_{i=1}^{n} (1 + X\tau_i) \, \mathbf{ZBC}\left(t, t_{i-1}, t_i, \frac{1}{1 + X\tau_i}\right)$$

2.2 The Shifted Squared Vasicek Model

The Shifted Squared Vasicek model serves as an extension of the squared Vasicek model, also known as SSV, with $\theta = 0$ and a constant μ outside the dynamics. For the original model, the distribution would be chi-squared rather than non-central chi-squared. Then this model is shifted with a constant outside the dynamics, which is given by

$$\begin{aligned} dx(t) &= -kx(t)dt + \sigma dW(t), \quad x(0) = x_0 \\ r(t) &= x(t)^2 + \mu \end{aligned}$$

where μ is a constant. Recall Itô's formula [17],

$$\mathrm{d}\varphi\left(t, X_t\right) = \frac{\partial \varphi}{\partial t} \, \mathrm{d}t + \frac{\partial \varphi}{\partial X} \, \mathrm{d}X_t + \frac{1}{2} \frac{\partial^2 \varphi}{\partial X^2} \, \mathrm{d}X_t \, \mathrm{d}X_t$$

with three key rules

$$\mathrm{d}t \, \mathrm{d}t = 0, \quad \mathrm{d}t \, \mathrm{d}W_t = 0, \quad \mathrm{d}W_t \, \mathrm{d}W_t = \mathrm{d}t$$

Using Itô's formula, the solution to the squared Vasicek is computed as follows:

$$\begin{aligned} dY(t) &= d\left(X_t\right)^2 \\ &= 2X_t dX_t + \langle dX_t \rangle^2 \\ &= 2X_t \cdot (-kX_t dt + \sigma dW_t) + \sigma^2 dt \\ &= -2kX_t^2 dt + 2X_t \sigma dW_t + \sigma^2 dt \\ &= 2k\left(\frac{\sigma^2}{2k} - Y_t\right) dt + 2\sigma\sqrt{Y_t} dW_t \end{aligned}$$

Compared to the CIR model

$$dx(t) = k(\theta - x(t))dt + \sigma\sqrt{x(t)}dW(t),$$

The SV model may be thought of as a specialised version of the CIR model that has undergone certain parameter transformations as follows:

$$k_{CIR} = 2k_{SSV}, \quad \sigma_{CIR} = 2\sigma_{SSV}, \quad \theta_{CIR} = \frac{\sigma_{SSV}^2}{2k_{SSV}}$$

As a result, through the method mentioned in the previous section, it is easy for us to derive the formulas of a zero-coupon bond, the spot rate, European call option, caps, and so on by these parameter transformations and taking into account the constant shift μ.

There is one more significant aspect that has to be brought out. In the Vasicek model, it is a given that we make the assumption that the rate follows a Gaussian distribution. Recall the definition of chi-squared distribution, this kind of distribution (also chi-squared distribution) with k degrees of freedom is the distribution of a sum of the squares of k independent standard normal random variables [18]. And as a consequence of this, the squared Vasicek model follows a chi-squared distribution with one degree of freedom, which is intuitively given by the definitions. At the same time, let's have a look at the steps involved in coming up with formulas. The European Call option price, which is given by (2.1.3), states that the degree of freedom, in the CIR++ model, is equal to $\frac{4k\theta}{\sigma^2}$. By certain transformation of parameters,

$$\frac{4k_{CIR}\theta_{CIR}}{\sigma_{CIR}^2} = \frac{4 \cdot 2k_{SSV} \cdot \frac{\sigma_{SSV}^2}{2k_{SSV}}}{(2\sigma_{SSV})^2} = 1$$

which is consistent with our intuitive results. In conclusion, the distribution in this scenario would be chi-squared rather than non-central chi-squared, which offers substantial advantages in terms of calculation.

2.3 Calibration of Exogenous Model

The basic concept of model calibration has previously been presented in Section 1.3. When it comes to the specific calibration processes, endogenous models and exogenous models are quite different from one another. The figure of the exogenous model, which we learned in the lecture notes [1, Page 176-177], will now be shown.

Figure 2.1 basically demonstrates the key steps that make up the calibration:

- Firstly, a market zero coupon curve of interest rates at time 0 is given to us, such as $T \mapsto L^M(0,T)$.
- Additionally, we have also been provided with a variety of vanilla options' volatilities, which generally consist of caps and a few swaptions. This is the "market volatility" shown in blue.
- The blue arrow illustrates how we now utilize a time-dependent "parameter" $\vartheta(t)$ or a general shift extension which is outside the dynamics $\varphi(t)$ to precisely match the zero curve.

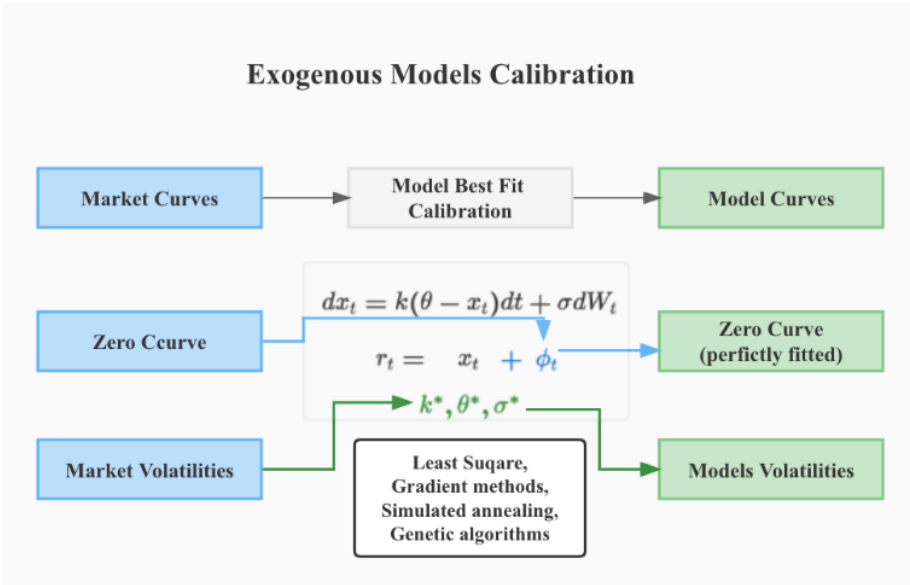

Figure 2.1: Calibration of exogenous models

- Then, we try to get the best possible fit for the data pertaining to the options using r_0, k, θ, σ in x_t, a part of r_t, and this step is indicated by the green arrow.

- To fit the option data as well as possible, we could utilize a variety of optimization techniques, including both local and global ones. The least-square technique, the gradient method, simulated annealing, and genetic algorithms are some examples of common approaches. The major approach used in the implementation of this paper is least square.

- In most cases, the outcomes of fit are not very good. Generally speaking, only when we fit few options, the result will improve.

2.4 Implied Cap Volatility

2.4.1 Market Cap Volatility

When the Black formula is applied to the current option price [19], it generates an output called implied volatility(shortly referred to as vol). As mentioned in section 1.2, [3, Page 17]

$$\begin{aligned}\mathbf{CPL}\left(0, T_1, T_2, X := P\left(0, T_2\right) \tau E\left(F\left(T_1; T_1, T_2\right) - X\right)^{+}\right. \\ = P\left(0, T_2\right) \tau\left[F\left(0; T_1, T_2\right) \Phi\left(d_1\left(X, F\left(0; T_1, T_2\right), \sigma_2 \sqrt{T_1}\right)\right)\right. \\ \left.-X \Phi\left(d_2\left(X, F\left(0; T_1, T_2\right), \sigma_2 \sqrt{T_1}\right)\right)\right], \\ d_{1,2}(X, F, u) = \frac{\ln(F/X) \pm u^2/2}{u}\end{aligned} \tag{2.4.1}$$

using equation (2.4.1) we can price market caps. And in this formula, we quote σ instead of the market price. The following concept will now be described very clear: **the market cap(or caplet) vol** is therefore simply defined as the parameter σ that must be entered into the Black calculation (2.4.1) in order to pricing the correct market cap(or caplet).

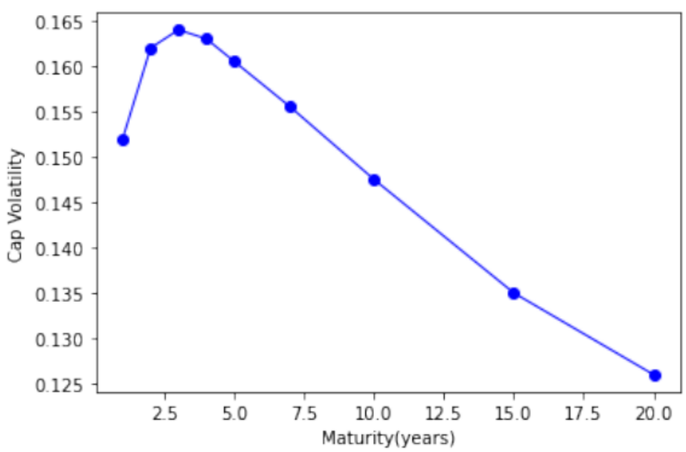

Figure 2.2: At-the money European cap volatility curve on February 13, 2001

Above is an example of a market cap vol curve. There are some different caps, all of which have implied vols that are quoted by the market. In this figure, there are two kinds of caps. In the first one, with a maturity of one year, $\alpha = 0$, T_0 is three months (the year fraction is equal to 0.25) and all other T_i's are three months. On the other hand, caps with maturities of 2, 3, 4, 5, 7, 10, 15 and 20 years, $\alpha = 0$, T_0 is three months, T_1 is after a three-month interval, and all other T_i's have six-month intervals.

2.4.2 Model-implied Cap Volatility

Then we move towards the concepts of **model-implied cap (or caplet) volatility**. Let's start with the implied caplet vol, which can be derived with the following steps.

When we look at the volatility data, we mean, for each caplet maturity T_2 and strike price X, the volatility σ_2 that matches the market price of the caplet when used in the above formula (2.4.1). Once we have a short rate model, we can calculate the price formula of an at-the-money caplet at time zero, and then we can get a mathematical expression $\mathbf{CPL}(0, T_1, T_2, F(0; T_1, T_2))$ with model parameters like k, x_0, θ, σ. For a function $T_2 \mapsto \sigma_2(X, T_2)$, each strike X is the at-the-money strike for that caplet, namely for a given T_2 we pick $X = F(0; T_1, T_2)$. Then we use formula (2.4.1) for a caplet with an expiry T_1 and pay date of T_2 to find the implied volatility.

In conclusion, the model-implied caplet vol, at time zero, is the solution σ^{Model} of the following [3, Definition 3.6.2, Page 90]:

$$P(0, T_2)\, \tau F(0; T_1, T_2) \left(2\Phi\left(\frac{\sigma^{\text{Model}} \sqrt{T_1}}{2} \right) - 1 \right) = \mathbf{CPL}\left(0, T_1, T_2, F(0; T_1, T_2)\right) \tag{2.4.2}$$

The model implied cap volatility is defined in the same way as above. Let the market Black's formula for a cap equal to the model cap price with strike $X = S_{\alpha,\beta}(0)$, where $S_{\alpha,\beta}(0)$ is the forward swap rate, and then we can find the implied volatility σ^{Model} by this expression [3, Definition 3.6.3, Page 90]:

$$\begin{aligned} &\sum_{m=\alpha+1}^{n} P(0, T_m)\, \tau_m \mathrm{Bl}\left(S_{\alpha,\beta}(0), F(0, \; T_{m-1}, T_m), \sigma^{\text{Model}} \sqrt{T_{m-1}}\right) \\ &= \mathbf{CAP}\left(0, \mathcal{T}_n, \bar{\tau}_n, S_{\alpha,\beta}(0)\right) \end{aligned} \tag{2.4.3}$$

Chapter 3

Implementation and Results

3.1 Fit Zero-coupon Curve

In the previous chapter, we have obtained some analytical formulae for both models. The two models investigated in this work, unlike endogenous models, are able to nearly precisely match the zero coupon curve, which will be shown next.

3.1.1 Fit Zero-coupon Curve by CIR++ Model

Using $\varphi^{CIR}(t;\alpha)$, which is derived by (2.1.1), no matter which values of $k, \theta, \sigma$and x_0 are used, this model always produces an accurate representation of the term structure from the market. Below is a fitting example with a random set of parameters, which is almost perfectly fitted. The market data we use is from AAA-rated bonds in the European market on January 10, 2019. The line created by the model clearly illustrates the inverted shape of the market data at the start.

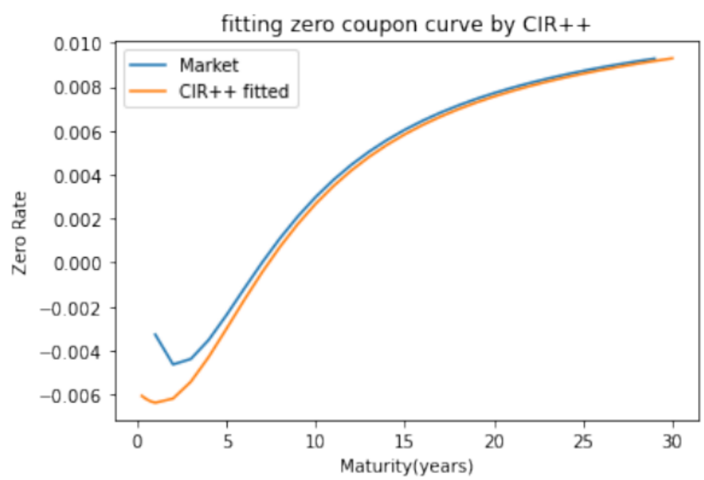

Figure 3.1: Fitting zero coupon curve by the CIR++ model

3.1.2 Fit Zero-coupon Curve by SSV Model

Repeating this step with a random set of parameters yields the curve fitted by the SSV model, as illustrated in Figure 3.2.

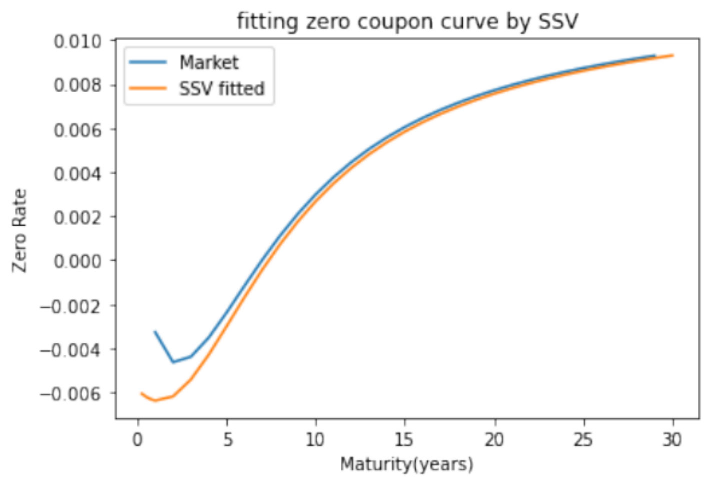

Figure 3.2: Fitting zero coupon curve by the SSV model

Although in this model, the shifted extension is a constant, it can still almost perfectly fit the market data. Because of this shift, we are able to utilise the market curve as an input throughout the process of fitting it, rather than only as an output for which we need to tweak parameters to approximate. Obviously, the accuracy of fitting the term structure by these two models is almost the same.

3.2 Calibration to Real Market Data though Analytical Formula

3.2.1 Case 1 : Calibration to Old Market Data

In this chapter, we calibrate our models through analytical approximation using true volatility data from the market. The at-the-money cap vol data from the European market on February 13, 2001 are the ones that we chose to use. The details are laid out in the Table 3.1.

The key steps of the calibration used are as follows:

- Firstly, use Black's formula (2.4.2) and the market volatility data shown in the table to derive the cap prices.

Maturity in years	Market Cap Vol
1	0.152
2	0.162
3	0.164
4	0.163
5	0.1605
7	0.1555
10	0.1475
15	0.135
20	0.126

Table 3.1: At-the money Cap Volatility in Euro Market on 02/13/2001

- Compute the analytical formula of cap price by the CIR++ model through the methods presented in Chapter 2.
- Calculate the difference in cap prices between market one and model one. Minimize the difference by the optimization method least square to obtain an optimal set of parametersk, x_0, θ, σ.
- Use the definitions and methods shown previously to calculate the implied cap vol by the CIR++ model. The specific calculation formula is shown in equation (2.4.2) and (2.4.3).
- Derive model implied cap vols with the optimal set of parameters k, x_0, θ, σ. Plot the model curve and the market one to observe the fitting accuracy.
- Experiment with various market data combinations, adjust the upper and lower bounds of the parameters and try to seek the ideal parameter set, which can minimize the gap between the market and the model prices while fitting the market cap vol curve as much as possible.
- Repeat the above steps in the SSV model and plot the curves together, which is shown in Figure 3.3.
- Another approach is also conceivable. Instead of computing the difference between cap prices, we can directly derive the formula for model-implied volatility, calculate the difference between it and market ones, and utilize the latter directly as an input to calibrate and discover the optimal parameters.

Figure 3.3 shows that both models can almost perfectly match the cap volatility curve. There is virtually no difference in fitting quality between the two models. θ is a particular parameter derived by $\frac{\sigma^2}{2k}$ in SSV model, which might explain the extremely minor difference.

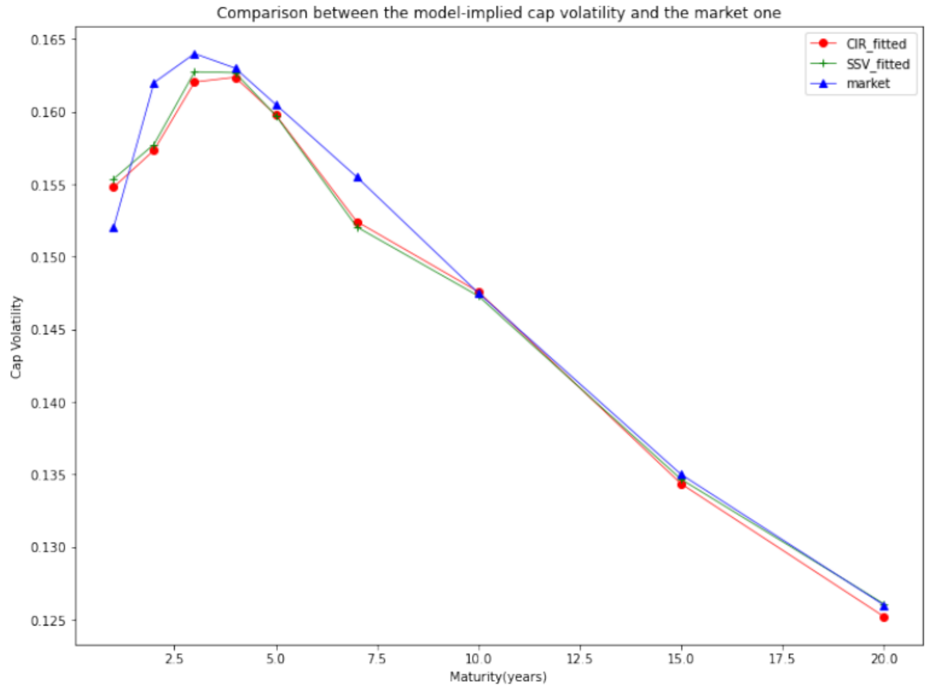

Figure 3.3: Comparison among cap volatility curves implied by the CIR++ model and the SSV model and that obtained in European market on 02/13/2001

CIR++ Model	SSV Model
k=0.13170899	k=0.05666197
x_0=0.00031363	x_0=0.00028553
θ=0.00122602	σ=0.0104549
σ=0.02312062	

Table 3.2: Calibrated parameters in the CIR++ and SSV models of 02/13/2001

The calibrated parameters in these two models are shown in Table 3.2.1. We may deduce from these parameter values that k_{CIR} is almost twice as large as k_{SSV}, σ_{CIR} is roughly twice as large as σ_{SSV}, x_0 is approximately the same in the two different models, and θ_{CIR} is nearly equal to $\frac{\sigma_{SSV}^2}{2k_{SSV}}$. These findings also verify the parameter transformations discussed in the preceding chapter.

3.2.2 Case 2 : Calibration to New Market Data

This research seeks to fit the new market data to assess the calibration quality of the two models after executing a good quality calibration on the old market data. Cap volatility from the European market as seen by Bloomberg on September 1, 2022, was used in this case, as shown in Table 3.3. The maturities of the caps utilised are 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 13, 15 (years), $\alpha = 0$, T_0 is three months, T_1 is after a three-month interval, and all subsequent T_i's are after a six-month gap.

Maturity in years	Market Cap Vol
1	0.157
2	0.163
3	0.155
4	0.143
5	0.132
6	0.121
7	0.114
8	0.107
9	0.101
10	0.095
12	0.091
13	0.086
15	0.086

Table 3.3: At-the-money Cap Volatility in Euro Market on 09/01/2022

Figure 3.4 presents the results of the fitting, in which the volatility curve predicted by the CIR++ model and the SSV model is compared with the market curve. Both models can still fit the most recent market data, which is plotted as a humped curve quite well, with little difference in quality. And the curve patterns they produce are very similar.

The calibrated parameters in these two models are shown in Table 3.2.2. This study explores different market data combinations while adjusting the upper and lower boundaries of the parameters to attain the optimal fitting result. Furthermore, because the least square method is a local optimization method, the optimal parameters may change depending on the bounds, and there may be several sets of parameters that can all produce a very good fitting effect, so the parameters obtained by the two models this time do not have the relationships shown in the previous case.

CIR++ Model	SSV Model
k=1.74312888	k=0.30970661
x_0=0.00423151	x_0=0.00064563
θ=0.17099215	σ=0.01514012
σ= 0.09059899	

Table 3.4: Calibrated parameters in the CIR++ and SSV models of 09/01/2022

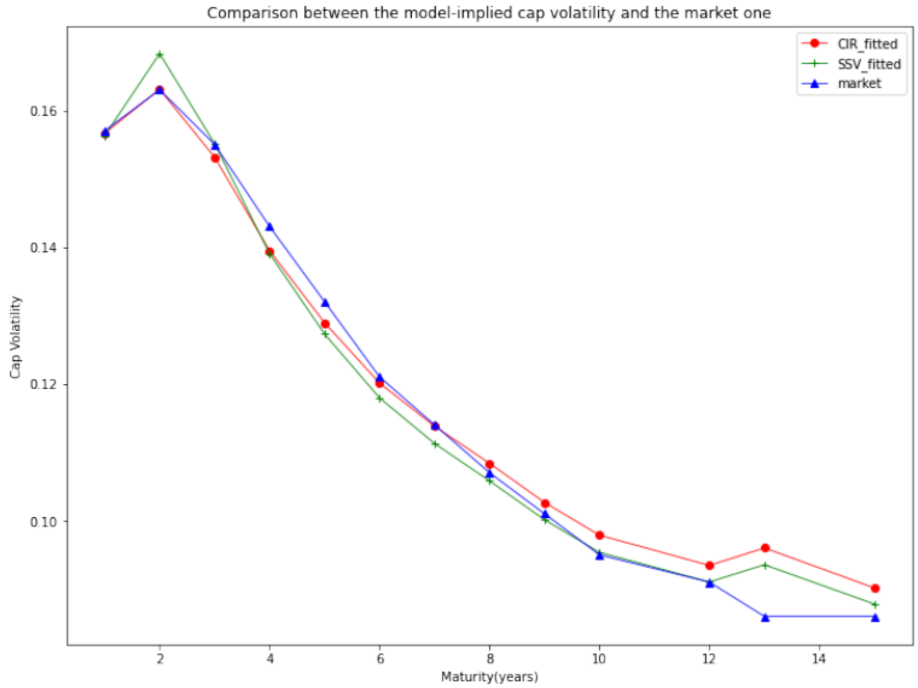

Figure 3.4: Comparison among cap volatility curves implied by the CIR++ model and the SSV model and that obtained in European market on 09/01/2022

One additional thing is noticeable with the above calibration. The shape of the market cap volatility curve we use is all humped, which is because the value of the first point is always lower than the value of the second point. We discovered that for both models, a satisfactory curve fitting may need low x_0 values, which corresponds to the theoretical derivation's finding: in the CIR++ model, at time zero, the volatility of the instantaneous forward rate is [11]

$$\sqrt{x_0}\frac{2h\sigma\exp(Th)}{[2h+(k+h)(\exp\{Th\}-1)]^2}$$

and there is a similar formula in the SSV model with certain parameter transformation.

Chapter 4

Monte Carlo Simulation

4.1 The Concept of Monte Carlo Simulation

There are products on the market with path-dependant payoffs and early exercise payoffs. When we intend to price derivatives whose payoff at final maturity T is a function of rates at a final time connected to the final maturity T as well as rates at earlier periods $t_i < T$, we say we have a path-dependent payout. To be more specific, this occurs when the payment cannot be decomposed into a total of payments, each of which references a single maturity rate at the time. In most cases, it may be required to price these path-dependent payouts using Monte Carlo simulation.

Suppose now we are going to price a path-dependent payoff with Euro exercise characteristics. The discounted payoff is as follows:[3, Page 114]

$$\sum_{j=1}^{n} \exp\left[-\int_0^{t_j} r(s)ds\right] H\left(r\left(t_1\right), \ldots, r\left(t_j\right)\right) \tag{4.1.1}$$

where

- t: the current time and $t = 0$.
- T: the final maturity to exercise products.
- $r\left(t_1\right), \ldots, r\left(t_n\right)$: the rate at time instants $0 < t_1 < t_2 < \ldots < t_n = T$.
- $H\left(r\left(t_1\right), \ldots, r\left(t_j\right)\right)$: the payoff we want to price, which is a function of $r\left(t_1\right), \ldots, r\left(t_n\right)$

The formula of $r\left(t_i\right)$ can be derived by

$$R\left(t_i, t_i + \tau, r\left(t_i\right)\right) := \frac{1}{\tau}\left[1/P\left(t_i, t_i + \tau\right) - 1\right]$$

where $R(t_i, t_i+\tau, r(t_i))$ is the spot rate. And once analytical formulas can be computed for such R, it is easier for us to use the Monte Carlo simulation method to price.

In order to price the term in equation 4.1.1, the common range for the number of short rate paths involved is between 100,000 and 1,000,000. Then we calculate the arithmetic mean of these values using equation 4.1.1 along each path. Additionally, it is very difficult for us to carry out the following term under the risk-neutral measure:

$$\exp\left[-\int_0^{t_j} r_s ds\right]$$

Instead, we can approximate under the T-forward measure because[3, Page 114]

$$\begin{aligned} E\left\{\exp\left[-\int_0^{t_j} r(s)ds\right] H\left(r\left(t_1\right), \ldots, r\left(t_j\right)\right)\right. \\ = P(0,T)E^T\left\{\frac{H\left(r\left(t_1\right), \ldots, r\left(t_j\right)\right)}{P\left(t_j, T\right)}\right\} \end{aligned} \tag{4.1.2}$$

where E^T means the expectation under the T-forward measure. Recall that in both the CIR++ model and the Shifted Squared Vasicek model, the formula for bond price can be derived. Thus, for $r(t_j)$, $P(t_j, T)$ is determined. At this point, there is no need for any more simulation, and we could start to get the paths. Finally, we calculate the payoff according to equation 4.1.2 along each path for r, which is the sum of process x and the deterministic function φ, and compute their average. Further specifics about the methods to get simulated paths can be found in the next section.

4.2 Approaches to get the Sample Paths

The concept and necessity of Monte Carlo simulation have been introduced. Now we turn our attention to the more detailed steps.

In the CIR++ model, φ is a deterministic function, and μ is a constant, which is certainly deterministic in the SSV model. As a result, simulating r is the same as simulating x. Denote $m+1$ sampling times with $0 = s_0 < s_1 < s_2 < \ldots < s_m = T$ including the dates t_j's. And define

$$\Delta s_i := s_{i+1} - s_i, i = 0, \ldots, m-1$$

Then, there are two methods we can use to get the sample paths [20]:

1. Under T-forward measure, approximate numerical solution of the SDE for x, through the Milstein scheme or the Euler.

2. Under T-forward measure [21], at each step, sample the transition density from $x(s_i)$ to $x(s_{i+1})$, $i = 0, \ldots, m-1$.

4.2.1 Analysis through the Milstein Scheme

The first method to get the simulated paths is to approximate the numerical solution of the stochastic differential equation for x, through the Milstein scheme [22].
For the SDE

$$\mathrm{d}X_t = \alpha(X_t)\,\mathrm{d}t + \beta(X_t)\,\mathrm{d}W_t, X_0 = x_0$$

where W_t represents the Brownian Motion. The Milstein scheme for the solution X is the Markov chain S, which will be obtained following the steps below:

- divide the time range $[0, T]$ with length $\Delta t > 0$ into N equal subranges and let $Y_0 = x_0$.

- for $1 \leq n \leq N$, define S_n as [23]:

$$S_{n+1} = S_n + \alpha(S_n)\,\Delta t + \beta(S_n)\,\Delta W_n + \frac{1}{2}\beta(S_n)\,\beta'(S_n)\left((\Delta W_n)^2 - \Delta t\right)$$

- Note that $\beta'(S_n)$ means the derivative of $\beta(S_n)$ wrt S_n and if $\beta'(S_n) = 0$, the diffusion term is not dependent on X_t, so this method equals the Euler method.

Then we calculate the Milstein Scheme of these two models. Take the CIR++ model as an example, the Milstein scheme for x is [3]

$$\begin{aligned} x(s_i + \Delta s_i) =& x(s_i) + \left[k\theta - \left(k + B(s_i, T)\,\sigma^2\right) x(s_i)\right] \Delta s_i \\ & - \frac{\sigma^2}{4}\left(\left(W^T_{s_i+\Delta s_i} - W^T_{s_i}\right)^2 - \Delta s_i\right) \\ & + \sigma\sqrt{x(s_i)}\left(W^T_{s_i+\Delta s_i} - W^T_{s_i}\right) \end{aligned}$$

where $W^T_{s_i+\Delta s_i} - W^T_{s_i}$ is an independent increment of Brownian motion, which follows a normal distribution whose mean equals zero and variance equals Δs_i at each step Δs_i. The following step is to sample this distribution. These normal samples are derived from variables that are independent at various periods. Thus, one can generate a matrix of normally distributed variables whose size is the number of paths times the number of sampling times.

In the specific simulation process, the parameters are set as follows:

- The number of paths p: 200,000
- The time horizon T: 30 (years)
- The number of time steps m: 1000

Then compute and accumulate the increments through the Milstein scheme of these two models using the calibrated parameters in case 1. The time we spend to do the simulation is shown as:

	CIR++ Model	SSV Model
Time (seconds)	52.292016	40.290508

Table 4.1: Comparison of simulation time

It is obvious that the SSV model offers greater benefits in the process of simulation since it requires less time.

4.2.2 Analysis through the Transition Density Form

Now, we analyze the simulation speeds of the two models through the second approach. In the CIR++ model, the dynamics of the short rate x_t under the T-forward measure are given by [11]

$$dx_t^\alpha = \left[k\theta - \left(k + B(t,T)\sigma^2\right)x_t^\alpha\right]dt + \sigma\sqrt{x_t^\alpha}dW_t^T.$$

And under T-forward measure [24], the transition distribution of the short rate x_{s_i} conditional on $x_{s_{i+1}}$ is [11]:

$$\begin{aligned} p^T_{x^\alpha_{s_{i+1}}|x^\alpha_{s_i}}(x) &= p_{\chi^2(v,\delta(s_{i+1},s_i))/l(s_{i+1},s_i)}(x) \\ &= l(s_{i+1},s_i)p_{\chi^2(v,\delta(s_{i+1},s_i))}(l(s_{i+1},s_i)x), \\ l(s_{i+1},s_i) &= 2[\rho(s_{i+1}-s_i)+\psi+B(s_{i+1},T)] \\ \delta(s_{i+1},s_i) &= \frac{4\rho(s_{i+1}-s_i)^2 x^\alpha_{s_i} e^{h(s_{i+1}-s_i)}}{l(s_{i+1},s_i)}, \\ v &= \frac{4k\theta}{\sigma^2} \end{aligned}$$

which means that at each step, it follows a non-central chi-squared distribution, which is a sum of the squares of $\frac{4k\theta}{\sigma^2}$ independent normal random variables. The density function of this kind of distribution can be written as [25]

$$p_X(x;v,\delta) = \frac{1}{2}e^{-(x+\delta)/2}\left(\frac{x}{\delta}\right)^{v/4-1/2} I_{v/2-1}(\sqrt{\delta x})$$

where $I_\nu(z)$ is a modified Bessel function of the first kind shown as

$$I_\nu(z) = (z/2)^\nu \sum_{i=0}^{\infty} \frac{(z^2/4)^i}{i!\Gamma(\nu + i + 1)}$$

In contrast, the SSV model follows a chi-squared distribution whose degree of freedom equals 1 (seen in Section 2.2). The density function is given by [18]

$$f(x;a) = \begin{cases} \frac{x^{\frac{a}{2}-1} e^{-\frac{x}{2}}}{2^{\frac{a}{2}} \Gamma\left(\frac{a}{2}\right)}, & x > 0 \\ 0, & \text{otherwise.} \end{cases}$$

where $\Gamma(a/2)$ represents the gamma function, which has closed-form values for integer a.

Based on the above, we can conclude that the SSV model is superior since the shifted chi square (taking into account the constant μ) is much more tractable than the non-central chi square. It also does not require special functions like the Bessel function, which raises the computational load greatly and slows down the simulation deeply.

Chapter 5

Conclusion

The aim of this paper is to compare a shifted squared Vasicek model to the CIR++ model for interest rates.

To make this paper readable for everyone, the basic concepts and definitions of interest rates, common derivative products, as well as the classic endogenous and exogenous short-rate models are presented in the introduction part. It is important to notice that, except for Vasicek, the CIR model is the only endogenous one with analytical formulas for both bonds and options with an instantaneous rate featuring tails that is similar to the square of normal distribution. This is the reason why we are looking at a shifted squared Vasicek model as an alternative to the CIR++ model.

In Chapter 2, the construction of the CIR++ model is firstly introduced. By adding a deterministic function determined by parameters to the dynamics as a general extension, the original CIR model becomes a time-dependent model, which can use the market zero-coupon curve as an input to perfectly fit the market. What follows is the derivation of analytical formulae of the shift φ, zero-coupon bond, spot rate, European Call and Put option, cap price, etc. Next, the structure of the Shifted Squared Vasicek model is presented. Based on a Squared Vasicek with $\theta = 0$, which is a special CIR model with some special parameter transformation, the SSV model is shifted with a constant outside the dynamics. In this model, the distribution of r is a shifted chi-squared, which is much better than a shifted non-central chi-squared distribution in the CIR++ model. With this model r, it is easy to price options, bonds, caps, and so on using the CIR++ model with the special parameter transformation coming from the calculations and taking into account the constant shift.

Then this work attempts a calibration to market data in some cases, comparing the Shifted Squared Vasicek model with the classic CIR++ model. Firstly, both models fit the zero-coupon curve through the models' shift almost perfectly, regardless of when the market curve is. What's more, calibrations to either the old or new market cap volatility curve using least square as an optimization method through the dynamic parameters are attempted. From the comparison among cap volatility curves implied by the SSV model and the CIR++ model and that obtained in the European market on February 13, 2001, and September 1, 2022, we conclude that these two models have a high level of consistency for the fitting quality. Additionally, by observing the calibrated parameters, we fine that for both models, a satisfactory curve fitting may need low x_0 values, which is consistent with the analytical formula.

Last but not least, this paper moves towards the Monte Carlo simulation. It is necessary to price path-dependent derivatives using Monte Carlo simulation. And to obtain the sample paths, there are two approaches: discretizing the SDE via the Milstein scheme and sampling the transition density at each step under T-forward adjusted measure. By analyzing through the two approaches, we find that the simulation speed of the SSV model is much faster than the other. The shifted chi-square is much more tractable than the shifted non-central chi-square since the latter one requires special functions like the Bessel function, which increases the computational burden and slows down the speed deeply. Thus, the SSV model is superior to the classical CIR++ model. Another point to note is that the CIR++ model and the SSV model have the advantage of allowing for negative rates through a negative shift φ or μ while in the original CIR model, rates are always non-negative.

In conclusion, the SSV model is a good alternative to the classical CIR++ model for interest rates. The advantage of shifted squared Vasicek is working with a chi-squared distribution as opposed to a non-central chi-squared. They can both fit a given set of caps with the same excellent efficiency and produce curve patterns similar to each other. By analyzing and comparing the simulation speed, we find that the SSV model is superior.

As for further work, we can attempt a calibration of a set of swaptions while this paper mainly focuses on the calibration of caps. It would be more interesting to model credit spreads using these two models and study this together with a Cheyette model (a special HJM model, popular in the industry) for interest rates, and to study the pricing of contingent CDS or CVA on a portfolio of swaps or rate products.

Bibliography

[1] Brigo D. Interest rate models with credit risk, collateral, funding liquidity risk and multiple curves. *MSc Mathematics and Finance, Imperial College London*, 2020.

[2] Pelsser A. *Efficient methods for valuing interest rate derivatives*. Springer-Verlag London, 2000.

[3] Brigo D. and Mercurio F. *Interest Rate Models - Theory and Practice*. Springer-Verlag London, 2006.

[4] D. Gatarek A., Brace and Musiela M. The market model of interest rate dynamics. *Mathematical Finance*, 7:127–155, 2002.

[5] Vasicek O.A. An equilibrium characterization of the term structure. *Journal of Financial Economics*, 1977.

[6] DOTHAN L.U. On the term structure of interest rates. *Journal Of Financial Eco- nomics*, 1978.

[7] ROSS S. COX, J.C. and M. RUBINSTEIN. A theory of the term structure of interest rates. *Econometrica*, 1985.

[8] FRANCIS A. LONGSTAFF K. C. CHAN, G. ANDREW KAROLYI and ANTHONY B. SANDERS. An empirical comparison of alternative models of the short-term interest rateh. *The Journal of Financet*, 47:1209–1227, 1992.

[9] HULL J. and WHITE A. Pricing interest-rate-derivative securities. *The Review Of Financial Studies*, 1990.

[10] BLACK F. and KARASINSKI P. Bond and option pricing when short rates are lognormal. *Financial Analysts Journal*, 1991.

[11] Brigo D. and Mercurio F. A deterministic–shift extension of analytically–tractable and time–homogeneous short–rate models. *Finance and Stochastics*, 5:369–387, 2001.

[12] Leippold M. and Wu L. Quadratic term structure models. *SSRN Electronic Journal*.

[13] Brigo D. and Mercurio F. The cir++ model and other deterministic- shift extensions of short-rate models. *Proceedings of the Columbia-JAFEE International Conference*, pages 563–584, 2000.

[14] Brigo D. and Mercurio F. On deterministic shift extensions of short rate models. *SSRN Electronic Journal*.

[15] Cairns A. *Interest rate models*. Princeton University Press, 2004.

[16] Baldeaux J. and Platen E. *Functionals of multidimensional diffusions with applications to finance*. Springer, 2013.

[17] Kiyosi Itô. On stochastic differential equations. *American Mathematical Society*, 4:1–51, 1951.

[18] Chi-squared distribution - wikipedia. https://en.wikipedia.org/wiki/Chi-squared distribution, 2022.

[19] J. D. MacBeth and L. J. Merville. An empirical examination of the black-scholes call option pricing model. *Journal of Finance*, 34:1173–1186, 1979.

[20] Wang L. Numerical solutions of stochastic differential equations. https://trace.tennessee.edu/utk graddiss/3974, 2016.

[21] Musiela M. and Rutkowski M. Continuous-time term structure models: Forward measure approach. *Finance Stochast*, 1:261–291, 1997.

[22] G. N. Mil'shtein. Approximate integration of stochastic differential equations. *Theory of Probability Its Applications*, 19:557–000, 1975.

[23] Milstein method - wikipedia. https://en.wikipedia.org/wiki/Milstein method, 2022.

[24] Daviso M. A note on the forward measure. *Finance and Stochastics*, 2:19–28, 1997.

[25] Noncentral chi-squared distribution - wikipedia. https://en.wikipedia.org/wiki/Noncentral chi-squared distribution, 2022.