Constant Maturity Swap Pricing

by Elissa Ibrahim

Submission date: 10-Sep-2019 02:51PM (UTC+0100)

Submission ID: 110634442

File name: Ibrahim_Elissa_01541406.pdf (1.73M)

Word count: 16441

Character count: 72217

Constant Maturity Swap Pricing

by

Elissa Ibrahim (CID: 01541406)

Department of Mathematics
Imperial College London
London SW7 2AZ
United Kingdom

Thesis submitted as part of the requirements for the award of the
MSc in Mathematics and Finance, Imperial College London, 2018-2019

Declaration

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

Signature and date: Elissa Ibrahim 10/09/2019

Acknowledgements

This thesis marks the end of my masters of Mathematics and Finance at Imperial College London that was a dream of mine. I would like to thank every person who helped me achieve my goal.

Thank you Mr. Piotr Karasinski for giving me the opportunity to work at the European Bank for Reconstruction and development (EBRD) and for teaching me a lot from your experience. My deepest gratitude goes to Guglielmo De Maria and Kaitai Chan for suggesting the topic and for their constant support. I was very lucky to be part of a team who was always ready to help. Thank you Professor Thomas Cass for giving me insightful advice that were useful both on the theoretical and the practical part of my thesis.

Last but not least, I would love to thank my family and friends in Lebanon, who were always by my side and constantly encouraging me to aim for the highest.
I dedicate my thesis to my mother and best friend, for without her endless heartening and cheering, I would have never made it that far and would have not achieved my dream.

Abstract

Constant Maturity Swap (CMS) is a swap where the interest rate on one leg is reset periodically with respect to a market swap rate, instead of being reset with respect to LIBOR. Pricing CMS-linked derivatives is more complicated than pricing a vanilla product because of the "unnatural" payment schedule. One of the most used pricing techniques to price CMS-linked products is pricing by replication. However, due to the lack of swaption market quotes for every possible strike, we need interpolation models to achieve the pricing. The most used smile-consistent interpolation methods in the market are the Vanna-Volga approach and SABR model, since they provide prices that are in line with swaptions market data. We will implement both techniques and compare them with the pricing done under Black's framework which assumes constant volatility. We will consider different bond-math approximation in our pricing and examine the impact of each one on the price.

Contents

1 Introduction

The interest rates market is continuously expanding, and it has undergone significant shifts especially in the last couple of years. It is now considered as the most active asset class with the largest amount of assets in terms of notional. The derivatives in this market are divided into three main categories depending on their complexity: the vanilla, the quasi-vanilla, and the exotic derivatives. Unlike vanilla instruments, quasi vanilla and exotic instruments are less liquid and may have different payment structures or expiration dates. They can be bespoke to a particular client or a specific market.

An example of a quasi vanilla interest rate derivative is the Constant Maturity Swap (CMS) which is a variation of the regular interest rate swap. It is characterised by the fact that its floating leg resets periodically against the rate of a fixed maturity instrument with a longer maturity than the length of the reset period. Whereas, a vanilla swap is a swap where the floating leg is set against a short term rate which is usually LIBOR. Due to the nature of their payment schedules, constant maturity swaps are exposed to changes in the long-term interest rate movements.
Different pricing techniques have been suggested by practitioners in the financial literature. The goal of this thesis is to combine them and compare them to provide the traders of CMS with a good insight into the characteristics of each method.

To understand the basics of CMS derivatives pricing, we start with the pricing of a CMS swap. The price of CMS swaps is the difference between the present value of the LIBOR leg and the present value of the CMS leg. Once the yield curve is known, valuing the LIBOR leg is rather straightforward because it is exactly the same as the case of a vanilla interest rate swap. Valuing the CMS leg, on the other hand, requires a deeper understanding. At each payment date, the CMS leg pays a CMS rate. The CMS rate is defined as the swap rate of an underlying swap with a tenor that is constant and does not depend on the reset frequency. For instance, a 30-year CMS leg depends on the 30-year CMS rate that is independent of whether the reset happens quarterly or semi-annually, which explains the term “constant maturity”. A CMS rate can be viewed as an alternative to LIBOR as a floating index because it allows investors to express their views on the future levels of rates. An important thing that will be emphasized in Section 2.2.2 is that the value of a CMS leg is not equal to the underlying forward swap rate, and we will need to do a change of probability measure in our pricing. Pelsser suggests in [6] that this change of measure should be accompanied by a convexity adjustment. From a modeling perspective, the use of a convexity adjustment occurs every time the underlying financial variable modelled is not a martingale under the pricing measure.

The most used way to calculate the convexity adjustment of a CMS-linked product is proposed by Hagan in [4] and is based on the pricing by replication technique. The idea behind this method is to replicate CMS swaps, caps and floors continuously using Payer and Receiver swaptions. This approach is very interesting because ideally, it would allow a static hedge of a CMS position based on a continuum of vanilla interest rate swaptions. However, swaptions are only quoted in the market for a discrete range of strikes, maturities and tenors. Hence we need a way that provides us with the rest of the swaption values.

Many authors have written around this subject. Hagan proposed bond-math approximations in [4] and provided closed form formulas for CMS coupons under Black's model based on the assumption of constant volatility. However, using this model is not consistent with the swaption smile or skew present in the market. The term "smile" is commonly used to refer to the pattern that results from plotting the implied volatility as a function of the strike. Hence, a better approach would be to use a smile-consistent interpolation method to infer the swaption implied volatilities for non-quoted strikes.

Castagna, Mercurio and Tarenghi suggest in [1] to use the Vanna-Volga approach to value CMS adjustments. This method was first introduced by Lipton and McGhee in [13]. It was mainly used by traders in the FX market, but it was extended to other derivatives markets and we will now apply it to CMS in particular. Although we will not implement it in our thesis, it is worth mentioning that Shkolnikov extended this empirical method further in [9] to include the interest-rate risk.

The Vanna-Volga method is often referred to as the "trader's rule of thumb". It is used to recover the implied volatility smile/skew for a given maturity using only three available quotes from the market. The main advantage of this technique is it is already calibrated to the market since it uses three market quoted volatilities, so it is easy and fast to implement. In addition to that, it produces accurate results that are consistent with the market smile. However, one of the drawbacks of this method is that it is not based on a robust foundation, so we can encounter cases where it may create arbitrage opportunities if it is not dealt with care.

Another way of valuing CMS was suggested by Mercurio and Pavallicini in [5]. They proposed to use the SABR functional form that was first developed in literature by Hagan, Kumar, Lesniewski, and Woodward in [7]. SABR is a stochastic volatility model that attempts to capture the market smile. For our purpose, it is used to get the swaptions implied volatility that are needed to calculate CMS convexity adjustments. The key step before using this model is to have a solid calibration procedure to produce accurate results. When calibrating the SABR model to swaption volatilities, we usually have one degree of freedom left: the parameter β. Practitioners usually fix this parameter to some prior belief about the market. For our purpose, we will calibrate the parameter β to the market CMS swap spreads quotes. Thus, we will use a joint-calibration procedure with swaption volatilities and CMS spreads data. The main advantage of the SABR volatility model is that it produces accurate volatilities that are in line with the market smile. Nevertheless, the calibration procedure should be dealt with care, in contrast with the Vanna-Volga approach that does not require a joint-calibration.

Clearly, each method has a precision/ease of use trade-off. In what follows, we aim to analyse the advantages of each approach, their differences and their ease of implementation and accuracy. To do so, we plan to organise the paper as follows.
We start with the basic definitions of CMS-linked derivatives. We then use the replication method to derive an analytical formula for the convexity adjustment of each of the products. We use bond-math approximations proposed by Hagan in [4] to model the yield curve and achieve the pricing. The first approximation we will base our work on is the standard one used by most practitioners in the literature. We first implement Black's model that assumes constant volatility. To decrease the computation time, we resort to a third-order Taylor Series approximation of the standard function, and we compare the results with the numbers obtained using the exact closed formula. We then want to implement interpolation methods that take into consideration the smile. We start with the Vanna-Volga method that has two approaches, and we provide the reader with a mathematical and a practical [1] justification that the two approaches are consistent. We resume our work by implementing the SABR model and calibrating the parameters to match both the swaption volatility smile and the CMS spreads market data. We try to provide the reader with the clear and detailed steps of the calibration as it is the core part before using the model to price CMS. Additionally, we compare the results obtained with the two interpolation methods and check whether they are consistent with each other.

[1] by application to market data

The main problem with what has previously been done in the literature is that the yield curve is assumed to be flat. However, the structure of CMS payments where LIBOR and spread are exchanged for an N-year CMS rate forces us to be aware of the importance of the shape and slope of the yield curve in the pricing. Hence as a further exploration, we study the impact of changing the usual yield curve approximation that is market practice on the price of a CMS swap. Although the first bond-math approximation is the standard market practice and is used in most of the papers in the literature, it would be interesting to relax the simplistic assumption of flat initial and final yield curve. The two other approximations we will use are also proposed by Hagan in the appendix of [4] but have not been implemented previously. The first one takes into consideration the initial shape of the yield curve instead of assuming it to be flat but only allows parallel shifts. The second one, on the other hand, takes into consideration the slope of the yield curve by allowing non-parallel shifts. Thus, the last part of the thesis will be the inspection of the two new approximations and their impact on CMS prices in specific.

Finally, we shall mention briefly the motivation for doing this project. Requests for CMS structures have been increasing lately. The bank where my internship is based does not have a static model for CMS-linked products. Hence, this project provides a way to benchmark and validate the valuation provided by the EBRD's pricing which is a vendor system.

2 Pricing scheme

The aim of this section is to introduce the important basics of pricing that are needed to value CMS-related products.

We consider the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$. We denote by $S_t^{a,b} = S(t, T_a, T_b)$ the swap-rate at time t of an underlying swap that starts at time T_a and matures at time T_b.

The tenor of the CMS coupon's underlying swap is defined as the difference between the maturity and the start date of the swap. Usually for a N-year indexed constant maturity swap, $T_a = t_{i-1}$ and $T_b = t_{i-1} + N$ years. The swap rate at a fixing time t is defined to be the fixed rate such that the swap's value is zero at time t.

Before providing the reader with the necessary background that is needed for the valuation, we start by introducing the most popular CMS linked products with their payoffs.

2.1 CMS linked products

The three most common CMS linked products are: CMS swaps, CMS caps and CMS floors. We will define each of these products individually in the following subsections. We assume a unit notional amount.

2.1.1 CMS Swap

A CMS swap is a contract involving two legs: a CMS leg and a floating leg. We denote by $t_1, t_2, \ldots, t_m$ the payment dates of the CMS leg.

The CMS leg consists of payments at time t_i of the N-year swap rate $S_{\tau_i}^{t_{i-1}, t_{i-1}+N}$. τ_i is usually spot lag business days before the interval begins or before it ends depending on the type of the contract. In fact, there are two types of contracts. CMS legs can be either set-in-advance or set-in-arrears. Set-in advance CMS legs are the most common in the market, but there are also contracts where the CMS legs are set-in arrears. The definitions below are based on Hagan's paper [4, p.1].

Definition 2.1 (set-in-advance)**.** A CMS leg is set-in-advance if the underlying swap of the swap rate S_i begins at t_{i-1} and ends N years later for each time period i. The fixing date τ_i of S_i is spot lag business days before the interval begins at t_{i-1}.

Definition 2.2 (set-in-arrears)**.** A CMS leg is set-in-arrears if the underlying swap of the swap rate S_i that begins at the end date t_i for each time period i. S_i is fixed on a date τ_i that is spot lag business days before the interval ends at t_i.

Thus, each payment of the CMS leg has the following payoff:

$$\theta_i S_{\tau_i}^{t_{i-1},t_{i-1}+N}$$

where θ_i is the coverage (denoted by cvg in [4]) for each period i. It represents the day count fraction specified by the contract for each period i and is given by:

$$\theta_i := \theta_{t_{i-1},t_i} = cvg(t_{i-1}, t_i, dcb_{sw}) \tag{2.1}$$

and dcb_{sw} is the standard swap day count basis.

The floating leg consists of payments at each time t_i with payoff

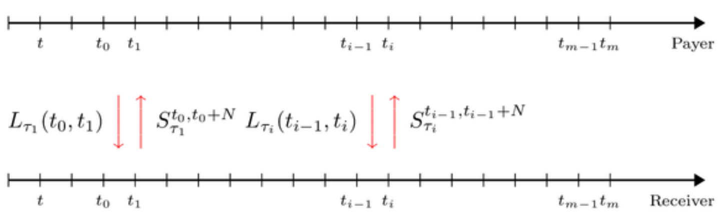
$$\theta_i(L(\tau_i, t_{i-1}, t_i) + m')$$

where $L(\tau_i, t_{i-1}, t_i)$ denotes the forward LIBOR at time τ_i for the interval $[t_{i-1}, t_i)$ and m' is called the CMS spread. Usually, CMS swaps are quoted in the market as the value of the spread m' such that the present value of a CMS swap is zero.

The value of a CMS swap at time t is defined as the difference between the present value of the CMS leg and the present value of the LIBOR leg, namely:

$$PV_{Swap}^{CMS}(t) = \sum_{i=1}^{m} P(t,t_i)\theta_i \mathrm{E}_t^{t_i}\left[S_{\tau_i}^{t_{i-1},t_{i-1}+N}\right] - \sum_{i=1}^{m} P(t,t_i)\theta_i \mathrm{E}_t^{t_i}\left[L(\tau_i, t_{i-1}, t_i) + m\right]$$

We will illustrate the cash flows of a CMS swap to makes it clearer to the reader. For simplicity, we will assume that both legs have the same payment schedules.



We shall define also the level of a swap (also known as the annuity) as it will be used later in the thesis.

Definition 2.3 (level of a swap (annuity))**.** The level of a swap is also known as the acronym DV01 which stands for dollar value of a basis point. It refers to the change in the present value of a swap when there is a shift of the swap curve [2] by one basis point. In our framework where the underlying swap rate is $S_{a,b}$, the annuity is defined as

$$L^{a,b}(t) = \sum_{j=1}^{n} \Theta_j P(t, t_j)$$

where Θ_j is as in (2.1).

2.1.2 CMS caps and floors

A CMS cap is a variation of the interest rate cap in which the rate used is a swap rate with constant maturity instead of the LIBOR rate. It can be seen as a call option on a swap rate where each exchange payment is executed only if it has positive value. As mentioned in [11], a CMS cap is made of a series of caplets just like the usual interest rate vanilla cap,

CMS caps can be useful for hedging positions in the long term. Investors who believe that interest rates will be increasing rapidly and faster than what is predicted by the yield curve are more likely to be buying a CMS cap since it will protect them from rises in the CMS rate above the predetermined level K.

The discounted payoff of a CMS cap with strike K is:

$$\sum_{i=1}^{m} D(t, t_i)\theta_i(S(t, t_{i-1}, t_i) - K)^+$$

.

A CMS floor is a variation of the interest rate floor in which the rate used is a swap rate with a constant maturity. A floor can be seen as a put option on a swap rate where each exchange payment is executed only if it has positive value. [11] Just like a usual interest rate vanilla cap, a CMS cap is made of a series of floorlets.

CMS floors can be useful for hedging positions in the long term, specially for investors who believe that there will be a downward trend in the interest rates and predict that they will go down faster than what is predicted by the yield curve. The discounted payoff of a CMS floor with strike K is:

$$\sum_{i=1}^{m} D(t, t_i)\theta_i(K - S(t, t_{i-1}, t_i))^+$$

[2] We can think of the swap curve as the the name given to the swap's equivalent of a yield curve.

2.2 No arbitrage valuation

While pricing derivatives, one needs to make sure that the market is free of arbitrage opportunities. We require the absence of strategies which, starting from zero initial capital, create a positive cash inflow with positive probability and never creates a loss or negative flow. To be able to talk about the fair value of an option, we must make sure that money cannot be generated freely.

2.2.1 Fundamentals of pricing

Before we proceed with any pricing, we shall state the basic definitions and theorems that will be used in our work. For this purpose, we will refer to [23] and [24] in what follows.

We start by stating the Radon-Nikodym theorem that will allow us to change the probability measure while pricing.

Theorem 2.4 (Radon-Nikodym theorem). *$\mathbb{P}$ and $\mathbb{Q}$ are equivalent probability measures if and only if there exists a positive random variable $Z \in \mathcal{L}^1$ such that $\mathbb{Q}(A) = \mathrm{E}^{\mathbb{P}}[Z\mathbf{1}_A]$. The random variable Z is unique and is denoted by $\frac{d\mathbb{Q}}{d\mathbb{P}}$ and it is called the Radon-Nikodym derivative.*

Before we state the change of numeraire theorem, we shall define the concepts of equivalent martingale measure and numeraire pair.

Definition 2.5 (EMM and numeraire pair). The measure $\mathbb{N}$ defined on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ is an equivalent martingale measure if $\mathbb{N}$ and $\mathbb{P}$ are equivalent and there exists a numeraire N such that the asset price vector normalized by the numeraire is an $\mathcal{F}_t$-martingale under $\mathbb{N}$.

Theorem 2.6 (Change of numeraire theorem). *Assume that the economy is arbitrage-free and admits a numeraire pair $(N, \mathbb{N})$. If $(M, \mathbb{M})$ is another numeraire pair for this economy, then the Radon-Nikodym derivative connecting $\mathbb{M}$ and $\mathbb{N}$ is given by a ratio of their numeraires. For all $t \in [0, T]$,*

$$\left.\frac{d\mathbb{M}}{d\mathbb{N}}\right|_{\mathcal{F}_t} = \frac{M_t}{N_t} \cdot \frac{N_0}{M_0}$$

As a result, arbitrage-free prices are invariant by change of numeraire. For instance, Suppose A is an $\mathcal{F}_T$ measurable random variable representing a payoff at maturity T. Then its arbitrage-free price at time t is

$$M_0\mathbb{E}_t^{\mathbb{M}}\left[\frac{A}{M_T}\right] = M_0\mathbb{E}_t^{\mathbb{N}}\left[\frac{A}{M_T}\left.\frac{d\mathbb{M}}{d\mathbb{N}}\right|_{\mathcal{F}_T}\right] = M_0\mathbb{E}_t^{\mathbb{N}}\left[\frac{A}{M_T}\frac{M_T}{N_T}\cdot\frac{N_0}{M_0}\right] = N_0\mathbb{E}_t^{\mathbb{N}}\left[\frac{A}{N_T}\right]$$

2.2.2 No arbitrage valuation applied to CMS

We fix a maturity T_a and a set of time $\Gamma_{a,b} = [T_{a+1}, \dots, T_b]$ with a strictly positive constant year fraction denoted by τ_1.

We consider a CMS linked product coupon where the CMS leg pays $g(S_{a,b}(T_a))$ at time $T_a + \delta$, δ denoting the accrual period. We want to find the value of the coupon's CMS leg at time 0. We start by taking the risk neutral expectation of the discounted payoff.

$$\begin{aligned} V_0 &= \mathrm{E}^{\mathbb{Q}}[D(0, T_a + \delta)g(S_{a,b}(T_a)] \\ &= P(0, T_a + \delta)\mathrm{E}^{T_a+\delta}\left[g(S_{a,b}(T_a))\right] \end{aligned}$$

We used the change of numeraire theorem (2.6) to get the second equality. The $T_a + \delta$ forward measure is associated with the numeraire $P(\cdot, T_a + \delta)$.

We shall first take the case of a CMS swap where $g(S_{a,b}(T_a)) = S_{a,b}(T_a)$.

$S_{a,b}(T_a)$ is not a martingale under the $T_a + \delta$ forward measure. The measure under which the forward swap rate $S_{a,b}(T_a)$ is a martingale is called the annuity measure denoted by $\mathbb{Q}^{a,b}$, with the annuity of the underlying swap as numeraire. Since we are computing the expectation of the CMS rate under a forward measure that differs from its natural martingale measure, this will lead us to define a convexity adjustment.

The term convexity in finance often refers to any non-linear behavior of the price as a function of the interest rate. From a mathematical and probability theory perspective, the convexity adjustment emerges from Jensen equality which states that the expected value of a convex function is greater than or equal to the function of the expected value. From a modeling perspective, it arises when the underlying asset is not a martingale under the probability measure used in the pricing.

The convexity adjustment of a CMS swap at time 0 is defined as follows:

$$CA := \mathrm{E}^{T_a+\delta}[S_{a,b}(T_a)] - S_{a,b}(0).$$

$$\begin{aligned}
\mathrm{E}^{T_a+\delta}[S_{a,b}(T_a)] &= \mathrm{E}^{T_a+\delta}\left[\frac{P(0,T_a+\delta)}{P(T_a,T_a+\delta)}\frac{P(T_a,T_a+\delta)}{P(0,T_a+\delta)}S_{a,b}(T_a)\right] \\
&= \mathrm{E}^{a,b}\left[\frac{L(0)}{L(T_a)}\frac{P(T_a,T_a+\delta)}{P(0,T_a+\delta)}S_{a,b}(T_a)\right] \\
&= \frac{\sum_{i=a+1}^{b}\theta_i P(0,t_i)}{P(0,T_a+\delta)}\mathrm{E}^{a,b}\left[\frac{P(T_a,T_a+\delta)}{\sum_{i=a+1}^{b}\theta_i P(T_a,t_i)}S_{a,b}(T_a)\right]
\end{aligned}$$

The second equality uses the change to the annuity measure $\mathbb{Q}^{a,b}$ with annuity L [3] as numeraire. Multiplying by $P(0, T_a + \delta)$ and using the tower property of the conditional expectation we get:

$$\begin{aligned}
V_0 &= P(0,T_a+\delta)\frac{L(0)}{P(0,T_a+\delta)}\mathrm{E}^{a,b}\left[\frac{P(T_a,T_a+\delta)}{L(T_a)}S_{a,b}(T_a)\right] \\
&= L(0)\mathrm{E}^{a,b}\left[\mathrm{E}^{a,b}\left(\frac{P(T_a,T_a+\delta)}{L(T_a)}\middle| S_{a,b}(T_a)\right)S_{a,b}(T_a)\right] \\
&= L(0)\mathrm{E}^{a,b}[\bar{f}(S_{a,b}(T_a))S_{a,b}(T_a)]
\end{aligned}$$

where, for any t ≥ 0 :

$$\bar{f}(S_{a,b}(t)) := \mathrm{E}^{a,b}\left[\frac{P(t,T_a+\delta)}{L(t)}\middle| S_{a,b}(t)\right] \tag{2.2}$$

$\bar{f}$ is called the annuity mapping function as is suggested in [20]. to find an approximation of the conditional expectation in (2.2), Hagan proposes in [4] to model the yield curve in a way that enables us to rewrite the level of the swap $L(t)$ and the zero-coupon bond $P(t, T_a+\delta)$ as a function of the swap rate $S_{a,b}(t)$. We will provide the details of the yield curve modeling in Section 6.1.

At time t=0, $S_{a,b}(0)$ is known so:

$$\bar{f}(S_{a,b}(0)) := \mathrm{E}^{a,b}\left[\frac{P(0,T_a+\delta)}{L(0)}\middle| S_{a,b}(0)\right] = \frac{P(0,T_a+\delta)}{L(0)}$$

and thus, we get:

$$\mathrm{E}^{T_a+\delta}[(S_{a,b}(T_a)] \approx \frac{1}{\bar{f}(S_{a,b}(0))}\mathrm{E}^{a,b}[\bar{f}(S_{a,b}(T_a))S_{a,b}(T_a)] \tag{2.3}$$

where the standard market practice annuity mapping is given by:

$$\bar{f}(x) := \frac{1}{G_{a,b}(x)(1+\tau_1 x)^{\frac{\delta}{\tau_1}}}, \tag{2.4}$$

$$G_{a,b}(x) := \sum_{j=1}^{b-a}\frac{\tau_1}{(1+\tau_1 x)^j} = \begin{cases} \frac{1}{x}\left[1-\frac{1}{(1+\tau_1 x)^{b-a}}\right] & \text{if } x>0 \\ \tau_1(b-a) & \text{if } x=0 \end{cases}$$

$\frac{\delta}{\tau_1}$ is the fraction of the period between the swap's start date and the pay date.

This model assumes that the initial and final yield curves are flat and that the coverages for the swaptions are approximately equal to the schedule.

[3] We write L(t) to represent $L^{a,b}(t)$ to simplify the notations

As mentioned above, the modeling details that lead to this function will be discussed in Section 6. For the first part of the thesis, we will do our implementation using the function in (2.4) and the second part of the thesis will study the impact of changing this approximation on the price of CMS.

3 Pricing CMS by replication

The value of each of the CMS swap, CMS cap and CMS floor is the sum of the values of each discounted payment. For CMS caplets, floorlets, and swaplets, the payoff $g(S_\tau^{a,b})$ specializes respectively to:

$$g(S_\tau^{a,b}) = (S_\tau^{a,b} - K)^+$$

$$g(S_\tau^{a,b}) = (K - S_\tau^{a,b})^+$$

$$g(S_\tau^{a,b}) = S_\tau^{a,b}$$

The purpose of pricing by replication is to replicate European options with payoff $f(\cdot)$ using Call and Put vanilla options. The main theorem behind pricing by replication is a very well-known result derived by Carr-Madan in [19].

3.1 Carr-Madan formulae

Theorem 3.1 (Carr-Madan formulae). *Let $f : \mathbb{R} \to \mathbb{R}$ be a $\mathcal{C}^2$ function. Then $\mathbb{P} - a.s$,*

$$f(x) = f(a) + f'(a)(x-a) + \int_0^a (v-x)^+ f''(v)dv + \int_a^\infty (x-v)^+ f''(v)dv \tag{3.1}$$

where x is strictly positive and a is a positive constant.

In the special case where $a = 0$, we have $\mathbb{P} - a.s$

$$f(x) = f(0) + f'(0)x + \int_0^\infty (x-v)^+ f''(v)dv \tag{3.2}$$

Proof. The proof follows closely the original paper [19].

By the Fundamental Theorem of Calculus,

$$\begin{aligned}
f(x) &= f(a) + \mathbf{1}_{\{x>a\}} \int_a^x f'(u)du - \mathbf{1}_{\{x<a\}} \int_x^a f'(u)du \\
&= f(a) + \mathbf{1}_{\{x>a\}} \int_a^x \left[f'(a) + \int_a^u f''(v)dv \right] du - \mathbf{1}_{\{x<a\}} \int_x^a \left[f'(a) - \int_u^a f''(v)dv \right] du \\
&= f(a) + f'(a)(x-a) + \mathbf{1}_{\{x>a\}} \int_a^x \int_v^x f''(v)dudv + \mathbf{1}_{\{x<a\}} \int_x^a \int_x^v f''(v)dudv \\
&= f(a) + f'(a)(x-a) + \mathbf{1}_{\{x>a\}} \int_a^x f''(v)(x-v)dv + \mathbf{1}_{\{x<a\}} \int_x^a f''(v)(v-x)dv \\
&= f(a) + f'(a)(x-a) + \int_a^\infty (x-v)^+ f''(v)dv + \int_0^a (v-x)^+ f''(v)du
\end{aligned}$$

The 3^{rd} equality uses Fubini's theorem and that $f'(a)$ does not depend on u. □

We will see that this result is useful from a hedging point of view. It provides us with a way of replicating a European payoff $f(\cdot)$ that is $\mathcal{C}^2$ by using a continuum of Call and Put Options for a given maturity.

We shall apply the replication technique to CMS linked products in what follows.

3.2 Pricing by replication of a CMS swaplet

The value of the CMS swap is just the sum of the values of each discounted payment. We start by valuing a single coupon at time t. We consider $\tau > t$ to be the fixing date of the swap rate.

$$V^{CMS}_{swaplet}(t) \approx L(t) \mathrm{E}^{a,b}_t \left[\bar{f}(S^{a,b}_\tau) S^{a,b}_\tau \right]$$

where $\bar{f}$ is given by (2.4).

We apply Carr-Madan formulae to the function $h(x) = x\bar{f}(x)$ which is $\mathcal{C}^2$ since it is the product of $\mathcal{C}^2$ functions [4]. Indeed, in our modeling it is easy to see that $\bar{f}$ given by (2.4) is $\mathcal{C}^2$ by noticing that it is the composition of $\mathcal{C}^2$ functions on the domain $[0, \infty)$.

A straightforward computation of the first and second order derivatives gives: $h'(x) = x\bar{f}'(x) + \bar{f}(x)$, and $h''(x) = x\bar{f}''(x) + 2\bar{f}'(x)$.

By a direct application of Carr-Madan formulae (3.2) at $x = S^{a,b}_\tau$, we get:

$$\begin{aligned}
h(S^{a,b}_\tau) &= h(0) + h'(0) S^{a,b}_\tau + \int_0^\infty (S^{a,b}_\tau - K)^+ h''(K) dK \\
&= \bar{f}(0) S^{a,b}_\tau + \int_0^\infty \left(2\bar{f}'(K) + \bar{f}''(K)K \right) (S^{a,b}_\tau - K)^+ dK
\end{aligned}$$

[4] x is a polynomial so it is $\mathcal{C}^2$

$$
\begin{aligned}
V_{swaplet}^{CMS}(t) &\approx L(t)\mathrm{E}_t^{a,b}\left[h(s_\tau^{a,b})\right] \\
&= L(t)\left(\mathrm{E}_t^{a,b}\left[\bar{f}(0)S_\tau^{a,b}\right] + \int_0^\infty \left(2\bar{f}'(K) + \bar{f}''(K)K\right)\mathrm{E}_t^{a,b}\left[(S_\tau^{a,b} - K)^+\right] dK\right)
\end{aligned}
$$

Now using martingale pricing theory, $S_\tau^{a,b}$ is a martingale under the $\mathbb{Q}^{a,b}$ probability measure. Hence, we will use the following:

$$
\mathrm{E}_t^{a,b}\left[S_\tau^{a,b}\right] = S_t^{a,b}
$$

$$
\mathrm{E}_t^{a,b}\left[(S_\tau^{a,b} - K)^+\right] = C_{a,b}(K) \tag{3.3}
$$

where $C_{a,b}(K)$ denotes the time t- price of a Payer swaption with strike K, divided by its annuity term. Namely, it is a Call option on the swap rate $S_{a,b}$, divided by $L^{a,b}(t)$.

Hence, we get:

$$
V_{swaplet}^{CMS}(t) \approx L(t)\left[\bar{f}(0)S_t^{a,b} + \int_0^\infty C_{a,b}(K)\left(K\bar{f}''(K) + 2\bar{f}'(K)\right) dK\right]
$$

To interpret the idea of "replication" behind this formula, we will discretise it.

$$
V_{swaplet}^{CMS}(t) \approx L(t)\left[\bar{f}(0)S_t^{a,b} + \lim_{n\to\infty}\sum_{i=1}^{n} C_{a,b}(K_i)(K_i\bar{f}''(K_i) + 2\bar{f}'(K_i))\Delta K_i\right]
$$

It would be interesting to interpret the result we just got from a hedging point of view to show the particularity of the replication method. "Ideally", to hedge the CMS position above, we buy a contract that pays a floating coupon (with LIBOR as the reference rate) with notional $\bar{f}(0)$ and frequency δ, and we buy Call swaptions with strikes K_i and notional $\left(K_i\bar{f}''(K_i) + 2\bar{f}'(K_i)\right)$ for i ranging from 1 to infinity and ΔK_i small enough. This is a very interesting aspect of the replication method that is specially useful for sell-side traders. Indeed, in the ideal case where vanilla swaptions quotes are provided for every strike, it would provide them with a hedge against their positions in the market.

Pricing by replication is considered to be very accurate as it is consistent with the prices of European options in the market. However, it is very computation intensive because as we can see, it requires to compute an integral to infinity.

3.3 Pricing by replication of CMS Caplets and CMS floorlets

In the caplet case, the payoff is not differentiable, so we cannot apply Carr-Madan formulae directly. We start by writing the payoff $g(S_\tau^{a,b}) = (S_\tau^{a,b} - K)^+ = (S_\tau^{a,b} - K)\mathbf{1}_{\{S_\tau^{a,b} > K\}}$

$$\begin{aligned} V_{caplet}^{CMS}(t) &\approx L(t)\mathrm{E}_t^{a,b}[\bar{f}(S_\tau^{a,b})g(S_\tau^{a,b})] \\ &= L(t)\mathrm{E}_t^{a,b}[\bar{f}(S_\tau^{a,b})(S_\tau^{a,b} - K)\mathbf{1}_{\{S_\tau^{a,b} > K\}}] \end{aligned}$$

By letting $h(x) = \bar{f}(x)(x - K)$, we get by following [4, p.4]:

$$h(S_\tau^{a,b})\mathbf{1}_{\{S_\tau^{a,b} > K\}} = h'(K)(S_\tau^{a,b} - K)^+ + \int_K^\infty (S_\tau - x)^+ h''(x)dx \tag{3.4}$$

with $h'(x) = \bar{f}(x) + \bar{f}'(x)(x - K)$ and $h''(x) = 2\bar{f}'(x) + \bar{f}''(x)(x - K)$

$$\begin{aligned} V_{caplet}^{CMS}(t) &\approx L(t)\mathrm{E}_t^{a,b}\left[\bar{f}(K)(S_\tau^{a,b} - K)^+ + \bar{f}'(K)(K - K)(S_\tau^{a,b} - K)^+\right] \\ &+ L(t)\mathrm{E}_t^{a,b}\left[\int_K^\infty (S_\tau^{a,b} - x)^+ \left(2\bar{f}'(x) + \bar{f}''(x)(x - K)\right) dx\right] \\ &= L(t)C_{a,b}(K)\bar{f}(K) + L(t)\int_K^\infty \mathrm{E}_t^{a,b}\left[\int_K^\infty (S_\tau^{a,b} - x)^+ \left(2\bar{f}'(x) + \bar{f}''(x)(x - K)\right)\right] dx \end{aligned}$$

The second equality is done by using (3.3) and Fubini's theorem. Hence,

$$V_{caplet}^{CMS}(t) \approx L(t)\left(C_{a,b}(K)\bar{f}(K) + \int_K^\infty C_{a,b}(x)(2\bar{f}'(x) + \bar{f}''(x)(x - K))dx\right) \tag{3.5}$$

The pricing of the CMS floorlet is very similar to the pricing of the CMS caplet done previously. The payoff is

$$g(S_\tau^{a,b}) = (K - S_\tau^{a,b})^+ = (K - S_\tau^{a,b})\mathbf{1}_{\{S_\tau^{a,b} < K\}}$$

The only difference is in step (3.4) above where we now use instead:

$$h(S_\tau^{a,b})\mathbb{1}_{\{S_\tau^{a,b} < K\}} = -h'(S_\tau^{a,b})(K - S_\tau^{a,b})^+ + \int_{-\infty}^K (x - S_\tau^{a,b})^+ h''(x)dx$$

Completing the computations, we get:

$$V_{floorlet}^{CMS}(t) \approx L(t)\left(\bar{f}(K)P_{a,b}(K) - \int_{-\infty}^K P_{a,b}(x)\left(2\bar{f}'(x) + \bar{f}''(x)(x - K)\right) dx\right)$$

where we used the following:

$$\mathrm{E}_t^{a,b}\left[(K - S_\tau^{a,b})^+\right] = P_{a,b}(K)$$

and $P_{a,b}(K)$ is the time t- price of a receiver swaption with strike K, divided by the annuity at time t. Namely, it is a Put option on the swap rate $S_{a,b}$ divided by L(t).

4 Pricing under Black's dynamics

As we can see, in all three cases of a CMS swap, cap or floor, we have integrals going to infinity. If we had all the possible strikes with their corresponding volatilities quoted in the market, then the pricing would be directly and easily achieved. However, this is not the case in reality. The most liquid swaption volatilities quoted in the market are the at-the-money swaptions. Apart from these, we can find few quotes for in-the-money and out-of-the money strikes. This being said, we need an interpolation method to continue the pricing of CMS-linked products.

As a first approach to achieve the pricing, Hagan proposes to use Black's model. This model is characterized by holding all the swaption volatilities constant, setting them equal to the at-the-money Black's implied volatility σ_{ATM}.

4.1 Black's model exact form

We start with the valuation of a CMS swap at time t=0 where $g(S_{a,b}(T_a)) = S_{a,b}(T_a)$.
Following the pricing by replication done in (2.3), we have:

$$\Pi := \mathrm{E}^{T_a+\delta}[S_{a,b}(T_a)] \approx \frac{1}{\bar{f}(S_{a,b}(0))}\mathrm{E}^{a,b}\left[\bar{f}(S_{a,b}(T_a)S_{a,b}(T_a)\right]$$

where

$$E^{a,b}\left[\bar{f}(S_{a,b}(T_a)S_{a,b}(T_a)\right] = \frac{V^{CMS}_{swaplet}(0)}{L(0)}$$

Hence,

$$\Pi \approx \frac{1}{\bar{f}(S_{a,b}(0))}\left[\bar{f}(0)S_{a,b}(0) + \int_0^\infty \left(x\bar{f}''(x) + 2\bar{f}'(x)\right)C_{a,b}(x)dx\right] \tag{4.1}$$

We plug in Black's (Call) Swaption formula:

$$C^{BS}_{a,b}(K) := Bl(K, S_{a,b}(0), \sigma_{ATM}\sqrt{T_a})$$

where

$$Bl(K, S, v) := S\Phi(d_1(K)) - K\Phi(d_2(K))$$
$$d_1(K) = \frac{\log(S/K) + v^2/2}{v}$$
$$d_2(K) = \frac{\log(S/K) - v^2/2}{v}$$

$\Phi(\cdot)$ denotes the standard Normal cumulative distribution. Hence, we get:

$$\Pi^{BS} \approx \frac{1}{\bar{f}(S_{a,b}(0))} \left[\bar{f}(0) S_{a,b}(0) + \int_0^\infty \left(x\bar{f}''(x) + 2\bar{f}'(x) \right) C_{a,b}^{BS}(x) dx \right] \tag{4.2}$$

4.2 Black's model with approximation

When implementing the integral in (4.2), the computation time is high. So to speed up the pricing, it is proposed in the literature to approximate the function $\bar{f}$ using its Taylor expansion around the known value $\bar{f}(S_{a,b}(0))$.

The n^{th} order Taylor Series approximation of $\bar{f}$ is

$$\bar{f}(x) \approx \tilde{f}_n(x) := \sum_{i=0}^{n} \frac{\bar{f}^{(i)}(S_{a,b}(0))}{i!} [x - S_{a,b}(0)]^i \tag{4.3}$$

A typical choice of n would be n=3 because it gives an accurate approximation very close to $\bar{f}$. The graph below shows that $\bar{f}$ and $\tilde{f}_3$ are indeed very close.

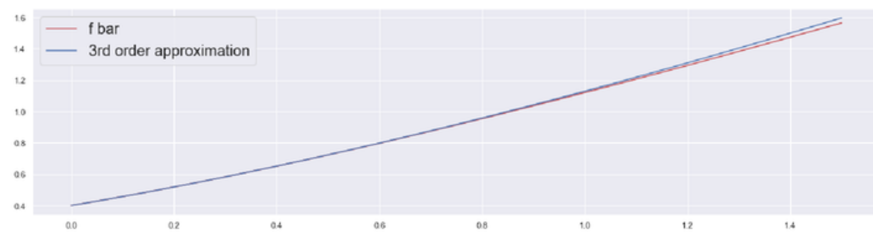


Figure 1: Comparison between $\bar{f}$ and its 3^{rd} order approximation

We need to compute the first, second and third order derivatives of $\bar{f}$. The calculations are already done in [1, p.9] and lead to:

$$\bar{f}'(x) = A(x)\bar{f}(x)$$
$$\bar{f}''(x) = B(x)\bar{f}(x)$$
$$\bar{f}'''(x) = C(x)\bar{f}(x)$$
$$A(x) = \frac{\theta(x)}{x}$$
$$B(x) = \frac{\theta^2(x) - \gamma(x)}{x^2}$$
$$C(x) = \frac{\theta^3(x) - \theta(x)\gamma(x) + 2\eta(x)}{x^3}$$

where setting $n := b - a$ and $T(x) := (1+\tau_1 x)^n - 1$

$$\theta(x) := 1 - \frac{\tau_1 x}{1+\tau_1 x}\left(\frac{\delta}{\tau_1} + \frac{n}{T(x)}\right)$$
$$\gamma(x) := 1 - \frac{(\tau_1 x)^2}{(1+\tau_1 x)^2}\left(\frac{\delta}{\tau} + \frac{n+n^2}{T(x)} + \frac{n^2}{T^2(x)}\right)$$
$$\eta(x) := 1 - \frac{(\tau x)^3}{(1+\tau_1 x)^3}\left(\frac{\delta}{\tau_1} + \frac{n+1.5n^2+0.5n^3}{T(x)} + \frac{1.5(n^2+n^3)}{T^2(x)} + \frac{n^3}{T^3(x)}\right)$$

Once we get the derivatives, we plug them in (4.3) to get $\tilde{f}_3$. Then, we plug $\tilde{f}_3$ in the integral (4.2) to get the final results of the approximation in this setting.

Another way to calculate the price under Black's framework using the approximation function $\tilde{f}_3$ is to evaluate the following integral that uses the Breedon-Litzenberger formulaes as suggested in [10]:

$$\Pi^{BS} \approx \frac{1}{\bar{f}(S_{a,b}(0))}\int_0^\infty \tilde{f}_3(x) x \frac{d^2 C_{a,b}^{BS}(x)}{dx^2} dx \tag{4.4}$$

where

$$\frac{d^2 C_{a,b}^{BS}(x)}{dx^2} = \frac{1}{x\sigma_{ATM}}\sqrt{T_a}\phi\left(\frac{\log(\frac{S_{a,b}(0)}{x}) - \frac{1}{2}\sigma_{ATM}^2 T_a}{\sigma_{ATM}\sqrt{T_a}}\right).$$

The way to look at (4.4) is as follows:

$$C_{a,b}(K) = \mathrm{E}_t^{a,b}\left[(S_\tau^{a,b} - K)^+\right] = \int_K^\infty (x-K)\phi(\tau,x)$$

where $\phi(\tau,\cdot)$ is the distribution function of the random variable $S_\tau^{a,b}$ under the $\mathbb{Q}^{a,b}$-measure. Taking the first derivative with respect to K we get:

$$\frac{\partial C^{a,b}(K)}{\partial K} = 0 - \int_K^\infty \phi(\tau,x)dx$$

Taking again the second derivative with respect to K and using the fundamental theorem of calculus we get:

$$\frac{\partial^2 C^{a,b}(K)}{\partial K^2} = \phi(\tau,x)$$

Hence we can now use that:

$$\mathrm{E}^{a,b}\left[x\bar{f}(x)\right] \approx \mathrm{E}^{a,b}\left[x\tilde{f}_3(x)\right] = \int_0^\infty x\tilde{f}_3(x)\frac{d^2 C_{a,b}^{BS}(x)}{dx^2}dx$$

Computing the integral (4.4) is equivalent to the computation of the first four moments [5] of a log

[5] The k^{th} moment of a random variable X is defined to be $\int_0^\infty x^k\phi(x)dx$ where $\phi(\cdot)$ is the distribution function of x if it exists

normal density. Indeed, using that the integral of a sum of functions is equal to the sum of their integrals, we get:

$$\begin{aligned}\Pi^{BS} &\approx \frac{1}{\bar{f}(S_{a,b}(0))}\int_0^{\infty}\sum_{i=0}^{3}\frac{f^{(i)}(S_{a,b}(0))}{i!}[x-S_{a,b}(0)]^i x\frac{d^2C^{BS}_{a,b}(x)}{dx^2}dx\\ &= \frac{1}{\bar{f}(S_{a,b}(0))}\sum_{i=0}^{3}\int_0^{\infty}\frac{f^{(i)}(S_{a,b}(0))}{i!}[x-S_{a,b}(0)]^i\frac{1}{\sigma_{ATM}}\sqrt{T_a}\phi\left(\frac{\log(\frac{S_{a,b}(0)}{x})-\frac{1}{2}\sigma^2_{ATM}T_a}{\sigma_{ATM}\sqrt{T_a}}\right)\end{aligned}$$

The calculations of the moments is lengthy but straightforward so we will not include them in the thesis. This leads us to the final result presented in [1, p.10]:

$$\begin{aligned}\Pi^{H,BS} \quad &\approx \tfrac{1}{\bar{f}(S_{a,b}(0))}\Big[S_{a,b}(0)\Big(1-A\left(S_{a,b}(0)\right)S_{a,b}(0)+0.5B\left(S_{a,b}(0)\right)S^2_{a,b}(0)-\tfrac{1}{6}C\left(S_{a,b}(0)\right)S^3_{a,b}(0)\Big)\Big]\\ &+2A\left(S_{a,b}(0)\right)I_1+B\left(S_{a,b}(0)\right)\left(3I_2-2S_{a,b}(0)I_1\right)+C\left(S_{a,b}(0)\right)\Big(2I_3-3S_{a,b}(0)I_2+S^2_{a,b}(0)I_1\Big)\end{aligned} \tag{4.5}$$

where

$$\begin{aligned}I_1 &= \frac{1}{2}S^2_{a,b}(0)e^{\sigma^2T_a}\\ I_2 &= \frac{1}{6}S^3_{a,b}(0)e^{3\sigma^2T_a}\\ I_3 &= \frac{1}{12}S^4_{a,b}(0)e^{6\sigma^2T_a}\end{aligned}$$

4.3 Results and discussion

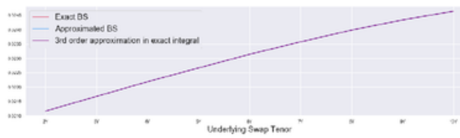
To test the formulas above, we need to apply them to real market data. The main goal is to examine how good the approximation is. We do not want to use any approximation just for the purpose of decreasing the computation complexity. We should also make sure that the results are accurate and consistent with the closed form. We consider 5 different one-look USD CMS coupons [6]: CMS-2Y, CMS-5Y, CMS-10Y, CMS-20Y and CMS-30Y. We also consider 9 different underlying swap tenors for each one of them, ranging from 2Y to 10Y. The underlying swap of the CMS leg has quarterly payments as per market convention for USD. The market data we will use is as of May-20-2019 [7].

To make sure that $\tilde{f}_3$ is indeed a good approximation of $\bar{f}$, we need to check that plugging $\tilde{f}_3$ or plugging $\bar{f}$ in (4.2) give very similar results. Another thing to check is that plugging $\tilde{f}_3$ in (4.2)

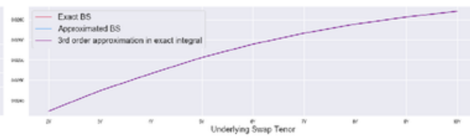
[6] We are considering same currency CMS

[7] All data used is from the EBRD database

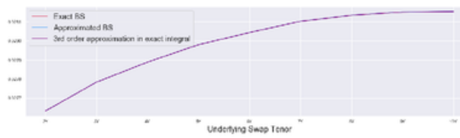
and using the formula given by (4.5) lead exactly to the same values since both are using the 3^{rd} order approximation of $\bar{f}$. We illustrate the results obtained in the following graphs.



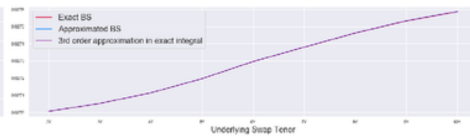
(a) 2Y-CMS coupon.



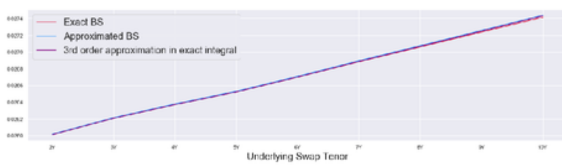
(b) 5Y-CMS coupon.



(c) 10Y-CMS coupon.



(d) 20Y-CMS coupon.



(e) 30Y-CMS coupon.

Figure 2: Comparison of Π^{BS} exact and Π^{BS} approximated

The blue curve represents the implementation of (4.5). The purple curve represents the implementation of the integral (4.2) after plugging $\tilde{f}_3$ and the red curve represents the same integral but with exact $\bar{f}$. We can see that the three curves coincide for a given CMS coupon across all different underlying swap tenors which is exactly as desired. This shows that using a 3^{rd} order Taylor Series approximation gives accurate results, while increasing the computation time considerably.

However, as discussed earlier, calculating Π^{BS} does not take into consideration the market smile for swaption volatilities. The reason behind the smile is that in usual market conditions, options that are out-of-the money and in-the-money are riskier than the at-the-money options. The risk premium is what causes to see a smile in the implied volatility plot as a function of the strike to appear. This being said, we need to consider models that are consistent with the smile. We will introduce in the next chapter two smile-consistent interpolation methods and use them to price CMS swaps and caps.

5 Interpolation methods and smile

5.1 Vanna-Volga

Vanna-Volga is an interpolation method that produces option prices in a market where three basic options quotes are available for a given maturity T. This method is mostly useful in cases where practitioners need to have the value of an option that is not liquid (for example, the strike could be very in-the-money or very out-of-the money). The problem with non-liquid options is that we can hardly find them quoted in the market. Hence we need to interpolate the data that is already present to get a value for these options.

5.1.1 First approach

We consider a payer swaption with strike x and maturity T_a where the underlying is the swap rate $S_{a,b}$. to use Vanna-Volga, we must have three available quotes for the swaption volatilities. The three strikes we need are K_i and their corresponding implied volatilities are σ_i for $i = 1, 2, 3$. K_2 is the ATM strike which is equal to the forward swap rate at time 0. K_1 and K_3 are two away-from-the-money strikes such that $K_1 < K_2 < K_3$. We set $\sigma = \sigma_2 = \sigma_{ATM}$. The Vanna-Volga price of a payer swaption with strike x and maturity T_a (divided by the annuity) is defined in [1, p.5] as:

$$C_{a,b}^{VV}(x) = C_{a,b}^{BS}(x) + \sum_{i=1}^{3} w_i(x)[C_{a,b}^{MKT}(K_i) - C_{a,b}^{BS}(K_i)] \tag{5.1}$$

where :

$$C_{a,b}^{BS}(x) = Bl(x, S_{a,b}(0), \sigma_{ATM}\sqrt{T_a})$$

$$C_{a,b}^{MKT}(K_i) = Bl(K_i, S_{a,b}(0), \sigma_i\sqrt{T_a})$$

$$Bl(K, S, v) := S\Phi(d_1(K)) - K\Phi(d_2(K))$$

$$d_1(K) = \frac{\log(S/K) + \frac{1}{2}v^2}{v}$$

$$d_2(K) = \frac{\log(S/K) - \frac{1}{2}v^2}{v}$$

The weights $w_i(x)$ are the solution of the following system:

$$\begin{cases} \frac{\partial C_{a,b}^{BS}(x)}{\partial \sigma} = \sum_{i=1}^{3} w_i(x) \frac{\partial C_{a,b}^{BS}(K_i)}{\partial \sigma} \\ \frac{\partial^2 C_{a,b}^{BS}(x)}{\partial \sigma^2} = \sum_{i=1}^{3} w_i(x) \frac{\partial^2 C_{a,b}^{BS}(K_i)}{\partial \sigma^2} \\ \frac{\partial^2 C_{a,b}^{BS}(K)}{\partial \sigma \partial S_{a,b}(0)} = \sum_{i=1}^{3} w_i(x) \frac{\partial^2 C_{a,b}^{BS}(K_i)}{\partial \sigma \partial S_{a,b}(0)} \end{cases}$$

The idea behind the method described above is to construct a portfolio made of $w_i(K)$ units of payer swaptions with strikes K_i. This portfolio is created such that it creates a perfect hedge to the payer swaption with strike K under the Black and Scholes settings. In fact, we are equating the Vega , Volga and Vanna of the portfolio created respectively with the Vega, Volga and Vanna of the payer swaption with strike K.

Vega represents the sensitivity of the price of an option with respect to a change in the implied volatility of the underlying asset. Its closed form [8] is:

$$\mathcal{V}(x) := \frac{\partial C_{a,b}^{BS}(x)}{\partial \sigma} = S_{a,b}(0)\sqrt{T_a}\phi(d_1(x))$$

Volga is the sensitivity of the Vega with respect to a change of the implied volatility . Volga has the following closed formula:

$$\frac{\partial^2 C_{a,b}^{BS}(x)}{\partial \sigma^2} = \frac{\mathcal{V}(x)}{\sigma} d_1(x) d_2(x)$$

Vanna represents the amount the Vega changes in reaction to a percentage change in S. The closed form is as follows:

$$\frac{\partial^2 C_{a,b}^{BS}(x)}{\partial \sigma \partial S_{a,b}(0)} = \frac{-\mathcal{V}(x)}{S_{a,b}(0)\sigma\sqrt{T_a}} d_2(x)$$

The figure below shows a comparison between the Black-Scholes Call price and the Vanna-Volga Call price as a function of the strikes.

[8] The closed form of the 3 Black-Scholes Greeks for a payer swaption with strike x at time 0 are given divided by the annuity

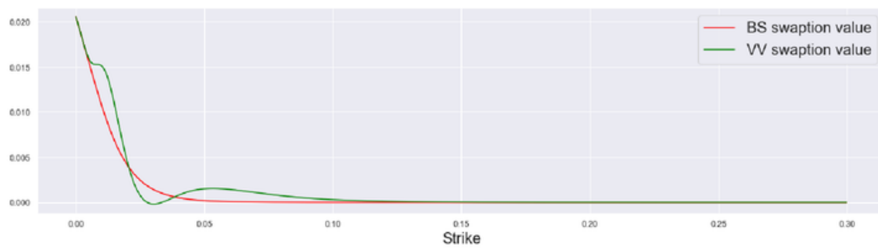


Figure 3: Black-Scholes and Vanna-Volga European swaption

We can interpret the difference between the two prices shown in figure 3 as the cost of the portfolio that hedges the risks associated with the greeks described above. This discrepancy is called the "smile cost" since it is the difference between the price computed holding the volatility constant and the price that includes the volatility smile.

Proposition 5.1. *The weights $w_i(x)$ are given by:*

$$\begin{cases} w_1(x) = \frac{\mathcal{V}(x)}{\mathcal{V}(K_1)} \frac{\log \frac{K_2}{x} \log \frac{K_3}{x}}{\log \frac{K_2}{K_1} \log \frac{K_3}{K_1}} \\ w_2(x) = \frac{\mathcal{V}(x)}{\mathcal{V}(K_2)} \frac{\log \frac{x}{K_1} \log \frac{K_3}{x}}{\log \frac{K_2}{K_1} \log \frac{K_3}{K_2}} \\ w_3(x) = \frac{\mathcal{V}(x)}{\mathcal{V}(K_3)} \frac{\log \frac{x}{K_1} \log \frac{x}{K_2}}{\log \frac{K_3}{K_1} \log \frac{K_3}{K_2}} \end{cases}$$

where

$$\mathcal{V}(x) := \frac{\partial C_{a,b}^{BS}(x)}{\partial \sigma} = S_{a,b}(0)\sqrt{T_a}\phi(d_1(x))$$

Proof. $$\begin{pmatrix} \frac{\partial C_{a,b}^{BS}(x)}{\partial \sigma} \\ \frac{\partial^2 C_{a,b}^{BS}(x)}{\partial \sigma^2} \\ \frac{\partial^2 C_{a,b}^{BS}(K)}{\partial \sigma \partial S_{a,b}(0)} \end{pmatrix} = \begin{pmatrix} \frac{\partial C_{a,b}^{BS}(K_1)}{\partial \sigma} & \frac{\partial C_{a,b}^{BS}(K_2)}{\partial \sigma} & \frac{\partial C_{a,b}^{BS}(K_3)}{\partial \sigma} \\ \frac{\partial^2 C_{a,b}^{BS}(K_1)}{\partial \sigma^2} & \frac{\partial^2 C_{a,b}^{BS}(K_2)}{\partial \sigma^2} & \frac{\partial^2 C_{a,b}^{BS}(K_3)}{\partial \sigma^2} \\ \frac{\partial^2 C_{a,b}^{BS}(K_1)}{\partial \sigma \partial S_{a,b}(0)} & \frac{\partial^2 C_{a,b}^{BS}(K_2)}{\partial \sigma \partial S_{a,b}(0)} & \frac{\partial^2 C_{a,b}^{BS}(K_3)}{\partial \sigma \partial S_{a,b}(0)} \end{pmatrix} \begin{pmatrix} w_1(x) \\ w_2(x) \\ w_3(x) \end{pmatrix}$$

We are solving a system of equations of the form Ax=b.

By Cramer's rule it has a unique solution if and only if A is invertible, or equivalently $\det(A) \neq 0$. The proof below shows that A has a non-zero determinant using the formulas of the

Greeks Vega, Volga and Vanna above.

$$\begin{aligned}
\det(A) &= \mathcal{V}(K_1)\det\begin{pmatrix} \frac{\mathcal{V}(K_2)d_1(K_2)d_2(K_2)}{\sigma} & \frac{\mathcal{V}(K_3)d_1(K_3)d_2(K_3)}{\sigma} \\ \frac{-\mathcal{V}(K_2)d_2(K_2)}{\sigma S_{a,b}(0)\sqrt{T_a}} & \frac{-\mathcal{V}(K_3)d_2(K_3)}{\sigma S_{a,b}(0)\sqrt{T_a}} \end{pmatrix} \\
&- \mathcal{V}(K_2)\det\begin{pmatrix} \frac{\mathcal{V}(K_1)d_1(K_1)d_2(K_1)}{\sigma} & \frac{\mathcal{V}(K_3)d_1(K_3)d_2(K_3)}{\sigma} \\ \frac{-\mathcal{V}(K_1)d_2(K_1)}{\sigma S_{a,b}(0)\sqrt{T_a}} & \frac{-\mathcal{V}(K_3)d_2(K_3)}{\sigma S_{a,b}(0)\sqrt{T_a}} \end{pmatrix} \\
&+ \mathcal{V}(K_3)\det\begin{pmatrix} \frac{\mathcal{V}(K_1)d_1(K_1)d_2(K_1)}{\sigma} & \frac{\mathcal{V}(K_2)d_1(K_2)d_2(K_2)}{\sigma} \\ \frac{-\mathcal{V}(K_1)d_2(K_1)}{\sigma S_{a,b}(0)\sqrt{T_a}} & \frac{-\mathcal{V}(K_2)d_2(K_2)}{\sigma S_{a,b}(0)\sqrt{T_a}} \end{pmatrix} \\
&= \mathcal{V}(K_1)\frac{\mathcal{V}(K_3)\mathcal{V}(K_2)d_2(K_2)d_2(K_3)\left[d_1(K_3)-d_1(K_2)\right]}{\sigma^2 S_{a,b}(0)\sqrt{T_a}} \\
&- \mathcal{V}(K_2)\frac{\mathcal{V}(K_1)\mathcal{V}(K_3)d_2(K_1)d_2(K_3)\left[d_1(K_3)-d_1(K_1)\right]}{\sigma^2 S_{a,b}(0)\sqrt{T_a}} \\
&+ \mathcal{V}(K_3)\frac{\mathcal{V}(K_1)\mathcal{V}(K_2)d_2(K_1)d_2(K_2)\left[d_1(K_2)-d_1(K_1)\right]}{\sigma^2 S_{a,b}(0)\sqrt{T_a}} \\
&= \frac{\mathcal{V}(K_1)\mathcal{V}(K_2)\mathcal{V}(K_3)}{\sigma^2 S_{a,b}(0)\sqrt{T_a}}\log\frac{K_2}{K_1}\log\frac{K_3}{K_1}\log\frac{K_3}{K_2}
\end{aligned}$$

The last equality is by straightforward algebraic simplifications. Hence we have $\det(A) > 0$ because the strikes are chosen such that $K_1 < K_2 < K_3$.

So the system admits a unique solution and we get the desired solution by:

$$x = A^{-1}b \text{ with } b = \begin{pmatrix} \frac{\partial C^{BS}_{a,b}(x)}{\partial \sigma} \\ \frac{\partial^2 C^{BS}_{a,b}(x)}{\partial \sigma^2} \\ \frac{\partial^2 C^{BS}_{a,b}(K)}{\partial \sigma \partial S_{a,b}(0)} \end{pmatrix} \qquad \square$$

After providing the reader with the necessary background behind the first the Vanna-Volga approach, we shall use the method in the pricing of a CMS swap. We plug (5.1) in (4.1) to get:

$$\Pi^{VV} \approx \frac{1}{\bar{f}(S_{a,b}(0))}\left[\bar{f}(0)S_{a,b}(0) + \int_0^\infty \left(x\bar{f}''(x) + 2\bar{f}'(x)\right) C^{VV}_{a,b}(x)dx\right] \tag{5.2}$$

5.1.2 Second approach

The first Vanna -Volga approach consists of pricing CMS linked products by replication following the steps in section 3, and then applying the Vanna-Volga method to get the Payer swaption price for a given strike K denoted by $C_{a,b}(K)$. The second approach that will be detailed in this section is to apply Vanna-Volga directly on the European payoff H to find $E^{a,b}[H]$. For example, H could be the payoff of a CMS cap:

$$H(S_{a,b}(T_a)) = \bar{f}(S_{a,b}(T_a))(S_{a,b}(T_a) - K)^+$$

.

$$\frac{1}{\bar{f}(S_{a,b}(0))}E^{a,b}[\bar{f}(S_{a,b}(T_a))(S_{a,b}(T_a)-K)^+)] = \Pi^{H,VV} := \Pi^{H,BS} + \sum_{i=1}^{3} w_i^H [C_{a,b}^{MKT}(K_i) - C_{a,b}^{BS}(K_i)]$$

$\Pi^{H,BS}$ is the time 0-Black and Scholes price of the payoff H payed at time $T_a + \delta$. We can compute it the exact same way as we did in section 4. The weights x_i^H are the unique solution of the following system:

$$\begin{cases} \frac{\partial \Pi^{H,BS}}{\partial \sigma} = \sum_{i=1}^{3} x_i^H \frac{\partial C_{a,b}^{BS}(K_i)}{\partial \sigma} \\ \frac{\partial^2 \Pi^{H,BS}}{\partial \sigma^2} = \sum_{i=1}^{3} x_i^H \frac{\partial^2 C_{a,b}^{BS}(K_i)}{\partial \sigma^2} \\ \frac{\partial^2 \Pi^{H,BS}}{\partial \sigma \partial S_{a,b}(0)} = \sum_{i=1}^{3} x_i^H \frac{\partial^2 C_{a,b}^{BS}(K_i)}{\partial \sigma \partial S_{a,b}(0)} \end{cases}$$

In what follows, the proofs follow closely the work [9] done in the appendix of [3].

Lemma 5.2. *: Let H be the payoff of a CMS swap. The relation between the weights of the first Vanna -Volga approach and the second Vanna-Volga approach is as follows:*

$$x_i^H = \int_0^\infty (\bar{f}''(K)K + 2\bar{f}'(K))w_i(K)dK$$

Proof.

$$\Pi^{H,BS} = \frac{1}{\bar{f}(S_{a,b}(0))}\left[\bar{f}(0)S_{a,b}(0) + \int_0^\infty (\bar{f}''(K)K + 2\bar{f}'(K))C_{a,b}^{BS}(K)dK\right]$$

Computing the partial derivatives to get the Vega, Volga and Vanna of Π^{BS}, we get:

$$\begin{cases} \frac{\partial \Pi^{H,BS}}{\partial \sigma} = \left[\int_0^\infty \left(\bar{f}''(K)K + 2\bar{f}'(K)\right) dK\right] \frac{\partial C_{a,b}^{BS}(K_i)}{\partial \sigma} \\ \frac{\partial^2 \Pi^{H,BS}}{\partial \sigma^2} = \left[\int_0^\infty \left(\bar{f}''(K)K + 2\bar{f}'(K)\right) dK\right] \frac{\partial^2 C_{a,b}^{BS}(K_i)}{\partial \sigma^2} \\ \frac{\partial^2 \Pi^{H,BS}}{\partial \sigma \partial S_{a,b}(0)} = \left[\int_0^\infty \left(\bar{f}''(K)K + 2\bar{f}'(K)\right) dK\right] \frac{\partial^2 C_{a,b}^{BS}(K_i)}{\partial \sigma \partial S_{a,b}(0)} \end{cases}$$

But the weights $w_i(K)$ are defined as the solution to the system:

$$\begin{cases} \frac{\partial C_{a,b}^{BS}(K)}{\partial \sigma} = \sum_{i=1}^{3} w_i(K) \frac{\partial C_{a,b}^{BS}(K_i)}{\partial \sigma} \\ \frac{\partial^2 C_{a,b}^{BS}(K)}{\partial \sigma^2} = \sum_{i=1}^{3} w_i(K) \frac{\partial^2 C_{a,b}^{BS}(K_i)}{\partial \sigma^2} \\ \frac{\partial^2 C_{a,b}^{BS}(K)}{\partial \sigma \partial S_{a,b}(0)} = \sum_{i=1}^{3} w_i(K) \frac{\partial^2 C_{a,b}^{BS}(K_i)}{\partial \sigma \partial S_{a,b}(0)} \end{cases}$$

[9]We did them separately at first but we then found them in the appendix of [3].

So we plug in and we get:

$$\begin{cases} \frac{\partial \Pi^{H,BS}}{\partial \sigma} = \sum_{i=1}^{3} \left[\int_0^\infty \left(\bar{f}''(K)K + 2\bar{f}'(K)\right) w_i(K)dK\right] \frac{\partial C_{a,b}^{BS}(K_i)}{\partial \sigma} \\ \frac{\partial^2 \Pi^{H,BS}}{\partial \sigma^2} = \sum_{i=1}^{3} \left[\int_0^\infty \left(\bar{f}''(K)K + 2\bar{f}'(K)\right) w_i(K)dK\right] \frac{\partial^2 C_{a,b}^{BS}(K_i)}{\partial \sigma^2} \\ \frac{\partial^2 \Pi^{H,BS}}{\partial \sigma \partial S_{a,b}(0)} = \sum_{i=1}^{3} \left[\int_0^\infty \left(\bar{f}''(K)K + 2\bar{f}'(K)\right) w_i(K)dK\right] \frac{\partial^2 C_{a,b}^{BS}(K_i)}{\partial \sigma \partial S_{a,b}(0)} \end{cases}$$

However, we know that the x_i^H are defined to be the **unique** solution to the system

$$\begin{cases} \frac{\partial \Pi^{BS}}{\partial \sigma} = \sum_{i=1}^{3} x_i^H \frac{\partial C_{a,b}^{BS}(K_i)}{\partial \sigma} \\ \frac{\partial^2 \Pi^{BS}}{\partial \sigma^2} = \sum_{i=1}^{3} x_i^H \frac{\partial^2 C_{a,b}^{BS}(K_i)}{\partial \sigma^2} \\ \frac{\partial^2 \Pi^{BS}}{\partial \sigma \partial S_{a,b}(0)} = \sum_{i=1}^{3} x_i^H \frac{\partial^2 C_{a,b}^{BS}(K_i)}{\partial \sigma \partial S_{a,b}(0)} \end{cases}$$

Hence by identification, $x_i^H = \int_0^\infty (\bar{f}''(K)K + 2\bar{f}'(K))w_i(K)dK$ □

We use lemma (5.2) to prove the following proposition:

Proposition 5.3. *The first Vanna-Volga approach and the second Vanna -Volga approach described above are equivalent. This means that the price of a CMS linked product is the same using both approaches.*

For instance, for a CMS swap, we have: $PV_1^{VV} = PV_2^{VV}$ *where :*

$$PV_1^{VV} = \bar{f}(0)S_{a,b}(0) + \int_0^\infty (\bar{f}''(K)K + 2\bar{f}'(K))C_{a,b}^{VV}(K)dK$$

,

$$PV_2^{VV} = PV^{BS} + \sum_{i=1}^{3} w_i^H [C_{a,b}^{MKT}(K_i) - C_{a,b}^{BS}(K_i)]$$

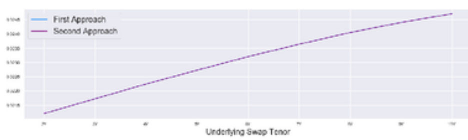
Proof. By plugging $C_{a,b}^{VV}(x) = C_{a,b}^{BS}(x) + \sum_{i=1}^{3} w_i(x)[C_{a,b}^{MKT}(K_i) - C_{a,b}^{BS}(K_i)]$ into PV_1^{VV} we get:

$$\begin{aligned} PV_1^{VV} &= \bar{f}(0)S_{a,b}(0) + \int_0^\infty (\bar{f}''(K)K + 2\bar{f}'(K)) \sum_{i=1}^{3} w_i(K) \left[C_{a,b}^{MKT}(K_i) - C_{a,b}^{BS}(K_i)\right] dK \\ &= PV^{BS} + \sum_{i=1}^{3} \left[C_{a,b}^{MKT}(K_i) - C_{a,b}^{BS}(K_i)\right] \int_0^\infty (\bar{f}''(K)K + 2\bar{f}'(K))w_i(K)dK \\ &= PV^{BS} + \sum_{i=1}^{3} w_i^H \left[C_{a,b}^{MKT}(K_i) - C_{a,b}^{BS}(K_i)\right] = PV_2^{VV} \end{aligned}$$

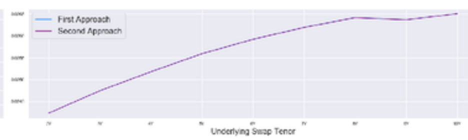
The last equality uses the lemma (5.2) proved above. □

5.1.3 Results and Discussion

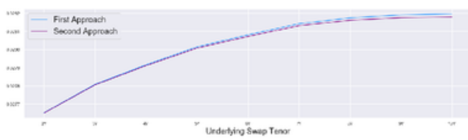
In all what follows, we will use the third order approximation of $\bar{f}$ to increase the computation time since we made sure in the previous section that it produces accurate results.



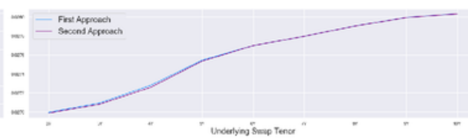
(a) CMS-2Y.



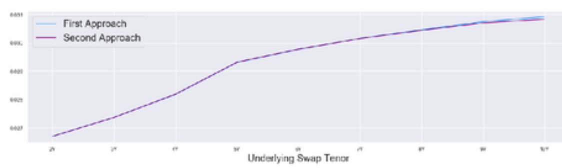
(b) CMS-5Y.



(c) CMS-10Y.



(d) CMS-20Y.



(e) CMS-30Y.

Figure 4: Comparison of the Vanna-Volga price computed with the first approach and the second approach.

The blue curve shows the price calculated with the first approach where we compute the Vanna-Volga payer swaption price and use the replication method to price. The purple curve applies Vanna-Volga directly on the payoff of the CMS swap. Figure 4 provides a good assertion of our claim that the two Vanna-Volga approaches are consistent and give the same prices.

We would like to implement another CMS pricing method that takes the smile into consideration to evaluate how accurate is the Vanna -Volga approach. Thus in the following subsection, we will introduce the SABR model and implement it to price a CMS swap.

5.2 SABR model

The Stochastic Alpha, Beta, Rho (SABR) model is a stochastic volatility model which, in contrast to Black's model, does not assume that the volatility is constant. It is used to match the volatility smile or skew by modeling one forward rate with stochastic volatility dynamics.

5.2.1 Model definition and dynamics

A constant volatility model, as the name indicates, does not capture any part of the volatility smile present in the market. The assumption of constant volatility is therefore unrealistic and needs to be relaxed. Thus the urge to switch to a model that is consistent with the observable market skew. At first, we might think of choosing a local volatility model where the volatility is a function of the current forward level $S_{a,b}(t)$ and the time t. The reason we might think it is a good choice is that such a model can be calibrated to have a perfect match with the market smile. However, if we use a local volatility model, $S_{a,b}$ and the market smile will move in opposite directions, contrary to what is supposed to happen. To resolve this problem, Hagan et al. derived the stochastic model SABR in [7] which correlates the forward asset with the volatility, and most importantly, matches correctly the market volatility smile. It assumes that the dynamics' type of the forward-asset is CEV (Constant-Elasticity-of-Variance).

Using SABR model, the forward-rate dynamics under the associated forward swap measure $\mathbb{Q}^{a,b}$ are modelled as follows:

$$\begin{aligned} dS_{a,b}(t) &= V(t)S_{a,b}(t)^{\beta}dZ^{a,b}(t), \\ dV(t) &= \epsilon V(t)dW^{a,b}(t), \\ V(0) &= \alpha \end{aligned}$$

where $Z^{a,b}$ and $W^{a,b}$ are $\mathbb{Q}^{a,b}$-standard Brownian motions that are correlated by ρ:

$$dZ^{a,b}(t)dW^{a,b}(t) = \rho dt$$

[5]. It is complicated to work with the probability distribution of the SABR model as it is hard to find an analytical closed formula of the implied volatility $\sigma^{imp}(K, S_{a,b}(0))$ of the swaption with maturity T_a and strike K. Hence, Hagan et al used singular perturbation techniques and derived

the following approximation [7]:

$$\sigma^{imp}(K, S_{a,b}(0)) \approx \frac{\frac{\alpha}{(S_{a,b}(0)K)^{\frac{1-\beta}{2}}\left[1+\frac{(1-\beta)^2}{24}\log^2(\frac{S_{a,b}(0)}{K})+\frac{(1-\beta)^4}{1920}\log^4(\frac{S_{a,b}(0)}{K})\right]}\frac{z}{x(z)}.}{\left[1+\left(\frac{(1-\beta)^2\alpha^2}{24(S_{a,b}(0)K)^{1-\beta}}+\frac{\rho\beta\epsilon\alpha}{4(S_{a,b}(0)K)^{\frac{1-\beta}{2}}}+\epsilon^2\frac{2-3\rho^2}{24}\right)T_a\right]}. \tag{5.3}$$

where

$$z := \frac{\epsilon}{\alpha}(S_{a,b}(0))^{\frac{1-\beta}{2}} z \log\left(\frac{S_{a,b}(0)}{K}\right)$$

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

In the special at-the-money case where K=$S_{a,b}(0)$, the implied volatility $\sigma^{imp}(K, S_{a,b}(0))$ reduces to:

$$\sigma^{imp}(S_{a,b}(0), S_{a,b}(0)) \approx \frac{\alpha}{(S_{a,b}(0))^{1-\beta}}\left[1+\left(\frac{(1-\beta)^2\alpha^2}{24(S_{a,b}(0))^{2-2\beta}}+\frac{1}{4}\frac{\rho\beta\epsilon}{(S_{a,b}(0))^{1-\beta}}+\frac{1}{24}(2-3\rho^2)\epsilon^2\right)T_a\right] \tag{5.4}$$

5.2.2 SABR parameters

Before starting with the model calibration to market data, it is important to study the effect of each parameter individually on the volatility smile.

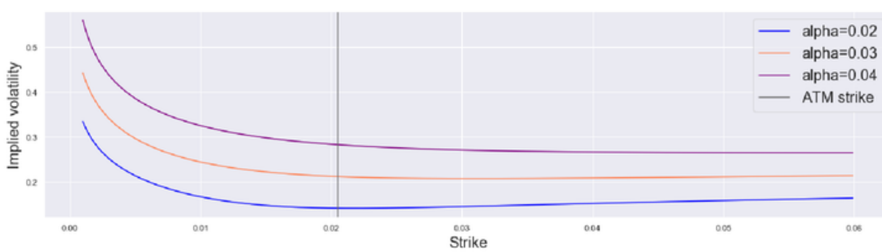


Figure 5: Effect of α on implied volatility

We notice that as we increase α, the implied volatility shifts upward. The explanation of this is rather straightforward because we know that $V(0) = \alpha$ so α is the time 0-expectation of the stochastic volatility process V. It is also good to see that around the ATM-strike which is represented by the gray vertical line in the figure above, α has a larger effect than the shift it produces at the extreme strikes. So α has a more substantial effect on the ATM volatility which can also be directly seen analytically by comparing the two formulas in (5.3) and (5.4).

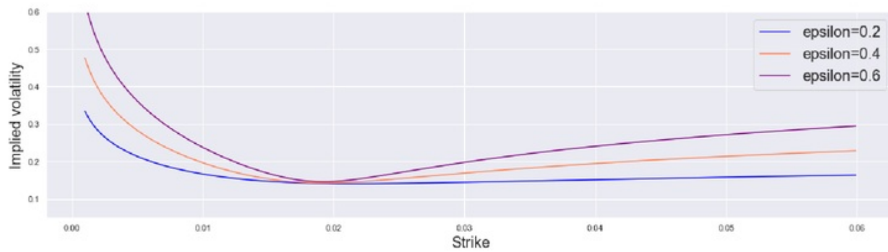


Figure 6: Effect of ϵ on the implied volatility

The figure above shows that when we manipulate ϵ, the curvature of the volatility smile changes. More specifically, increasing ϵ leads to a smile that is more convex.

The next step is to consider the impact of ρ on the volatility smile. It is important to take into account that ρ can take negative values so comparing two values of rho with opposite sign makes no sense. Thus, we will look at the impact of ρ in two separate cases.

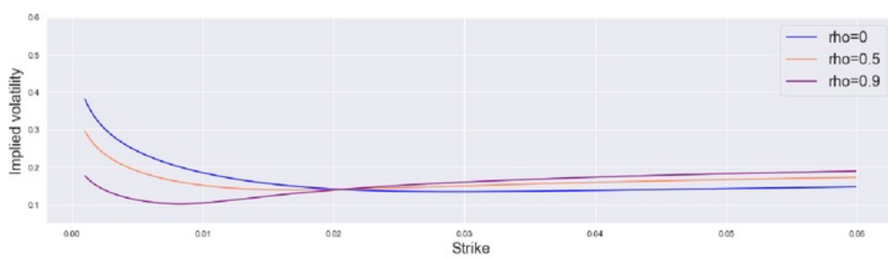


Figure 7: Effect of ρ on the implied volatility (for positive ρ)

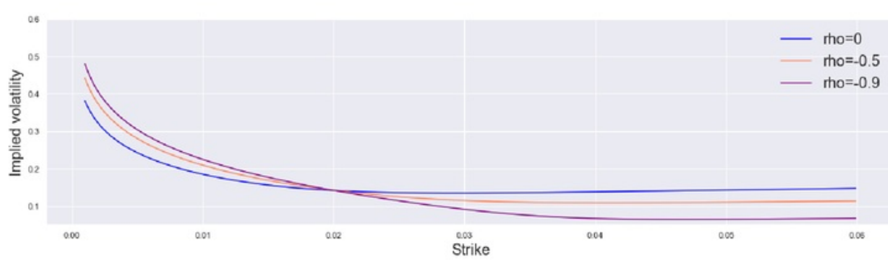


Figure 8: Effect of ρ on the implied volatility (for negative ρ)

As is shown in the two figures above, we can see that when the correlation is positive, increasing ρ leads to a volatility smile that is steeper. Whereas it is the opposite in the negative case as increasing the ρ produces a smile that is flatter.

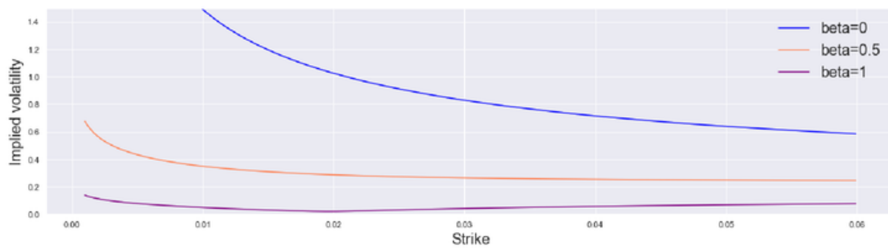


Figure 9: Effect of β on implied volatility

We can see in figure 9 that when we decrease the value β, the volatility smile becomes steeper. We notice that β has the same impact on the volatility smile as ρ, in the sense that they both affect the steepness of the smile. Since it is undesirable to fit parameters having the same effect simultaneously, we will see in the next section that β is fixed during the model's calibration to the swaption volatilities data. Aside from affecting the steepness, we also notice that increasing β causes a drop in the volatility level which is rather intuitive since β is the exponent of the forward swap rate in:

$$dS_{a,b}(t) = V(t)S_{a,b}(t)^{\beta}dZ^{a,b}(t). \tag{5.5}$$

ranging between 0 and 1. We look at the diffusion term $V(t)S_{a,b}(t)^{\beta}$ in the SDE (5.5) and we interpret it as the volatility of the forward rate process. So, if we decrease β, the volatility is expected to become higher. Another thing to observe is that modifying β has more impact on the volatility smile of swaptions with lower strikes. In fact, we notice in figure 8 that the "difference" between the two curves is larger for lower strikes than for higher strikes.

Having discussed the individual impact of each parameter on the volatility smile, we are now ready to calibrate the model.

5.2.3 Model Calibration

Model Calibration is particularly crucial before we take any further step in the pricing. The steps usually depend on the purpose of use of the model. For instance, the calibration of a SABR model that is used to price CMS linked products is different than the usual calibration process where SABR is used for other vanilla instruments. Ideally, to price CMS swaps using SABR model, we should calibrate the parameters to CMS swaps spreads market data. However, there are only a few CMS spreads quoted in the market so these are not enough to compute directly the convexity adjustments. Hence, this motivates a joint calibration with swaption volatilities data as suggested by Mercurio and Pavaliccini in [5].

While attempting to fit the SABR model to match the swaptions market data, an important step is the determination of the parameter β. There is a standard approach suggested in [7] that is often implemented by practitioners to determine this parameter. This approach uses historical observations of the "backbone". The term backbone is used to describe the at-the-money volatility as a function of the forward price. In this procedure, β is found from the gradient of the plot of historical values of σ_{ATM} against the forward price. Additionally, it is claimed in the financial literature that any value of β results in a good match with the market smile. Indeed, we can choose two different values of β and have a good fit of the volatility smile with both values.

On top of that, we observed in the previous section that the parameter β has a similar effect on the volatility smile as the parameter ρ which usually represents the "market noise". This shows that calibrating β using market data is not desirable.

For our purpose, the calibration of the SABR model is done in two parts. First, we fix β to a random value and we calibrate the three parameters α, ϵ and ρ to the swaption volatilities data. Each swap rate is thus associated with different parameters α, ϵ and ρ. We then calibrate β to CMS swap spreads data. The parameter β is assumed to be equal across different maturities and tenors.

We shall provide the reader with the detailed steps of our calibration.

For each couple (a,b) corresponding to the start date and the maturity date of the swaption's underlying swap, we will use five available quotes in the market: the at-the-money strike ($K_2 = K_{ATM}$) with its corresponding volatility $\sigma_2 = \sigma_{ATM}$ and 4 other quotes K_1, K_3, K_4, K_5 with their corresponding volatilities σ_1, σ_3, σ_4 and σ_5. Two of the quotes are in-the-money (K_1, K_4) and the other two are out- of-the-money (K_3, K_5). The goal is to minimize the square error between the implied volatility given by the SABR model and the implied volatility given by the market.

So for each couple (a,b), we have the following optimization problem:

$$\min_{\alpha,\epsilon,\rho} \sum_{i=1}^{5} \left(\sigma_i^{SABR}\left(K_i, S_{a,b}(0)\right) - \sigma_i^{MKT}\left(K_i, S_{a,b}(0)\right)\right)^2$$

subject to

$0 \leq \alpha < \infty,$

$0 \leq \epsilon < \infty,$

$-1 \leq \rho \leq 1,$

Once we get the parameters that are solution to the above optimization problem, we calculate the implied volatility given by (5.3) and then use it to get $C_{a,b}$. We then plug $C_{a,b}$ in (4.1) and subtract $S_{a,b}(0)$ to get the convexity adjustment for each couple (a,b).

The second part of the calibration is for the determination of the parameter β and is related to the CMS spread data. For every couple (n,c) where n represents the last payment date of the CMS swap and c corresponds to the c-year swap rate payed by the CMS swap, we need to calculate the CMS swap spread using the parameters $\beta, \alpha(\beta), \epsilon(\beta)$ and $\rho(\beta)$. The definition of a CMS swap spread leads to the formula below given in terms of the convexity adjustment in [5, p.11]:

$$X_{n,c}^{SABR} = \frac{\sum_{i=1}^{n}(S'_{i,c}(0) + CA(S'_{i,c}(0), \delta))P(0, T'_i)}{\sum_{i=1}^{n} P(0, T'_i)} - \frac{1 - P(0, T'_n)}{\delta \sum_{i=1}^{n} P(0, T'_i)} \tag{5.6}$$

The implementation of the formula (5.6) is not very straightforward as the payment dates and frequencies can be tricky. We are considering USD CMS swap spreads where the payment dates T'_i are semi-annual, whereas the frequency of the underlying swap's payments is quarterly (as per market convention for USD). So we be aware of this difference. Another thing to note is that very often, we might encounter cases where we need the convexity adjustment of a swap rate that we did not calculate in the step above. Indeed, the swap rate's underlying swap could have a different maturity or tenor due to the payment schedule discussed above. A different maturity or tenor means that it does not correspond to any couple (a,b) considered in the previous calibration step. Thus, to get the value of these convexity adjustments, we must do a cubic spline interpolation of the CMS adjustments calculated in earlier.

For our purpose, will use the quoted CMS swap spreads in the market, with CMS swap maturities and tenors equal to 5,10,15 and 20 years. We have to solve again a minimization problem which is given by:

$$\min_{\beta} \sum_{n} \sum_{c} (X_{n,c}^{SABR} - X_{n,c}^{MKT})^2$$

subject to

$$0 \leq \beta \leq 1$$

So the idea is to minimize the square error between the CMS spreads given by the model and those quoted in the market. After calibrating the SABR model, we are now ready to use it for our purpose of CMS pricing.

5.2.4 Application to market data

As a first step, we will treat the trivial case of SABR with constant volatility to compare with Π^{BS} calculated earlier in (4.2). Indeed, if we take $\beta = 1$, $\epsilon = 0$ and $\alpha = \sigma^{ATM}$, we should be able to get that the price calculated with SABR matches the price calculated under the Black-Scholes framework. The figure below shows the CMS swaps prices for a 5-year CMS coupon and underlying swap tenors ranging from 2 to 10 years. We can see that the two prices match, which proves that our calculations in the first part were correct.

Figure 10: 5Y-CMS swap coupon values under the assumption of constant volatility

We would like now to compare the results obtained with the Vanna-Volga approach and SABR model. We start by fixing the parameter $\beta = 0.5$ to calibrate the three remaining parameters α, ρ and ϵ to the swaption volatilities. We will use the same volatilities quotes as the ones we used in Vanna-Volga approach but with two additional quotes (K_4 and K_5) to improve the precision of the calibration. The minimization is done using sequential least squares programming ('SLSQP') in Python. 'SLSQP' is an iterative method that solves a sequence of constrained optimization problems. We will not detail this method in our thesis as it is not our main goal, but interested readers can read about it in [14, chapter 2].

We start with an initial vector guess $w_0 = [\alpha_0, \epsilon_0, \rho_0]$, and we suggest to replace it at each iteration by the calibrated parameters of the previous couple maturity/tenor. The reason for that is that we want to avoid using a random initial guess for every single iteration, so we thought of a way that minimizes the randomness involved in each iteration. We tried using a new random initial guess at every step in the iteration but the results were not as accurate as using the way that we suggested. The justification behind this choice of implementation is that we are assuming that there is a certain "complementarity" between the parameters when the swaptions' maturities and tenors are close.

The tables below show the results of the first calibration to swaption volatilities data.

Table 1: Results showing the value of α calibrated

T_a \ $T_b - T_a$	2Y	3Y	4Y	5Y	6Y
2Y	0.066533	0.065210	0.063488	0.061291	0.059268
5Y	0.053901	0.052881	0.051871	0.050849	0.049634
10Y	0.045263	0.045010	0.045226	0.045226	0.044570
20Y	0.040695	0.041097	0.041596	0.042097	0.041686

T_a \ $T_b - T_a$	7Y	8Y	9Y	10Y
2Y	0.057407	0.055690	0.054130	0.052703
5Y	0.048533	0.047464	0.020707	0.019738
10Y	0.044030	0.043448	0.042878	0.042343
20Y	0.041279	0.040877	0.040478	0.040083

Table 2: Results showing the value of ϵ calibrated

T_a \ $T_b - T_a$	2Y	3Y	4Y	5Y	6Y
2Y	0.01327	0.01405	0.01563	0.01696	0.01630
5Y	0.01390	0.012881	0.014812	0.010793	0.013294
10Y	0.032032	0.032031	0.032031	0.032030	0.032017
20Y	0.031157	0.031166	0.031169	0.031172	0.031166

T_a \ $T_b - T_a$	7Y	8Y	9Y	10Y
2Y	0.014503	0.016594	0.017331	0.017269
5Y	0.015823	0.014654	0.034167	0.034141
10Y	0.033109	0.033091	0.033074	0.033058
20Y	0.032004	0.031993	0.031981	0.031970

Table 3: Results showing the value of ρ calibrated

T_a \ $T_b - T_a$	2Y	3Y	4Y	5Y	6Y
2Y	0.33292	0.33289	0.33285	0.33281	0.33277
5Y	0.33171	0.33170	0.33168	0.33167	0.33165
10Y	0.31105	0.31089	0.31073	0.31073	0.31048
20Y	0.29605	0.29598	0.29589	0.29580	0.29563

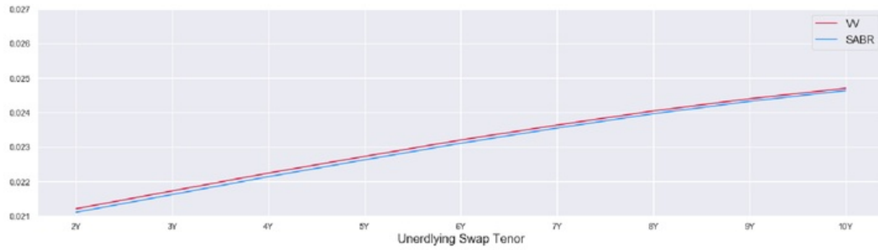
T_a \ $T_b - T_a$	7Y	8Y	9Y	10Y
2Y	0.33274	0.33271	0.033268	0.033266
5Y	0.33164	0.33162	0.36460	0.36619
10Y	0.31028	0.31006	0.30986	0.30966
20Y	0.29546	0.29529	0.29512	0.29495

We use the parameters in tables 3,4 and 5 above to calibrate the β parameter to CMS swap spreads quotes as discussed earlier. We get that $\beta = 0.66998453$.
Using the above results together, we will find CMS swaps and caps prices given by the SABR model.

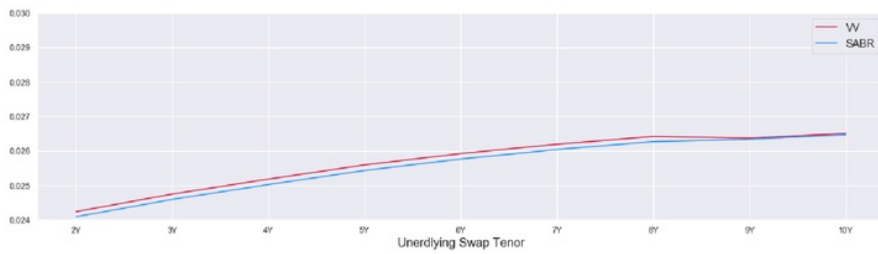
5.3 Comparison of the two interpolation methods

5.3.1 Results

We shall now compare the results of the two interpolation-methods. We illustrate the CMS swap prices in the following graphs.

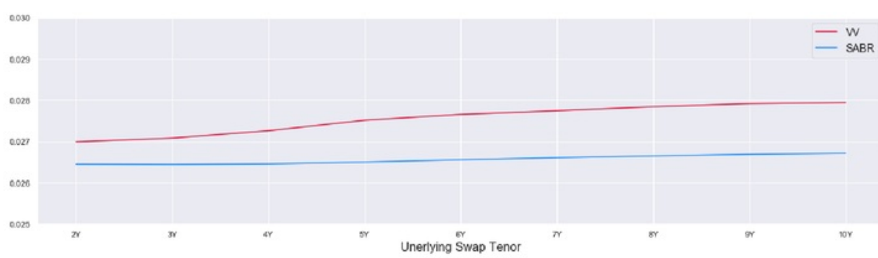


(a) 2Y-CMS coupon



(b) 5Y-CMS coupon

(c) 10Y-CMS coupon



(d) 20Y-CMS coupon

Figure 11: Comparison of Π^{VV} and Π^{SABR}.

We will plot the absolute difference in percentage in the figure below to have a clearer comparison. In all the bar plots that follow, we will denote by maturity the CMS coupon payment date. In fact, each CMS coupon is related to a series of swaptions with maturity equal to the payment date.

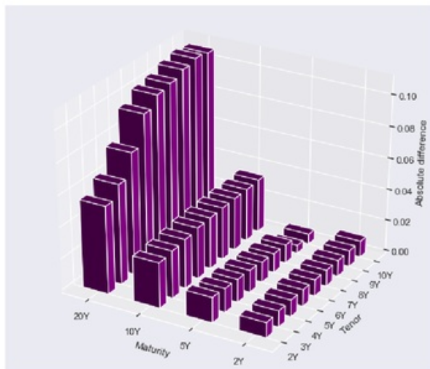
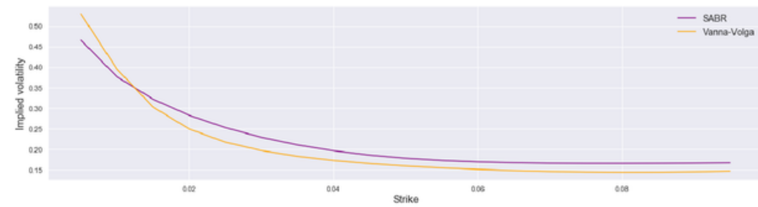
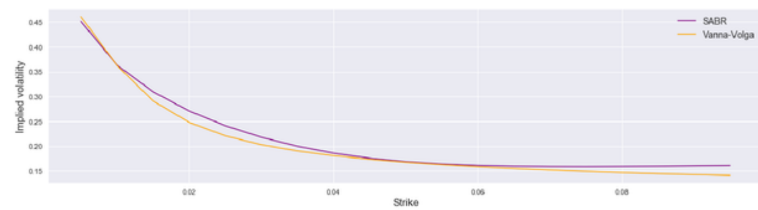


Figure 12: Absolute difference (in percentage) of the SABR and the Vanna-Volga CMS swap value

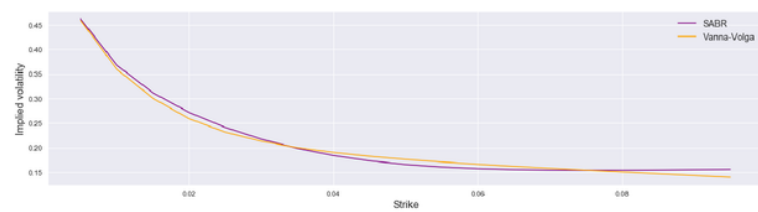
As we can see, both interpolation methods give results that are close. The difference is very negligible for short maturities and short tenors. Figure 12 shows that the difference increases if we fix the tenor and increase the maturity. It also increase if we fix the maturity and increase the tenor. The difference hits its maximum of 0.1% for a maturity of 30 years and a tenor of 10 years, which is still considered a small discrepancy. This shows that the two methods' results are in line with each other. We would like to try to justify our observation that the difference increases as the maturity and the tenor increase. We will do so by plotting the smiles obtained with both interpolation methods. Our hypothesis is that the smiles obtained from the two methods are closer for a shorter maturity and a shorter tenor.



(a) Volatility smile (Maturity 30Y Tenor 9Y)



(b) Volatility smile (Maturity 30Y Tenor 2Y)



(c) Volatility smile (Maturity 5Y Tenor 2Y)

Figure 13: Comparison between the swaption implied volatilities generated by Vanna-Volga and SABR.

The subfigures 13.a and 13.b show that the volatility smiles given by the two methods are less matching for a higher underlying swap's tenor (keeping the maturity constant). Additionally, figures 13.b and 13.c show that by increasing the maturity, the difference between the implied volatility using SABR and Vanna-Volga increases too. These results combined together are a good explanation of the barplot in figure 12. Hence, this confirms our hypothesis about the difference between the implied volatility smile given by SABR and by Vanna-Volga.

Since the prices we got for Vanna -Volga are consistent with the results we got from implementing SABR, we could think that we can use either of the methods with no particular preference. However, each method has its own characteristics that create an advantage or a disadvantage for traders. We shall discuss these.

5.3.2 Discussion

Clearly, our work shows that Vanna -Volga method is easier to implement than the SABR model as it does not need any further calibration to the market data, whereas the SABR model requires a joint calibration to give accurate results. So, in terms of ease of implementation, Vanna-Volga has an advantage over SABR. This good characteristic of the Vanna-Volga comes from the market information on three swaption volatilities that are used to get the swaption price for a given strike K. Nevertheless, we should be cautious while using the Vanna-Volga approach since it is only an empirical "rule of thumb" and it is not built on a stable foundation. More specifically, there is nothing in the way this method has been constructed that guarantees the absence of arbitrage. This being said, before using this approach to price CMS linked products, we should make sure to eliminate possible arbitrage opportunities. The first thing that should be checked is that the value of a swaption given by (5.1) must be positive. The latter condition is not always satisfied. We will provide the reader with an example in figure 14 below that illustrates a case where the value of a payer swaption is negative which can lead to an arbitrage. We would like to give possible conditions that could lead to inappropriate results in this method. First, we must ensure that we have a good choice of the wings K_1 and K_3. We cannot choose any two strikes that satisfy $K_1 < K_2 < K_3$. The two strikes respectively lower and higher than the ATM have to be chosen carefully such that the absolute value of the delta [10] of the payer and receiver swaptions is as close as possible to a certain target level (typically 25% for liquidity purposes) [1]. On another note, evaluating the price of a swaption with a strike that is different than the input at-the-money and the delta pillars will not necessarily lead to a correct value. Indeed, the inputs of this method consist of only three volatilities so we are only sure of having good results for the aforementioned strikes. Hence while using the Vanna-Volga approach, as mentioned in [15], one should not be optimistic about retrieving correct values for swaptions that are other than the at-the-money and target-Δ swaption. Indeed the nature of this method does not guarantee the accurate reproduction of the entire volatility matrix based on only three input volatilities. A final thing to note is that the Vanna-Volga method oftens gives good interpolation values between the range K_1 and K_3 but the extrapolation outside of the wings chosen is not always good.

[10] Δ is here used with a slight abuse of terminology as it does not include the annuity

(a) VV Payer swaption. (b) VV Payer Swaption zoomed in.

Figure 14: Negative Vanna-Volga swaption price.

The figure above clearly shows that the Vanna-Volga method can indeed lead to a negative payer swaption price.

We shall now provide the reader with the properties of SABR model. The first advantage is the presence of an efficient approximation of the implied volatility as a function of the 4 parameters given by (5.3). Although it is only an approximation, it is often used by practitioners as an exact closed form formulae because it allows for a relatively quick calibration to the swaption volatilities market data. However, it is worth mentioning that Hagan et al. used assymptotic expansions in [7] to derive this closed formula which can lead to a negative probability density function of the forward rate for very low strikes. Many authors in the financial literature attempted to improve this weakness of the SABR but we will not discuss this area of research in this thesis. Interested readers can find the details in [21].

Another common weakness of this model that is more relevant to our work is the stability of the 4 parameters. It would be very inconvenient to have major bounces in the SABR parameters across maturities and tenors. This is why we tried to overcome this issue by replacing the initial guess at each iteration in the calibration with the calibrated parameters of the previous couple maturity/tenor. On top of that, we might encounter an additional problem in the calibration which is the presence of more than one local minima in the square error that is minimized. This leads us to say that using the SABR model could result in an uncertainty in the parameters, which is not faced in the Vanna-Volga approach.

5.4 CMS Cap Results

Strike	Cap Value VV	Cap Value SABR	Cap Value BS
2%	93.74	93.21	91.87
4%	36.96	36.32	34.95
10%	7.10	6.93	3.19

Table 4: Results showing the values in basis points of 7-years CMS caps on the CMS-20Y index, calculated in the three ways

Table 4 again shows that the prices of CMS caps are very close using both interpolation methods SABR and Vanna -Volga. The discrepancy is bigger with the price calculated under Black's model which is very expected since the latter does not take into consideration the market smile/skew.

6 Different yield curve models

From an econometric stand point, the dynamics of the yield curve can be modeled using three factors loading. The first factor is called the level of the yield, the second one corresponds to the slope and the third one represents the curvature. As an example, this is well-presented in the Nelson-Siegel model [11] given in [18, chap. 2] by:

$$y(\tau) = \beta_1 + \beta_2 \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} \right) + \beta_3 \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right).$$

The figure below is taken from [18] and it shows the loadings plotted as a function of the maturity T-t.

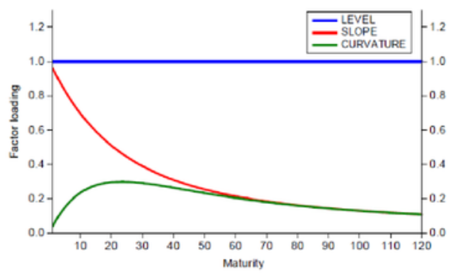


Figure 15: Loadings as a function of the maturity T-t

Following this, we would like to change the standard bond-math approximation used by most practitioners by considering different yield curves modeling. The first function which is market practice assumes a flat yield curve. The second one models the yield curve with parallel shifts which is linked to the first loading factor. The third one assumes non parallel shifts so it takes into consideration the slope of the yield curve, involving the second loading factor.

In this section, we denote by $R_s(t)$ the swap rate of an underlying swap that starts at time s_0 and ends at s_n where usually $s_n = s_0 + N$. [12] We also denote by τ_1 the strictly positive constant year fraction. We will start by explaining the rationale behind the standard model used by the practitioners.

[11] This is not the 1987 original version of the model, it has been updated in 2006

[12] $R_s(t)$ was denoted by $S_t^{s_0, s_0+N}$ in the previous sections but we change the notation in this section just to be consistent with the notations in the reference paper [4].

6.1 Standard Model with flat yield curve

The standard "bond math model" used for the computation of convexity adjustments assumes flat initial and final yield curves. On top of that, it assumes that there is a perfect correlation between the rates of different maturities. We will justify the steps behind the function in (2.4).

For all $j \leq n$, we have:

$$\begin{aligned}\frac{P(t,s_0)}{P(t,s_j)} &= \frac{P(t,s_0)}{P(t,s_1)}\frac{P(t,s_1)}{P(t,s_2)}\cdots\frac{P(t,s_{j-2})}{P(t,s_{j-1})}\frac{P(t,s_{j-1})}{P(t,s_j)} \\ &\approx (1+\tau_1 R_s(t))^j\end{aligned}$$

The approximation above is obtained by discounting all cash flows after time t at the same rate given by the swap rate $R_s(t)$, due to the assumption of flat yield curve. This justifies why the j ratios that we decomposed above are all equal. Hence,

$$\frac{P(t,s_j)}{P(t,s_0)} \approx \frac{1}{(1+\tau_1 R_s(t))^j}$$

We plug this in L(t) to get:

$$\begin{aligned}L(t) = \sum_{j=1}^{n}\theta_j P(t,s_j) &= P(t,s_0)\sum_{j=1}^{n}\theta_j\frac{P(t,s_j)}{P(t,s_0)} \\ &\approx P(t,s_0)\sum_{j=1}^{n}\frac{\tau_1}{(1+\tau_1 R_s(t))^j} \\ &= \frac{P(t,s_0)}{R_s(t)}\left[1-\frac{1}{(1+\tau_1 R_s(t))^n}\right]\end{aligned}$$

The second step is by considering that $\theta_j = \tau_1$. The last equality uses the sum of a geometric sequence.

Similarly, we can write:

$$\frac{P(t,t_p)}{P(t,s_0)} \approx \frac{1}{(1+\tau_1 R_s(t))^{\Delta}}.$$

where Δ is the fraction of the period between the swap's start date and the pay date given by:

$$\Delta = \frac{t_p - s_0}{s_1 - s_0} = \frac{t_p - s_0}{\tau_1}.$$

Hence, we get:

$$\frac{P(t,t_p)}{L(t)} \approx R_s(t)\frac{\left(\frac{1}{1+\tau_1 R_s(t)}\right)^{\frac{t_p-s_0}{\tau_1}}}{1-\left(\frac{1}{1+\tau_1 R_s(t)}\right)^{b-a}} \tag{6.1}$$

(6.1) shows that $\frac{P(t,t_p)}{L(t)}$ can be expressed in terms of $R_s(t)$, so once $R_s(t)$ is given, $\frac{P(t,t_p)}{L(t)}$ is just a constant and the expectation of a constant is the constant itself. Hence,

$$\bar{f}(R_s(t)) := \mathrm{E}^{a,b}\left[\frac{P(t,t_p)}{L(t)}\middle|R_s(t)\right] \approx R_s(t)\frac{\left(\frac{1}{1+\tau_1 R_s(t)}\right)^{\frac{t_p-s_0}{\tau_1}}}{1-\left(\frac{1}{1+\tau_1 R_s(t)}\right)^{b-a}}.$$

Rewriting it according to our notations in the sections above where δ denotes the accrual period, this will retrieve the function in (2.4) as desired.

A flat yield curve is a sign that there is very little difference between short-term and long-term rates for bonds having the same credit quality. Although this assumption is very simplistic and rather unrealistic, practitioners still use it in their work. For example, Castagna, Mercurio and Tarenghi used it in their implementation for the Vanna-Volga approach and SABR model for CMS pricing. We are interested in seeing how relaxing some of these simplistic assumptions would contribute to the prices of CMS.

6.2 Yield curve with parallel shifts

We will implement the first model that takes into consideration the shape of the initial yield curve. The model still assumes parallel shifts of the yield curve which is rather a simplistic assumption. Parallel shifts usually happen when the interest rates on all fixed-income maturities change by the same number of basis points regardless if the maturity is short term, intermediate term or long term. The slope and the shape of the yield curve thus do not change, we just have a shift of the data points. Hagan proposes the following approximation in [4]:

$$\frac{P(t,s_j)}{P(t,s_0)} \approx \frac{P(0,s_j)}{P(0,s_0)}e^{-(s_j-s_0)x} \qquad (6.2)$$

where x represents the amount of the parallel shift depending on t. However, the paper does not justify the rationale behind this approximation. Hence, we will try to provide the reader with a possible justification. Our starting point is:

$$P(t,T) = e^{-\int_t^T f(t,s)ds} \qquad (6.3)$$

where due to the assumption of parallel shifts, we suggest to model the instantaneous forward rate as follows:

$$f(t,s) \approx f(0,s) + x_t \qquad (6.4)$$

It is important to notice that the amount of parallel shift in (6.4) depends on t only and not on s. [13] Plugging (6.4) in (6.3) leads to:

$$
\begin{aligned}
\frac{P(t,s_j)}{P(t,s_0)} &\approx \frac{e^{-\int_t^{s_j} f(0,s)ds-(s_j-t)x}}{e^{-\int_t^{s_0} f(0,s)ds-(s_0-t)x}} \\
&= \frac{e^{-\int_0^t f(0,s)ds}}{e^{-\int_0^t f(0,s)ds}} \cdot \frac{e^{-\int_t^{s_j} f(0,s)ds}}{e^{-\int_t^{s_0} f(0,s)ds}} e^{-(s_j-s_0)x} \\
&= \frac{P(0,s_j)}{P(0,s_0)} e^{-(s_j-s_0)x}
\end{aligned}
$$

We continue by following the same steps as the subsection above and we get:

$$
L(t) \approx P(t,s_0) \sum_{j=1}^{n} \theta_j \frac{P(0,s_j)}{P(0,s_0)} e^{-(s_j-s_0)x}
$$

$$
\begin{aligned}
R_s(t) &= \frac{P(t,s_0) - P(t,s_n)}{L(t)} \\
&\approx \frac{P(0,s_0) - P(0,s_n) e^{-(s_n-s_0)x}}{\sum_{j=1}^{n} \theta_j P(0,s_j) e^{-(s_j-s_0)x}}
\end{aligned}
$$

The new model gives the following function:

$$
\frac{P(t,t_p)}{L(t)} \approx \frac{R_s e^{-(t_p-s_0)x}}{1 - \frac{P(0,s_n)}{P(0,s_0)} e^{-(s_n-s_0)x}} \tag{6.5}
$$

where x which is the amount of the parallel shift and is determined implicitly in terms of $R_s(t)$ by:

$$
R_s(t) \sum_{i=1}^{n} \theta_j P(0,s_j) e^{-(s_j-s_0)x} + P(0,s_n) e^{-(s_n-s_0)x} = P(0,s_0) \tag{6.6}
$$

Again, (6.5) expresses $\frac{P(t,t_p)}{L(t)}$ in terms of $R_s(t)$ so using the same argument as before we get:

$$
\bar{f}(R_s) \approx \frac{R_s e^{-(t_p-s_0)x}}{1 - \frac{P(0,s_n)}{P(0,s_0)} e^{-(s_n-s_0)x}} \tag{6.7}
$$

As mentioned earlier, the drawback of this function is that it only assumes parallel shifts. Nevertheless, taking into consideration the initial shape is quite an important addition to the previous standard model. It mostly contributes to a big difference whenever we are working in an environment where the yield curve is steep. If the initial yield curve is already flat, then relaxing the assumption of flat initial curve will not bring any difference.

[13] We will denote x_t by x in what follows to simplify the notations.

The implementation of this model is not as simple as than the first one because we do not have one single closed form that works. Indeed, we need to solve the equation (6.6) numerically multiple times, as we can clearly see in (6.6) that the amount of the parallel shift x depends on the swap's fixed leg maturities. So, x is not constant for every swap rate R_s.
The figure below illustrates the absolute difference in percentage between the prices of CMS swaps obtained using the standard model and the ones obtained using the new model. Each color represents the absolute difference of CMS swaps that have the same maturity and different underlying swap tenors.

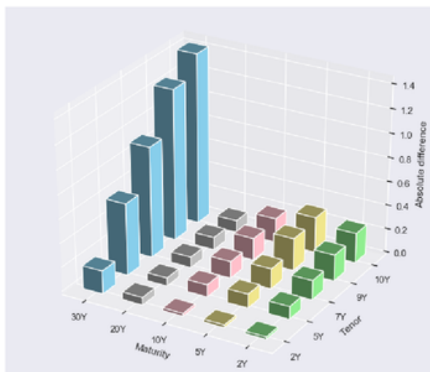


Figure 16: Absolute difference (in percentage) between the prices using the the flat yield curve model and the parallel shifts model

Figure 15 shows that there is a clear difference between the prices calculated with the two different models. Furthermore, the difference seems to increase with both the maturity and tenor . Indeed, we can see that for a CMS-2Y, the difference is very low specially for shorter tenors and it hits 0.2% for the higher tenor. Whereas if we take a CMS-30Y, the difference in prices hits approximately 1.4% for the largest 10Y tenor.
This difference could be explained by looking at the yield curve of 20 May 2019 which is the date we are using for our pricing throughout the thesis. We can see that the yield curve of this day is not flat so the assumption of initial flat yield curves will for sure lead to an error in the pricing. So this is a first explanation of the non-zero differences presented in the figure above. Intuitively, the farther we go from the starting point, the more we are dropping information about the shape of the yield curve by assuming a constant slope. So this could be a possible explanation of why the differences are increasing with respect to the tenors.

We would still like to use an improved model of the yield curve for our purpose of CMS pricing. We will detail this in the following subsection.

6.3 Yield curve with non parallel shifts

When there is a non parallel shift in the yield curve, this means that changes in yields for bonds are not constant for different maturities. Following similar steps as in the above subsection, the approximation is as follows:

$$\frac{P(t,s_j)}{P(t,s_0)} \approx \frac{P(0,s_j)}{P(0,s_0)} e^{-(h(s_j)-h(s_0))x}$$

where x is still the amount of the shift. However, instead of having the maturity s as in (6.1), it is now replaced by h(s) which is the effect of the shift on the maturity.

Again, similarly to the previous case, x is found implicitly in terms of R_s by solving:

$$R_s \sum_{i=1}^{n} \alpha_j P(0,s_j) e^{-(h(s_j)-h(s_0))x} + P(0,s_n) e^{-(h(s_n)-h(s_0))x} = P(0,s_0) \tag{6.8}$$

This model leads to the following function:

$$\bar{f}(R_s) \approx \frac{R_s e^{-(h(t_p)-h(s_0))x}}{1 - \frac{P(0,s_n)}{P(0,s_0)} e^{-(h(s_n)-h(s_0))x}}$$

We sill require to choose the function h to determine the shape of the non parallel shift. One of the suggestions proposed in [4] is to assume a constant mean reversion, which translates to:

$$h(s) - h(s_0) = \frac{1}{\kappa}\left[1 - e^{-\kappa(s-s_0)}\right] \tag{6.9}$$

where κ is the mean reversion constant.

We would like to see the effect of κ on the function $\bar{f}$.

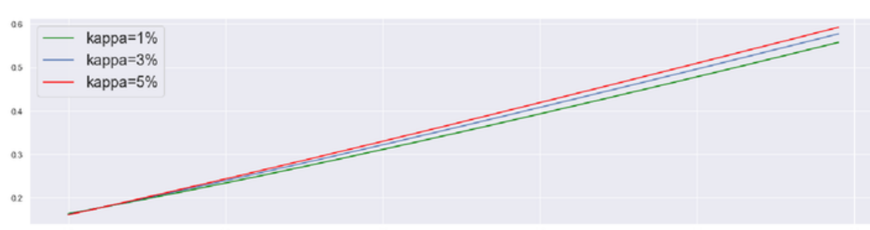


Figure 17: Effect of κ on $\bar{f}$

The figure above shows that increasing κ leads to an upward shift in $\bar{f}$. If we have a closer look, we can see that the upward shift is not parallel and the slope of $\bar{f}$ increases when κ increases. The slope being directly related to the first order derivative, this shows that increasing κ leads to an increase in $\bar{f}'$. Π^{VV} given by (5.2) depends directly on both $\bar{f}$ and $\bar{f}'$ so we should expect that if we increase κ, the value of Π^{VV} also increases. We compute the price of a 10-years CMS swap on the CMS-10Y index using $\kappa = 0.01$ and $\kappa = 0.03$. We get respectively 0.29723 and 0.03014 which indeed confirms our observation above.
We would like to compare the two new models implemented with each other to see the effect of the additional feature of non parallel shifts on the price. Following the advice in [12], the calibration of κ to vanilla swaptions shows that it is best chosen between -1% and 6%. We take $\kappa = 3\%$ as it is a sensible parameter with the market data.

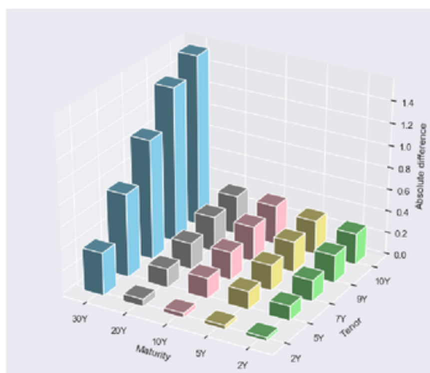


Figure 18: Absolute difference (in percentage) between CMS swap prices using the parallel shifts model and the non parallel shifts model

Figure (18) shows that for if we fix the maturity and increase the tenors, this will lead to higher absolute differences in the prices between the model with parallel shifts and non parallel shifts. We will provide a mathematical justification for this observation in what follows.
For shorter underlying swap tenors, $s - s_0$ is small. Furthermore, multiplying $s - s_0$ by $\kappa = 0.03$, the product is very close to 0. Hence we can use a second order Taylor series approximation:

$$e^x \approx 1 + x + \frac{x^2}{2}$$

. We get:

$$\begin{aligned} h(s) - h(s_0) &\approx \frac{1}{\kappa}\left[1 - \left(1 - \kappa(s - s_0) + \kappa^2(s - s_0)^2\right)\right] \\ &= (s - s_0) - \kappa(s - s_0)^2 \end{aligned}$$

The difference $s - s_0$ is small for shorter tenors, so if we square it multiply it by a small number (κ is usually strictly less than 0.06), we get that $\kappa(s - s_0)^2 \approx 0$. Thus,

$$h(s) - h(s_0) \approx (s - s_0)$$

Hence for shorter tenors, this model is very close to the model with parallel shifts, which explains the smaller differences in figure 15. However, for longer tenors, the value of $s - s_0$ is larger, so the error term $\kappa(s - s_0)^2$ is not close to 0. This being said, the argument above that uses a Taylor Series approximation does not apply and the difference between $h(s) - h(s_0)$ and $s - s_0$ is larger. Combining the arguments above together, this could be a possible explanation of our observation that the absolute difference between the two models increases with respect to the tenors.

7 Conclusion and Further Reseacrh

In this thesis, we have introduced the replication method and the notion of convexity adjustment that provide a good background for the reader, since they are widely used in the pricing of several derivatives in the financial market. We specified the use of these concepts for the CMS linked products, namely the CMS swap, cap and floor. We presented progressively the theoretical and the practical side of different approaches that are used to achieve the pricing of CMS. First, prices are obtained under Black's framework. The second method we implemented is the Vanna-Volga approach. Finally we used the SABR model for our purpose of CMS pricing. We provided the reader with all the details of the joint calibration procedure which is core part before the of this model. After implementing each method separately to price CMS swaps, we compared them and used the three approaches to price CMS caps. Prices were very close for the Vanna-Volga and the SABR model, but they were different than the price computed under Black's framework, due to the presence of a swaption volatility smile in the market.

Last but not least, we changed the standard bond-math approximation that is used by most practitioners to see how this would impact the prices of CMS. It would be good to take this contribution into consideration as the slope of the yield curve brings a lot of information especially for CMS-linked products. For instance, the yield curve is currently flattening which makes CMS contracts attractive today as mentioned in [16]. The reason is that the yield curve will eventually steepen in the future and the CMS swap's value will be positive compared to the market price. This situation is comparable to the period 2005-2007 when the yield curve was also flat. The peculiarity of the flat yield curve made the swaps receiving CMS prominent and trendy at that time [17]. However, despite being comparable to the 2005-2007 period, the current situation is not the same. The main difference is that the Federal Reserve is unlikely to cut short term rates in 2019. Hence, we should be aware that there is a mitigating risk that the slope of the yield curve may not increase again soon enough.

A further query is the announcement of the discontinuation of LIBOR by the end of 2021. A significant work should be undertaken to be prepared for a transition away of LIBOR. Practitioners sill do not have a clear idea of what will happen to trades that indirectly reference LIBOR such as CMS rates from LIBOR swaps, as is mentioned in [22]. Currently, there is no clear definition provided on how CMS payouts will be determined. Hence the main question we will consider in our future research is to look closely at what will happen to CMS LIBOR-based products after the discontinuation.

References

[1] A. Castagna, F. Mercurio and M. Tarenghi. *Smile consistent CMS adjustments in closed form: introducing the Vanna-Volga approach*, 2007.

[2] A. Castagna, F. Mercurio *The Vanna-Volga method for implied volatilities: tractability and robustness*, 2007.

[3] A. Castagna, F. Mercurio. *Consistent FX pricing*. Banca IMI,Milan, 2007.

[4] P.S. Hagan. *Convexity Conundrums: Pricing CMS Swaps, Caps, and Floors*. Wilmott magazine, 38-44, 2003.

[5] F. Mercurio and A. Pallavicini. *Swaption skews and convexity adjustments*. Banca IMI,Corso Matteotti, 20121 Milano, Italy, 2006.

[6] A. Pelsser. *Mathematical foundation of convexity correction. Quantitative Fi- nance* 3, 59-65,2003.

[7] P.S. Hagan, D. Kumar, A. Lesniewski and D.E. Woodward. *Managing Smile Risk*. Wilmott Magazine, 84-108, 2002.

[8] N. Karouzakis, J. Hatgioannides and C. Andriosopoulos. *Convexity Adjustment for Constant maturity Swaps in a Multi-Curve Framework*. Annals of Operations Research, 2017.

[9] Y. Shkolnikov. *Generalized Vanna-Volga Method and Its Applications*. NumeriX Quantitative Research, 2009.

[10] M. H.A. Davis. *The dupire formula*. Imperial College London.

[11] D. Brigo and F. Mercurio. *Interest Rates Models Theory and Practice with Smile*, Inflation and Credit, Springer,2007

[12] P.S. Hagan. *Evaluating and hedging exotic swap instruments via LGM*, 2015.

[13] A. Lipton. W. McGhee. *Universal barriers* Risk, p.81-85, 2002.

[14] T. Hastie. R. Tibshirani. J. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd edition, 2009

[15] U. Wystup. Vanna-Volga Pricing. MathFinance AG Waldems, Germany, 2008.

[16] S. Sohn. E. Jordahl. Constant Maturity Swaps: Is It Time to Reconsider Given the Current Flat Yield Curve? KaufmanHall Treasury and Capital Markets, 2019.

[17] E. Benhamou. *Swaps: Constant maturity swaps (CMS) and constant maturity Treasury (CMT) swaps.* Swaps Strategy, London, FICC, Goldman Sachs International, 2008.

[18] F. Diebold. G. Rudebusch. *Yield curve modeling and forecasting: The dynamic Nielson-Siegel approach.* Princeton University Press. Princeton and Oxford, 2013.

[19] P. Carr. D. Madan. *Optimal positioning in derivative securities.* Quantitative Finance Volume 1,19-37, 2001.

[20] L. Anderson. V. Piterbarg. *Interest Rate modeling, Volume 3: Products and Risk Management.* Atlantic Financial Press, 2010.

[21] P.S. Hagan, D. Kumar, A. Lesniewski and D. Woodward.*Arbitrage-Free SABR*, 2014.

[22] *Beyond Libor.* Risk.net Special report 2018.

[23] P. Hunt. J. Kennedy *Financial Derivatives in Theory and Practice*,2004.

[24] D. Williams. *Probability with Martingales.* Cambridge Univeristy Press.

Constant Maturity Swap Pricing

GRADEMARK REPORT

FINAL GRADE

/0

GENERAL COMMENTS

Instructor

PAGE 1

PAGE 2

PAGE 3

PAGE 4

PAGE 5

PAGE 6

PAGE 7

PAGE 8

PAGE 9

PAGE 10

PAGE 11

PAGE 12

PAGE 13

PAGE 14

PAGE 15

PAGE 16

PAGE 17

PAGE 18

PAGE 19

PAGE 20

PAGE 21

PAGE 22

PAGE 23

PAGE 24

PAGE 25

PAGE 26

PAGE 27

PAGE 28

PAGE 29

PAGE 30

PAGE 31

PAGE 32

PAGE 33

PAGE 34

PAGE 35

PAGE 36

PAGE 37

PAGE 38

PAGE 39

PAGE 40

PAGE 41

PAGE 42

PAGE 43

PAGE 44

PAGE 45

PAGE 46

PAGE 47

PAGE 48

PAGE 49

PAGE 50

PAGE 51

PAGE 52

PAGE 53

PAGE 54

PAGE 55

PAGE 56

PAGE 57

PAGE 58

PAGE 59