Delft University of Technology
Faculty of Electrical Engineering, Mathematics and Computer Science
Delft Institute of Applied Mathematics

The efficient pricing of CMS and CMS spread derivatives

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

for the degree

MASTER OF SCIENCE
in
APPLIED MATHEMATICS

by

Sebastiaan Louis Cornelis Borst
Delft, the Netherlands
September 2014

MSc THESIS APPLIED MATHEMATICS

"The efficient pricing of CMS and CMS spread derivatives"

Sebastiaan Louis Cornelis Borst

Delft University of Technology

Daily supervisor

Dr. N. Borovykh

Responsible professor

Prof. dr. ir. C.W. Oosterlee

Other thesis committee members

Dr. P. Cirillo

September 2014

Delft, the Netherlands

Abstract

Two popular products on the interest rate market are Constant Maturity Swap (CMS) derivatives and CMS spread derivatives. This thesis focusses on the efficient pricing of CMS and CMS spread derivatives, in particular the pricing of CMS and CMS spread options.

The notional values for these products are usually quite large, so even small errors when pricing these products can lead to substantial losses. Therefore, the pricing of these products has to be accurate. It is possible to use sophisticated models (e.g. Libor Market Model) to price these products, however the downside is that these models generally have high computational costs; they are not very efficient.

To efficiently price CMS options the Terminal Swap Rate (TSR) approach can be used. From this approach TSR models are obtained, we will consider four different TSR models. Two of these TSR models are established in the literature, the other two TSR models are developed in this thesis. The main advantages of a TSR model is that the computational costs are low and that it has good numerical tractability.

To price CMS spread options the copula approach is usually used. With the copula approach a pricing formula can be obtained for efficient valuations of CMS spread options. The copula that is considered in this thesis is the Gaussian copula. The TSR models are also a key component in the copula approach, because the marginal distributions are obtained with the help of a TSR model.

Furthermore, an alternative approach is considered for the pricing of CMS spread options. The CMS spread options are priced with a relatively simple stochastic volatility model, the displaced diffusion SABR model. The displaced diffusion SABR model is obtained by applying the Markovian projection method to a modification of a two-dimensional version of the well-established SABR model. The calibration of the two-dimensional SABR model is performed with the help of the TSR approach.

Acknowledgements

This thesis is the final product of my time as a student of the master study Applied Mathematics at Delft University of Technology. It is the end result of a collaboration of the TU Delft and Rabobank International. I would like to express my gratitude to those who have contributed in the process of writing this thesis.

First and foremost, I would like to thank the members of the examination committee. Dr. Natalia Borovykh, my daily supervisor at Rabobank International, for all helpful discussions and comments. I want to thank prof. dr. ir. Kees Oosterlee, my supervisor at TU Delft, for his excellent guidance during the last years of my study and for the useful comments on this thesis. Furthermore, I would like to thank dr. Pasquale Cirillo from TU Delft for being part of the thesis committee.

I would also like to thank the whole Derivatives Research and Validation (DR&V) team of Rabobank International. In particular, drs. Erik van Raaij and drs. Sacha van Weeren who hired me as an intern and provided me with an interesting topic for my master thesis.

Last but certainly not least, I would like to thank my family for the support they have given me such that I was able to finish my study.

Sebastiaan Borst
Utrecht, September 2014

Glossary

$A_{n,m}(t)$	Annuity factor or annuity
a.s.	Almost surely
α_0	Volatility parameter SABR model
$\alpha(\cdot)$	Annuity mapping function
β	Variance elasticity parameter SABR model
$B(t)$	Continuously compounded money-market account
$C(\cdot,\cdot)$	Bivariate copula function
$c(\cdot,\cdot)$	Bivariate copula density function
$\mathbb{E}^A[\cdot\|\mathcal{F}_t]$	Expectation operator given the information at time t with respect to the probability measure $\mathbb{Q}^A$
$\mathcal{F}_t$	Filtration
$G(\cdot,\cdot)$	Function of mean reversion
γ_{ij}	Cross-skew parameter displaced diffusion SABR model
K	Strike price
$L(t,T_n,T_{n+1})$, $L_n(t)$	Simply compounded forward Libor rate
Λ	Value of the convexity adjustment
ν	Volatility of volatility parameter SABR model
Ω	Sample space
$(\Omega,\mathcal{F},\mathbb{P})$	Probability space
$P(t,T)$	Time-t price of a zero-coupon bond with maturity T
$\mathbb{Q}^A$	Annuity measure
$\mathbb{Q}^B$	Risk neutral measure
$\mathbb{Q}^T$	T-forward measure
$\mathbf{R}$	Correlation matrix
$Re(z)$, $Im(z)$	Real and imaginary part of a complex number z
ρ	Correlation coefficient SABR model
$S_{n,m}(t)$	(Forward) swap rate at time t
$\sigma^{\mathrm{SABR}}(t)$	SABR implied volatility at time t
t	Time
T	Maturity time
$\tau(T_n,T_{n+1})$	Year fraction between time points T_n and T_{n+1}
V	Value of payoff function
ϱ	Correlation coefficient for two different swap rates
ς	Price difference CMS spread option computed with copula approach or DD SABR model vs reference model
1W, 1M, 1Y	1 week, 1 month, 1 year
ξ_{ij}	De-correlation parameter displaced diffusion SABR model
ζ	Price difference CMS option computed with TSR model vs reference model

AD	Anti-Dependence
ATM, ITM, OTM	At-The-Money, In-The-Money, Out-of-The-Money
bps	Basis points. A basis point is 1/100 of one percent, (1bp $= 10^{-4}$)
CDF	Cumulative Distribution Function
CMS	Constant Maturity Swap
CMSSO	CMS Spread Option
DD	Displaced Diffusion
FRA	Forward-Rate Agreement
GIC	Guaranteed Investment Contracts
Libor	London Interbank Offered Rate
LMM	Libor Market Model
MC	Monte Carlo
ODE	Ordinary Differential Equation
OTC	Over-The-Counter
PDE	Partial Differential Equation
PDF	Probability Density Function
SABR	Stochastic Alpha Beta Rho
SSE	Sum Squared Error
SMM	Swap Market Model
TSR	Terminal Swap Rate
SDE	Stochastic Differential Equation
USD	United States Dollar
ZCB	Zero-Coupon Bond

Contents

Chapter 1

Introduction

The global *Over-The-Counter (OTC)* market has increased at an incredible pace during the last decade. The asset class of interest rate contracts is the largest asset class of the OTC market by far. Shortly after the financial crisis in 2007-2008 the trading volume of OTC interest rate derivatives decreased. However, it wasn't long before the trading volume started to rise again.

In Table 1.1 the notional amount of the different asset classes for three different time periods is given; for more details we refer to [5].

	Notional amounts outstanding		
Risk Category/Instrument	Jun 2011	Jun 2012	Jun 2013
Foreign exchange contracts	64,698	66,672	73,121
Interest rate contracts	553,240	494,427	561,299
Equity-linked contracts	6,841	6,313	6,821
Commodity contracts	3,197	2,994	2,458
Credit default swaps	32,409	26,931	24,349
Other derivatives	46,498	42,059	24,860
Total contracts	706,884	639,396	692,908

Table 1.1: Amounts outstanding of over-the-counter (OTC) derivatives. By risk category and instrument. In Billions of USD.

The interest rates contracts at the end of June 2013 amounted to about 561 Trillion *United States dollar (USD)*, which is equivalent to over 81% of the total OTC traded derivatives market.

Although notional values are not necessarily very meaningful in the derivative markets for assessing the total exposure of a market, they give an indication for the trading volumes in specific derivative instruments. The notional values indicate somewhat the industry's interest in a certain type of derivative.

The majority of OTC derivative notional volumes are relatively simple products like interest rate swaps, interest rate options and forward rate agreements (FRAs). However, there are more exotic derivatives that are useful to companies and investors such as *Constant Maturity Swap (CMS) derivatives* and *CMS spread derivatives.*

1.1 Problem Exploration

CMS derivatives and CMS spread derivatives are very popular products nowadays because they enable investors to take a view on the level or the change in the level of the yield curve. The efficient pricing of CMS and CMS spread derivatives is the main objective of this thesis.

Some types of CMS derivatives are *CMS swaps*, *CMS caps* and *CMS floors*, these are options that are based on a *CMS rate*. The underlying is a *swap rate*, which is a long-term interest rate. CMS options are commonly traded in the market. CMS-based products are widely used by insurance companies and pension funds in their Asset & Liability management ([15] and [28]), because these institutions are very vulnerable to movements in the interest rates. CMS options provide suitable hedge requirements for insurance products like for instance *Guaranteed Investment Contracts (GICs)*. GICs are contracts that guarantee repayment of principal and a fixed or floating interest rate for a predetermined period of time. GICs are generally issued by life insurance companies and they are often bought for retirement plans. In particular, a CMS floor provides a hedge to GICs when interest rates are dropping and the insurance company has to make guaranteed fixed interest payments. Similarly, CMS caps provide a hedge in case interest rates are rising.

Some of the most common CMS spread derivatives are CMS spread options, *CMS spread caps* and *CMS spread floors*. A CMS spread option is similar to a regular cap/floor option. The difference is that whereas in a regular cap/floor the underlying is usually a reference rate, in a CMS spread cap/floor the underlying is the spread between two swap rates of different maturity. Banks typically use CMS spread options to hedge the CMS spread swaps that they have entered into with customers.

It is very important that the pricing of both CMS and CMS spread derivatives is efficient and accurate, since a small pricing error will lead to substantial losses due to the large notional values associated with these kind of products.

1.2 Research Question

The use of sophisticated models to price CMS and CMS spread derivatives is not always desirable due to too time-consuming calculations. Our aim is to develop models which can efficiently and accurately price CMS and CMS spread derivatives. Therefore, the following research question could be imposed:

Can one- or two-factor models be derived which can efficiently and accurately price CMS and CMS spread derivatives?

In order to answer this question, we will start by looking into the pricing of CMS derivatives. The CMS derivatives will be priced by making use of the *Terminal Swap Rate approach*, from which *Terminal Swap Rate models* are obtained. We will investigate the performance of several TSR models. After that we will focus on the pricing of CMS spread derivatives. The CMS spread derivatives will be priced by using the *copula approach* in which the earlier mentioned TSR models also play an important role. The CMS spread derivatives will also be priced with a stochastic volatility model. The results of the respective models will be compared, to infer which models could be possible interesting alternatives for the industry.

1.3 Outline

The outline of this thesis is as follows.

We start in Chapter 2 with some fundamentals we need when pricing interest rate derivatives. Readers that are already familiar with interest rate modeling may want to skip this chapter.

In Chapter 3 we look into the pricing of CMS derivatives with TSR models. We start with the explanation of CMS derivatives and the important concept of CMS convexity adjustment. Next, a replication method will be presented which can be used for the pricing of CMS derivatives. After that, the Terminal Swap Rate approach is explained together with another important concept the annuity mapping function. We will also discuss different TSR models. Two of the TSR models we will consider are described in the literature, the linear TSR model and the swap-yield TSR model. We will however also consider two new TSR models, we developed ourselves, that are based on interpolation. The performance of the respective TSR models will be investigated by means of several numerical experiments.

In Chapter 4 we look into the pricing of CMS spread derivatives by making use of the copula approach. We begin with the explanation of CMS spread derivatives, and discuss the pricing approach we are going to take to efficiently price CMS spread options. After that, copulas are discussed; in particular the Gaussian copula. Additionally, Sklar's Theorem which is a key component in the copula approach is presented. We will derive a one-dimensional pricing formula which can be used for the pricing of CMS spread options, as well as a simple Monte Carlo method that can be applied for the pricing of CMS spread options in case a Gaussian copula is used. The performance of the copula approach together with the TSR models will be investigated by performing several numerical experiments.

In Chapter 5 we will look deeper into the pricing of CMS spread options, when we will consider a stochastic volatility model. The stochastic volatility model that we will consider is the displaced diffusion SABR model. We first present a two-dimensional version of the SABR model that can be used for the pricing of CMS spread options. We present the Markovian projection method which is crucial to obtain the displaced diffusion SABR model. After that, the necessary steps to obtain the displaced diffusion SABR model from the two-dimensional SABR model are discussed in detail. The results of the displaced diffusion SABR model will be compared with the results obtained by the copula approach.

Chapter 6 summarizes the main results and conclusions that we have obtained regarding the efficient pricing of CMS and CMS spread derivatives.

Finally, Chapter 7 discusses possible further research directions that could be followed.

Chapter 2

Fundamentals of Interest Rate Modeling

2.1 Introduction

Before we can formulate the precise problem of pricing CMS and CMS spread derivatives we have to prepare ourselves with some fundamentals of interest rate derivative pricing to cover our upcoming needs. This chapter is organized as follows.

In Section 2.2 we start by reviewing some basic concepts of no-arbitrage pricing theory and we discuss the technique of *change of numéraire*, which as we will see plays a key role in the pricing of interest rate derivatives. In Section 2.3 we formulate some important definitions concerning interest rates. Next, in Section 2.4 we look at examples of pricing numéraires. In Section 2.5 we consider some basic interest rate derivatives. Finally, in Section 2.6 we discuss the concept of implied volatility.

This chapter is mainly based on [35, pp. 3-30, 167-207] and [9, pp. 1-40].

2.2 No-Arbitrage Pricing Framework

The concept of a numéraire is very important in the pricing of financial derivatives. The definition of a numéraire is given as follows:

Definition 2.2.1 (Numéraire). *A numéraire is any positive asset with price process $N(t)$ that pays no dividend. The numéraire $N(t)$ is used to discount other asset price processes, the relative price process of asset S is given by:*

$$S^N(t) \triangleq \frac{S(t)}{N(t)}. \tag{2.1}$$

So a numéraire can be seen as a reference asset that is chosen to normalize all other asset prices with respect to it. Another import concept is a so-called *equivalent martingale measure* that is defined as follows:

Definition 2.2.2 (Equivalent Martingale Measure). *Consider a continuous time framework within a compact time interval $[0, T]$. Let $\mathbb{Q}$ be a probability measure on $(\Omega, \mathcal{F})$, measure $\mathbb{Q}$ is called an equivalent martingale measure with numéraire $N(t)$, $t \in [0, T]$ if the following two properties are satisfied:*

i) Measure $\mathbb{Q}$ is equivalent to measure $\mathbb{P}$, that is $\mathbb{P}(A) = 0$ if and only if $\mathbb{Q}(A) = 0$, for every $A \in \mathcal{F}$.

ii) The relative price process $S^N(t)$ is a martingale under measure $\mathbb{Q}$, i.e.

$$\mathbb{E}^{\mathbb{Q}}\left[S^N(t) \middle| \mathcal{F}_s \right] = S^N(s), \quad s < t. \tag{2.2}$$

For more details we refer to [9, pp. 24-25]. Next we state the following two theorems:

Theorem 2.2.3 (**First Fundamental Theorem of Asset Pricing**)**.** *A financial market, on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, is arbitrage-free if and only if there exists at least one risk-neutral probability measure $\mathbb{Q}$, called an equivalent martingale measure, that is equivalent to the original (or actual) probability measure $\mathbb{P}$.*

Proof. For a proof of this theorem we refer to [42, pp. 228-232]. □

Theorem 2.2.4 (**Second Fundamental Theorem of Asset Pricing**)**.** *Let a financial market have at least one risk-neutral probability measure. Then the market model is complete if and only if the risk-neutral probability measure is unique.*

Proof. For a proof of this theorem we again refer to [42, pp. 232-234]. □

These two theorems provide the fundament for the no-arbitrage pricing framework, they ensure that prices are unique. The fundamental pricing formula presented in [35, pp. 9-11] is a result of these theorems and since it is such an important result we highlight it by listing it in the following lemma:

Lemma 2.2.5 (**Fundamental Pricing Formula**)**.** *Assume there exists an equivalent martingale measure $\mathbb{Q}$, then for each attainable*[1] *contingent claim $V(T)$, defined as a stochastic cash-flow at time T and modelled as an $\mathcal{F}_T$-measurable random variable*[2] *there exists a unique price $V(t)$, for each $0 \le t \le T$, given by*

$$V(t) = N(t)\mathbb{E}^N \left[\frac{V(T)}{N(T)} \middle| \mathcal{F}_t \right]. \tag{2.3}$$

Proof. For the proof we refer to [35, pp. 8-11]. □

Equation (2.3) enables us to calculate today's price of a derivative security in a no-arbitrage pricing framework.

Often it is possible to reduce the complexity of a pricing problem by an appropriate *measure transformation* by changing the numéraire. The price of any asset divided by a numéraire is a martingale (no drift) under the measure associated with that numéraire.

The so-called *Radon-Nikodym derivative* is the key concept to change from one measure to another. To change from one measure to another we will make use of Theorem 2.2.6.

Theorem 2.2.6 (**Change of Numéraire**)**.** *Let M be a numéraire and $\mathbb{Q}^M$ be the corresponding probability measure, equivalent to an initial measure $\mathbb{Q}^0$, such that the price of any traded asset, X, relative to M, is a martingale under measure $\mathbb{Q}^M$, i.e.*

$$\mathbb{E}^M \left[\frac{X(T)}{M(T)} \middle| \mathcal{F}_t \right] = \frac{X(t)}{M(t)}. \tag{2.4}$$

[1] This is a technical requirement. For a claim to be attainable there needs to exist a suitable self-financing replicating strategy. When the market is complete every contingent claim is attainable.

[2] The significance of V being $\mathcal{F}_T$-measurable is that it may depend on the whole path of the underlying in $[0, T]$, precisely because $\mathcal{F}_T$ contains all this information.

Let now N be an arbitrary numéraire. Then a probability measure $\mathbb{Q}^N$ exists, equivalent to $\mathbb{Q}^0$, such that the price of any attainable claim, Y, normalized by some quantity, N, is a martingale under measure $\mathbb{Q}^N$, i.e.

$$\mathbb{E}^N\left[\frac{Y(T)}{N(T)}\middle|\mathcal{F}_t\right] = \frac{Y(t)}{N(t)}. \tag{2.5}$$

The Radon-Nikodym derivative, which defines the measure $\mathbb{Q}^N$, is given by:

$$\left.\frac{d\mathbb{Q}^N}{d\mathbb{Q}^M}\right|_{\mathcal{F}_t} = \frac{N(T)M(t)}{N(t)M(T)}. \tag{2.6}$$

Proof. The proof is given in Appendix A.1. □

2.3 Definitions & Notations

Now we will define several concepts. The concepts that will be discussed are: tenor structure, zero-coupon bonds, Libor rates, swap rates and the money market account.

2.3.1 Tenor Structure

A *tenor structure* is an increasing sequence of maturity times (dates), usually roughly equidistantly spaced. We will consider the following tenor structure,

$$0 \leq T_0 < T_1 < T_2 < \ldots < T_N. \tag{2.7}$$

The *accrual factor* or *year fraction*, denoted by $\tau(t,T)$, measures the time between t and T (in years),

$$\tau(t,T) = T - t. \tag{2.8}$$

In practice there exists a great variety of so-called *day-count conventions* (e.g. Actual/365, 30/360, etc.), which differ according to product type and country. We refer to [9, pp. 4-6] for more information about day-count conventions. For clarity we will now simply use $\tau = T - t$, although later when we deal with real market data we will take the correct day-count conventions into account.

2.3.2 Zero-coupon bonds

So-called *zero-coupon bonds (ZCBs)*, also known as pure discount bonds, are the most basic products in the interest rate market. ZCBs have no payoff until their maturity date T and are therefore interesting tools for interest rate modeling and discounting future cash-flows. The formal definition of a ZCB is given as follows:

Definition 2.3.1 (Zero-Coupon Bond). *A zero-coupon bond (ZCB) is a basic interest rate product, whose value is denoted by $P(t,T)$, where $t \leq T$. A ZCB pays 1 currency unit at maturity T, i.e. $P(T,T) = 1$.*

The ZCB is one of the main building blocks for interest theory/models.

2.3.3 Libor Rates

Libor (London Interbank Offered Rate) rates are the benchmark interest rates at which banks can borrow unsecured funds from other banks in the interbank markets. The *simply compounded forward Libor rate* is defined as follows:

Definition 2.3.2 (Simply Compounded Forward Libor Rate). *Let the tenor structure be given by (2.7). The simply compounded forward Libor rate contracted at time t for the interval $[T_i, T_{i+1}]$ is given by:*

$$L(t, T_i, T_{i+1}) \triangleq \frac{1}{\tau(T_i, T_{i+1})} \left(\frac{P(t, T_i)}{P(t, T_{i+1})} - 1 \right). \tag{2.9}$$

We introduce the useful shorthand notation

$$L_i(t) = L(t, T_i, T_{i+1}). \tag{2.10}$$

We speak of a spot Libor rate denoted by $L_i(T_i) = L(T_i, T_i, T_{i+1})$ when $t = T_i$, i.e. when the present date coincides with the start of the period over which the interest rate is effective.

2.3.4 Swap Rate

Given the tenor structure defined in (2.7) the *annuity factor* or *annuity* is defined as follows

$$A(t) \triangleq A_{n,m}(t) = \sum_{i=n}^{n+m-1} P(t, T_{i+1})\tau_i, \tag{2.11}$$

where n, m are any two integers satisfying $0 \leq n < N$, $m > 0$ and $n + m \leq N$. Note that $\tau_i = T_{i+1} - T_i$ is the year fraction for the time interval $[T_i, T_{i+1}]$.

The *forward swap rate* or *swap rate* is defined as follows:

Definition 2.3.3 (Swap Rate). *The swap rate is given by:*

$$S(t, T_n, T_{n+m}) \triangleq S_{n,m}(t) = \frac{P(t, T_n) - P(t, T_{n+m})}{A(t)}, \tag{2.12}$$

where $A(t)$ given by (2.11). If no confusion arises the short-hand notation $S(t)$ is used instead.

For practical purposes it is sometimes more convenient to use a different expression.

Lemma 2.3.4 (Alternative Expression Swap Rate). *The swap rate $S_{n,m}(t)$ can be expanded into a weighted sum of forward rates to get the following expression*

$$S(t) = S_{n,m}(t) = \frac{\sum_{i=n}^{n+m-1} \tau_i P(t, T_{i+1}) L_i(t)}{A_{n,m}(t)}, \tag{2.13}$$

with $A_{n,m}(t)$ and $L_i(t)$ given by (2.11) and (2.10).

Proof. The proof is given in Appendix A.2. □

2.3.5 Money Market Account

The continuously compounded money market account can be seen as a deposit of 1 unit of currency that earns the instantaneous risk-free rate, r, and it satisfies the following stochastic differential equation (SDE)

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

where $r(t)$ is the short rate. Solving this equation we obtain

$$B(t) = e^{\int_0^t r(z)dz}.$$

2.4 Examples of Pricing Numéraires

There are several well-known numéraires which are very useful for interest rate derivative pricing. We are going to review some of these numéraires and their associated measures, since they play an important role in pricing CMS type products.

2.4.1 Money Market Account as the Numéraire

The first numéraire we will look into is the continuously compounded money market account, $B(t)$. This is the numéraire that defines the risk-neutral measure, denoted by $\mathbb{Q}^B$. The related expectation is denoted by $\mathbb{E}^B[\cdot]$. Using the fundamental pricing formula (2.3) we immediately get

$$V(t) = B(t)\mathbb{E}^B\left[\left.\frac{V(T)}{B(T)}\right| \mathcal{F}_t\right]. \tag{2.14}$$

Note that if we take $V(t) = P(t,T)$ then we get

$$P(t,T) = B(t)\mathbb{E}^B\left[\left.\frac{P(T,T)}{B(T)}\right| \mathcal{F}_t\right] \tag{2.15}$$

$$= \mathbb{E}^B\left[\left.\frac{B(t)}{B(T)}\right| \mathcal{F}_t\right] \tag{2.16}$$

$$= \mathbb{E}^B\left[\left.e^{-\int_t^T r(z)dz}\right| \mathcal{F}_t\right]. \tag{2.17}$$

Hence, we can formulate the following lemma presented in [35, pp. 172-173]:

Lemma 2.4.1 (Fundamental Bond Pricing Formula). *If there are no-arbitrage opportunities in the market, the time t price $P(t,T)$ of a T-maturity ZCB is given by*

$$P(t,T) = \mathbb{E}^B\left[\left.e^{-\int_t^T r(z)dz}\right| \mathcal{F}_t\right]. \tag{2.18}$$

2.4.2 ZCB as the Numéraire

The T-forward measure uses a T-maturity ZCB as the numéraire, the related expectation is denoted by $\mathbb{E}^T[\cdot]$. We can change from the risk-neutral measure to the T-forward measure by applying Theorem 2.2.6 and get

$$\begin{aligned} V(t) &= B(t)\mathbb{E}^B\left[\left.\frac{V(T)}{B(T)}\right| \mathcal{F}_t\right] \\ &= B(t)\mathbb{E}^T\left[\left.\frac{V(T)}{B(T)}\frac{d\mathbb{Q}^B}{d\mathbb{Q}^T}\right| \mathcal{F}_t\right] \\ &= B(t)\mathbb{E}^T\left[\left.\frac{V(T)}{B(T)}\frac{B(T)P(t,T)}{B(t)P(T,T)}\right| \mathcal{F}_t\right] \\ &= P(t,T)\mathbb{E}^T\left[V(T)|\, \mathcal{F}_t\right]. \end{aligned} \tag{2.19}$$

Note that by changing to the T-forward measure we have obtained an easier expression. The $1/B(T)$ term inside the expectation operator is replaced by the related bond price that is outside of the expectation operator. Measure $\mathbb{Q}^T$ is called the T-forward measure because forward rates are martingales under the T-forward measure. An example of a forward rate is the simply compounded forward Libor rate. We will now show that the simply compounded forward Libor rate is indeed a martingale under the T-forward measure.

Lemma 2.4.2 (Forward Libor Rate under T_{i+1}-Forward Measure). *In the absence of arbitrage the simply compounded forward Libor rate for time interval $[T_i, T_{i+1}]$, $L(t, T_i, T_{i+1})$, is a martingale under the T_{i+1}-forward measure, $\mathbb{Q}^{T_{i+1}}$, i.e.*

$$\mathbb{E}^{T_{i+1}}[L(t, T_i, T_{i+1})|\, \mathcal{F}_s] = L(s, T_i, T_{i+1}), \quad \textit{for all } \ 0 \leq s \leq t \leq T_i < T_{i+1}. \tag{2.20}$$

Proof. Let $L(t, T_i, T_{i+1})$ be defined as in (2.9), then

$$\begin{aligned}
\mathbb{E}^{T_{i+1}}\left[L(t, T_i, T_{i+1})|\, \mathcal{F}_s\right] &= \mathbb{E}^{T_{i+1}}\left[\frac{1}{\tau(T_i, T_{i+1})}\left(\frac{P(t, T_i)}{P(t, T_{i+1})} - 1\right)\middle|\, \mathcal{F}_s\right] \\
&= \frac{1}{\tau(T_i, T_{i+1})}\mathbb{E}^{T_{i+1}}\left[\frac{P(t, T_{i+1}) - P(t, T_i)}{P(t, T_i)}\middle|\, \mathcal{F}_s\right] \\
&= \frac{1}{\tau(T_i, T_{i+1})}\left(\frac{P(s, T_i) - P(s, T_{i+1})}{P(s, T_{i+1})}\right) \qquad (2.21) \\
&= \frac{1}{\tau(T_i, T_{i+1})}\left(\frac{P(s, T_i)}{P(s, T_{i+1})} - 1\right), \\
&= L(s, T_i, T_{i+1}), \qquad (2.22)
\end{aligned}$$

for all $0 \leq s \leq t \leq T_i < T_{i+1}$. Equation (2.21) is obtained using the fact that $P(t, T_i)$ and $P(t, T_{i+1})$ are both tradeable assets divided by the numéraire $P(t, T_{i+1})$, so they must be martingales under the T_{i+1}-forward measure. And naturally their difference is also a martingale under the T_{i+1}-forward measure. □

2.4.3 Annuity Factor as the Numéraire

Note that the annuity is a linear combination of ZCBs, so the annuity can be taken as a numéraire. The associated measure is the so-called *annuity measure* or *swap measure*, denoted by $\mathbb{Q}^A$. The swap rate is a martingale under the annuity measure.

Lemma 2.4.3 (Swap Rate under Annuity Measure). *In the absence of arbitrage the swap rate $S_{n,m}(t)$ is a martingale under the annuity measure, $\mathbb{Q}^A$, i.e.*

$$\mathbb{E}^A\left[S_{n,m}(t)|\, \mathcal{F}_s\right] = S_{n,m}(s). \tag{2.23}$$

Proof. Let $S_{n,m}(t)$ be defined as in (2.12), then

$$\begin{aligned}
\mathbb{E}^A\left[S_{n,m}(t)|\, \mathcal{F}_s\right] &= \mathbb{E}^A\left[\frac{P(t, T_n) - P(t, T_{n+m})}{A_{n,m}(t)}\middle|\, \mathcal{F}_s\right] \\
&= \frac{P(s, T_n) - P(s, T_{n+m})}{A_{n,m}(s)}, \qquad (2.24) \\
&= S_{n,m}(s). \qquad (2.25)
\end{aligned}$$

Equation (2.24) is obtained using the fact that $P(t, T_n)$ and $P(t, T_{n+m})$ are both tradeable assets divided by the numéraire $A_{n,m}(t)$, so they must be martingales under the annuity measure. And their difference is also a martingale under the annuity measure. □

2.5 Basic Interest Rate Derivatives

We will discuss three main derivative products of fixed-income markets, namely swaps, caps/floors and swaptions.

2.5.1 Interest Rate Swaps

A *swap* is a generic term for an OTC derivative in which two counterparties agree to exchange one stream of cash flows against another stream of cash flows. These streams are called the *legs* of the swap. When the fixed leg is paid the swap is usually called a *payer swap*[3], when the fixed leg is received the swap is called a *receiver swap*. Swaps of different maturities between interest rate dealers and financial institutions are often traded to adjust interest risk positions of the parties involved, or to simply make bets on the future direction of interest rates. Swaps are also used by corporates to transform fixed rate obligations into floating ones, or vice versa.

A *plain vanilla fixed-for-floating interest rate swap* (a *plain vanilla swap* or just a *swap* if there is no confusion) is a swap in which one leg is a stream of fixed rate payments and the other is a stream of payments based on a floating rate, most often Libor. To formally define a fixed-floating swap a tenor structure needs to be specified. We assume the tenor structure given by (2.7). In a fixed-floating swap with fixed rate K, one party (the fixed rate payer) pays the simple interest based on the rate K in return for simple interest payments computed from the Libor rate fixing on date T_n, for each period $[T_n, T_{n+1}]$, $n = 0, \dots, N-1$. The payments are exchanged at the end of each period, i.e. at time T_{n+1}. In practice, the payments are netted. This means that the cash flow only takes place in one direction each payment. From the perspective of the fixed rate payer, the next cash flow of the swap at time T_{n+1} is therefore given by[4]

$$\tau_n(L_n(T_n) - K), \tag{2.26}$$

corresponding to the interest rate $L(T_n, T_{n+1})$ fixing at time T_n. Dates when the Libor rates are observed are usually called *fixing dates*, dates when the payments occur are called *payment dates*.

By the fundamental pricing formula, Lemma 2.2.5, with the money market account, $B(t)$, as the numéraire the value of a swap is equal to the expected discounted value of its (netted) payment. We can formulate the following lemma:

Lemma 2.5.1 (Valuation Interest Rate Swap). *The value of the swap from the perspective of the fixed rate payer at time t, $0 \le t \le T_0$ is given by*[5]

$$\begin{aligned} V_{swap}(t) &= B(t) \sum_{n=0}^{N-1} \tau_n \mathbb{E}^B \left[\frac{L_n(T_n) - K}{B(T_{n+1})} \middle| \mathcal{F}_t \right], &(2.27)\\ &= A(t)(S(t) - K). &(2.28) \end{aligned}$$

Similarly,

$$V_{swap\text{-}rec}(t) = A(t)(K - S(t)), \tag{2.29}$$

where $A(t)$ is given by (2.11) and $S(t)$ is given by (2.13).

Proof. The proof is given in Appendix A.3. □

2.5.2 Caps and Floors

An *interest rate cap* is a derivative that allows one to benefit from low floating rates yet be protected from high rates. An *interest rate floor* on the other hand allows one to benefit from low floating rates yet be protected from high rates. Caps and floors are among the most liquidly traded interest rate derivatives in fixed-income markets.

---

[3] In this document when we talk about a swap we mean a payer swap unless specified otherwise.
[4] Here an unit notional is assumed, this assumption is made throughout the document.
[5] This is a somewhat idealized expression. For more details we refer to [35, pp. 224-226].

Formally a cap is a strip of *caplets*, call options on successive Libor rates, and similarly a floor is a strip of *floorlets*, put options on successive Libor rates. The time-T_{n+1} cash flows of caplets/floorlets are given by

$$V^n_{\text{caplet}} = \tau_n (L_n(T_n) - K)^+, \tag{2.30}$$
$$V^n_{\text{floorlet}} = \tau_n (K - L_n(T_n))^+. \tag{2.31}$$

Applying Lemma 2.2.5 with numéraire $B(t)$, the time-t value of the cap/floor covering the time interval $[T_0, T_N]$ is given by

$$V_{\text{cap}}(t) = B(t) \sum_{n=0}^{N-1} \tau_n \mathbb{E}^B \left[\left. \frac{(L_n(T_n) - K)^+}{B(T_{n+1})} \right| \mathcal{F}_t \right], \tag{2.32}$$
$$V_{\text{floor}}(t) = B(t) \sum_{n=0}^{N-1} \tau_n \mathbb{E}^B \left[\left. \frac{(K - L_n(T_n))^+}{B(T_{n+1})} \right| \mathcal{F}_t \right]. \tag{2.33}$$

To get easier expressions we will change numéraire. Changing to the T_{n+1}-forward measure for each period we get using Theorem 2.2.6,

$$\begin{aligned} V_{\text{cap}}(t) &= B(t) \sum_{n=0}^{N-1} \tau_n \mathbb{E}^{T_{n+1}} \left[\left. \frac{(L_n(T_n) - K)^+}{B(T_{n+1})} \frac{B(T_{n+1}) P(t, T_{n+1})}{B(t) P(T_{n+1}, T_{n+1})} \right| \mathcal{F}_t \right] \\ &= \frac{B(t)}{B(t)} \sum_{n=0}^{N-1} \tau_n \mathbb{E}^{T_{n+1}} \left[(L_n(T_n) - K)^+ P(t, T_{n+1}) \middle| \mathcal{F}_t \right] \\ &= \sum_{n=0}^{N-1} \tau_n P(t, T_{n+1}) \mathbb{E}^{T_{n+1}} \left[(L_n(T_n) - K)^+ \middle| \mathcal{F}_t \right]. \end{aligned} \tag{2.34}$$

Similarly,

$$V_{\text{floor}}(t) = \sum_{n=0}^{N-1} \tau_n P(t, T_{n+1}) \mathbb{E}^{T_{n+1}} \left[(K - L_n(T_n))^+ \middle| \mathcal{F}_t \right]. \tag{2.35}$$

As such we represent caps and floors as baskets of European calls (caplets) and puts (floorlets) on Libor forward rates.

2.5.3 Swaptions

European *swaptions* as the name suggests are European options on interest rate swaps. A European swaption gives the holder the right, but not the obligation, to enter a swap at a future date at a given fixed date. A *payer swaption*[6] is an option to pay the fixed leg on a fixed-floating swap; a *receiver swaption* is an option to receive the fixed leg.

If we assume the underlying swap starts on the expiry date T_0 of the option, the payoff for the swaption at time T_0 is given by

$$V_{\text{swaption}}(T_0) = (V_{\text{swap}}(T_0))^+ \tag{2.36}$$
$$= \left(\sum_{n=0}^{N-1} \tau_n P(T_0, T_{n+1}) (L_n(T_0) - K) \right)^+. \tag{2.37}$$

[6] When we talk about a swaption we mean a payer swaption unless specified otherwise.

As before, applying Lemma 2.2.5 with numéraire $B(t)$ gives us the following value at an intermediate time t, $t < T_0$:

$$V_{\text{swaption}}(t) = B(t)\mathbb{E}^B\left[\left(\frac{\sum_{n=0}^{N-1}\tau_n P(T_0,T_{n+1})(L_n(T_0)-K)}{B(T_0)}\right)^+ \middle| \mathcal{F}_t\right]. \tag{2.38}$$

Note that the payoff does not only depend on the evolution of the individual Libor rates, as was the case with caplets/floorlets, but also depends on the joint behavior of the rates.

Using expression (2.28) we get

$$V_{\text{swaption}}(t) = B(t)\mathbb{E}^B\left[\frac{1}{B(T_0)}A(T_0)(S(T_0)-K)^+ \middle| \mathcal{F}_t\right]. \tag{2.39}$$

We formulate the following useful lemma for the valuation of payer and receiver swaptions:

Lemma 2.5.2 (**Valuation Payer and Receiver Swaption**). *The value of the payer and receiver swaptions at time t can be written as:*

$$V_{\textit{swaption-pay}}(t) = A_{n,m}(t)\mathbb{E}^A\left[(S_{n,m}(T_n)-K)^+\middle|\mathcal{F}_t\right], \tag{2.40}$$
$$V_{\textit{swaption-rec}}(t) = A_{n,m}(t)\mathbb{E}^A\left[(K-S_{n,m}(T_n))^+\middle|\mathcal{F}_t\right]. \tag{2.41}$$

Proof. The proof is given in Appendix A.4. □

From expressions (2.40) and (2.41) it is clear that under the annuity measure a payer swaption is basically a call option on the forward swap rate, with the strike K being equal to the fixed rate of the swap. Similarly, a receiver swaption can be interpreted as a put option on the forward swap rate.

2.6 The Implied Volatility Smile

The market convention is to quote caplets/floorlets and swaption prices in terms of the implied value of volatility which sets the Black model price equal to the market price.

Specifically, suppose that the market price of an option with strike K and maturity T is known, denote this price by $V(t,S,T,K)$. The time t implied volatility function $\sigma_{\text{imp}}(t,S,T,K)$ is defined as the solution to

$$\text{Black}(S,K,\sigma_{\text{imp}}(t,S,T,K)\sqrt{T-t},w) = V(t,S,T,K), \tag{2.42}$$

where $\text{Black}(S,K,v,w)$ is given in the following lemma:

Lemma 2.6.1 (**Black's Formula**). *Let $F(t)$ be an asset with dynamics given by*

$$dF(t) = \sigma F(t)dW(t), \tag{2.43}$$

so the underlying follows a geometric Brownian motion. The price of a European call/put option, denoted by $V(t)$, with strike K, volatility σ and maturity T is given by $V(t) = Black(F,K,v,w)$; where Black's formula is given by

$$Black(F,K,v,w) \triangleq wF\Phi(wd_1(F,K,v)) - wK\Phi(wd_2(F,K,v)), \tag{2.44}$$

with $w = 1$ for call options and $w = -1$ for put options; and

$$
\begin{aligned}
d_1(F, K, v) &= \frac{\log(F/K) + v^2/2}{v}, \\
d_2(F, K, v) &= \frac{\log(F/K) - v^2/2}{v}, \\
v &= \sigma\sqrt{T - t}, \\
\Phi(x) &= \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du.
\end{aligned}
$$

Proof. For the proof we refer to [7]. □

The mapping $K \mapsto \sigma_{\text{imp}}(t, S, T, K)$ is called an *implied volatility smile*[7]. Market participants prefer to quote prices in terms of implied volatilities, as volatilities tend to be more stable over time. Implied volatilities are also used to price options with non-quoted strikes and to compute hedge parameters.

2.6.1 SABR Model and Hagan's Formula

The SABR model is a four parameter stochastic volatility model that is introduced to accommodate the volatility smile in derivatives markets, [20]. The name is an abbreviation of 'Stochastic Alpha Beta Rho' referring to the three key parameters of the model. In the interest rate market the SABR model has become an industrial standard for quoting, interpolating and extrapolating the prices of plain-vanilla products. Its popularity is due to

- the analytical approximations for the implied volatilities,
- the intuitive meaning of the parameters of the model.
- the capability to (re-)produce a wide range of skew/smile patterns,
- realistic implied volatility smile dynamics with respect to changes in the forward level.
- the capability to calculate hedge parameters for every strike.

The SABR model is defined by the following system of SDEs

$$
\begin{aligned}
dF_t &= \alpha_t F_t^{\beta} dW_t^1, \quad F_0 = F, \\
d\alpha_t &= \nu \alpha_t dW_t^2, \quad \alpha_0 = \alpha, \\
\langle dW_t^1, dW_t^2 \rangle &= \rho dt
\end{aligned}
\tag{2.45}
$$

where F_t is the forward price with F being today's forward price, α_t is the volatility with $\alpha > 0$, β is the variance elasticity with $0 \le \beta \le 1$, ν is the volatility of the volatility with $\nu > 0$ and ρ is the correlation coefficient.

From (2.45) Hagan derived a formula to calculate the Black implied volatility. The derivation is based on singular perturbation techniques, for the details we refer to [20]. The main attractive

[7] In case the smile is monotonically downward or upward sloping, i.e. U-shaped, it is often called a *volatility skew*. *Skew* then refers to the slope of the smile.

feature of the SABR model is exactly this asymptotic approximation formula for the implied volatility, also commonly referred to as *Hagan's formula*. Hagan's formula is given by:

$$\sigma^{\text{SABR}}(F,K) = \frac{\alpha}{(FK)^{(1-\beta)/2}\left(1+\frac{(1-\beta)^2}{24}\log^2\left(\frac{F}{K}\right)+\frac{(1-\beta)^4}{1920}\log^4\left(\frac{F}{K}\right)+\dots\right)}\cdot\frac{z}{x(z)}$$
$$\cdot\left[1+\left(\frac{(1-\beta)^2}{24}\frac{\alpha^2}{(FK)^{1-\beta}}+\frac{1}{4}\frac{\rho\beta\nu\alpha}{(FK)^{(1-\beta)/2}}+\frac{2-3\rho^2}{24}\nu^2\right)T+\dots\right], \qquad (2.46)$$

where

$$z = \frac{\nu}{\alpha}(FK)^{(1-\beta)/2}\log\left(\frac{F}{K}\right),$$

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

F denoting today's forward price and K denotes the strike. For the special case of at-the-money (ATM) options (options with $K=F$) the σ_{imp} formula reduces to

$$\sigma^{\text{SABR}}_{\text{ATM}} = \frac{\alpha}{F^{(1-\beta)}}\left[1+\left(\frac{(1-\beta)^2}{24}\frac{\alpha^2}{F^{2-2\beta}}+\frac{1}{4}\frac{\rho\beta\nu\alpha}{F^{(1-\beta)}}+\frac{2-3\rho^2}{24}\nu^2\right)T+\dots\right]. \qquad (2.47)$$

So the SABR model allows for an accurate analytical approximation for the volatility at a given strike and the market price and the market risks can be obtained immediately from Black's formula.

The option prices are usually quoted in the market in terms of implied Black-Scholes volatilities, the quotes are available for a set of strikes. Thus is order to calibrate the SABR model to the market the SABR parameters are adjusted such that the model implied volatilities (given by Hagan's formula (2.46)) are as close as possible to the market implied volatilities. Note that generally the beta parameter is fixed in advance (from e.g. historical consideration), alpha is used to fit the ATM volatility (see (2.47)) exactly, and the parameters rho and nu are used to fit the remaining market quotes using some optimisation method.

To get better insight in the influence of the parameters on the implied volatility smile a simple specific base scenario is chosen that generates an implied volatility smile. The parameters for the base scenario are given by

$$F_0 = 0.025, \quad \alpha_0 = 0.03, \quad \beta = 0.6, \quad \nu = 0.025, \quad \rho = -0.3, \qquad (2.48)$$

and $T=1$. Next each individual parameter is investigated by shifting the parameter both up and down. The resulting implied volatility smiles are given in Figure 2.1.

Analyzing Figure 2.1 we notice the following interesting results:

- Shifting F_0 up (F_0 high $= 0.030$) or down (F_0 low $= 0.015$) we see that the forward price and the smile move in the same direction. When the forward price increases the volatility shifts to higher prices; smile moves to the right. Similarly when the forward prices decrease the volatility smile shifts to lower prices; the smile moves to the left.

- Shifting α_0 up (α_0 high $= 0.05$) or down (α_0 low $= 0.01$) results in the entire smile moving up or down. So α does not really influence the shape of the smile, but rather the vertical location.

- Shifting ν up (ν high $= 0.5$) or down (ν low $= 0.1$) we see that the curvature of the smile is influenced. The higher ν is the more convex the smile becomes.

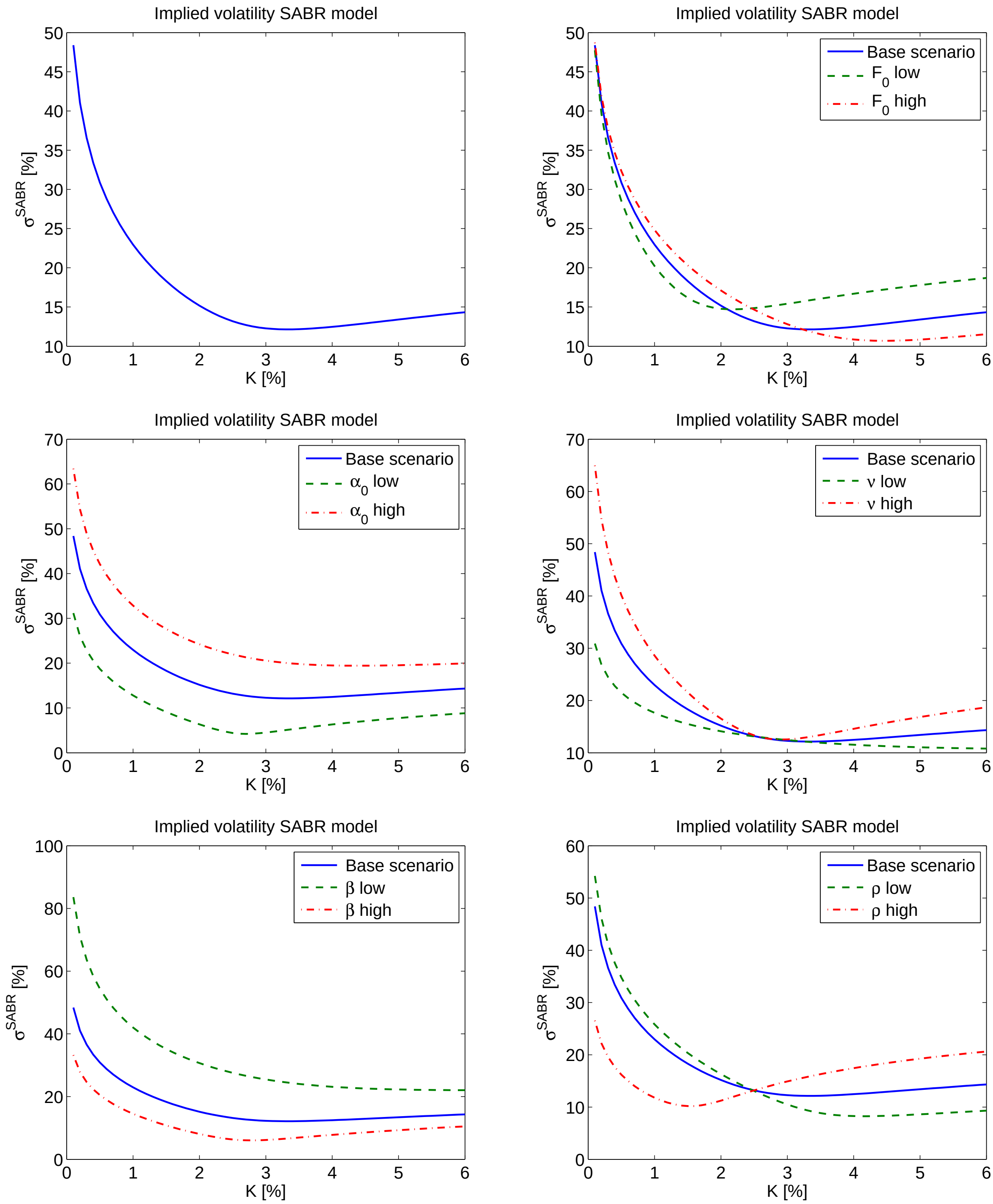


Figure 2.1: Impact of the SABR parameters on the implied volatility smile.

- Shifting β up (β high $= 0.60$) or down (β low $= 0.20$) we see that β also has an effect on the curvature of the smile, in particular on the side of the smile left from the ATM point. The higher β is, the steeper (the more convex) the smile becomes. We also note that the entire smile is shifted in the opposite direction of β.

- Shifting ρ up (ρ high $= 0.80$) or down (ρ low $= -0.80$) we see that changing the level of the correlation causes the smile to 'rotate' around the ATM point. Decreasing ρ causes a steeper smile, while increasing ρ causes the smile to flatten.

Thus, by making use of Hagan's formula we can obtain the market price for a given strike directly by substituting σ^{SABR} in Black's formula[8].

Remark 2.6.2. *Hagans's formula is only accurate for short time to maturity, [21] and [2].*

For more details about the SABR model we refer to [20].

2.6.2 Pricing Caps and Floors

It is common market practice to quote the value of a cap or a floor not in terms of its price but instead in terms of implied volatilities. Since we make use of Hagan's formula the implied volatilities are denoted by $\sigma_{n,N}^{\text{SABR}}$. Assuming the swap rate follows a geometric Brownian motion we can make use of Black's formula to obtain the time-t price of the cap/floor:

$$V_{\text{cap-Black}}(t) = \sum_{n=0}^{N-1} \tau_n P(t, T_{n+1}) \text{Black}(L_n(t), K, \sigma_{n,N}^{\text{SABR}} \sqrt{T_n - t}, 1), \tag{2.49}$$

$$V_{\text{floor-Black}}(t) = \sum_{n=0}^{N-1} \tau_n P(t, T_{n+1}) \text{Black}(L_n(t), K, \sigma_{n,N}^{\text{SABR}} \sqrt{T_n - t}, -1), \tag{2.50}$$

where $0 \leq t \leq T_n$. So for the price of the n-th caplet/floorlet we can write:

$$V^n_{\text{caplet-Black}}(t) = \tau_n P(t, T_{n+1}) \text{Black}(L_n(t), K, \sigma_{n,N}^{\text{SABR}} \sqrt{T_n - t}, 1), \tag{2.51}$$

$$V^n_{\text{floorlet-Black}}(t) = \tau_n P(t, T_{n+1}) \text{Black}(L_n(t), K, \sigma_{n,N}^{\text{SABR}} \sqrt{T_n - t}, -1). \tag{2.52}$$

2.6.3 Pricing Swaptions

From Lemma 2.4.3 we know that $S_{n,m}(t)$ is a martingale under the annuity measure, $\mathbb{Q}^{A_{n,m}}$. Same as with caps/floors it is common market practice to express market prices of swaptions in terms of implied volatilities, assuming the swap rate follows a geometric Brownian motion we can again make use of Black's formula and get

$$V_{\text{swaption-pay}} = A_{n,m}(t) \text{Black}(S_{n,m}(t), K, \sigma_{n,m}^{\text{SABR}}(S(t), K) \sqrt{T_n - t}, 1), \tag{2.53}$$

$$V_{\text{swaption-rec}} = A_{n,m}(t) \text{Black}(S_{n,m}(t), K, \sigma_{n,m}^{\text{SABR}}(S(t), K) \sqrt{T_n - t}, -1), \tag{2.54}$$

where $0 \leq t \leq T_n$.

[8] For each strike the same parameters are used in Black's formula

Chapter 3

Pricing CMS Derivatives with TSR models

3.1 Introduction

In this chapter we will focus on the pricing of CMS derivatives by making use of Terminal Swap Rate (TSR) models. This chapter is organized as follows.

We start in Section 3.2 with the explanation of CMS derivatives and the important concept of CMS convexity adjustment. Next, in Section 3.3, a replication method will be presented which can be used for the pricing of CMS derivatives. Section 3.4 introduces the TSR Approach. Next, in Section 3.5 we will consider several TSR models which can be used for the pricing of CMS derivatives. In Section 3.6 several numerical experiments are performed to investigate the performance of the respective TSR models. Section 3.7 concludes.

This chapter is mainly based on [35, pp. 206-207 and 336-338] and [37, pp. 709-739].

3.2 CMS Derivatives

A *CMS swap* is a fixed-for-floating swap, where, in contrast to a plain vanilla swap, the floating leg payments are based on CMS rates rather than on Libor rates.

For the pricing of CMS derivatives, it is necessary to compute the expectation of the future CMS rates under the forward measure that is associated with the payment date. However, the natural martingale measure of the CMS rate is the annuity measure. A so-called *convexity adjustment* arises because the expected value of the CMS rate under the forward measure differs from the expected value of the CMS rate under its natural swap measure with annuity as the numéraire.

Definition 3.2.1 (**CMS Convexity Adjustment**). *The CMS convexity adjustment is the difference between the expectation of the (function of the) CMS rate under the forward measure and the expectation of the (function of the) CMS rate under the annuity measure.*

Formally, let $S_{n,m}(\cdot)$ denote the m-period swap rate with first fixing date T_n, as defined in (2.13). Then an m-period (*payer*) CMS swap (linked to m-period rate) is given by

$$V_{\text{CMS-swap}}(t) = B(t) \sum_{n=0}^{N-1} \tau_n \mathbb{E}^B \left[\frac{S_{n,m}(T_n) - K}{B(T_{n+1})} \middle| \mathcal{F}_t \right]. \tag{3.1}$$

In order to simplify expression (3.1) we will change the valuation from the risk-neutral measure to the T_{n+1}-forward measure. Changing to the T_{n+1}-forward measure for each period we obtain using Theorem 2.2.6,

$$
\begin{aligned}
V_{\text{CMS-swap}}(t) &= B(t)\sum_{n=0}^{N-1}\tau_n\mathbb{E}^{T_{n+1}}\left[\frac{S_{n,m}(T_n)-K}{B(T_{n+1})}\frac{B(T_{n+1})P(t,T_{n+1})}{B(t)P(T_{n+1},T_{n+1})}\middle|\,\mathcal{F}_t\right] \\
&= \frac{B(t)}{B(t)}\sum_{n=0}^{N-1}\tau_n\mathbb{E}^{T_{n+1}}\left[\left(S_{n,m}(T_n)-K\right)P(t,T_{n+1})\middle|\,\mathcal{F}_t\right] \\
&= \sum_{n=0}^{N-1}\tau_n P(t,T_{n+1})\mathbb{E}^{T_{n+1}}\left[S_{n,m}(T_n)-K\middle|\,\mathcal{F}_t\right]. \qquad (3.2)
\end{aligned}
$$

CMS caps and *floors* are defined as strips of European options on CMS rates, just like regular caps and floors are strips of European options on Libor rates. The values of these derivatives are given by the following formulas,

$$V_{\text{CMS-cap}}(t) = \sum_{n=0}^{N-1}\tau_n P(t,T_{n+1})\mathbb{E}^{T_{n+1}}\left[\left(S_{n,m}(T_n)-K\right)^+\middle|\,\mathcal{F}_t\right], \tag{3.3}$$

$$V_{\text{CMS-floor}}(t) = \sum_{n=0}^{N-1}\tau_n P(t,T_{n+1})\mathbb{E}^{T_{n+1}}\left[\left(K-S_{n,m}(T_n)\right)^+\middle|\,\mathcal{F}_t\right]. \tag{3.4}$$

While plain vanilla swaps, caps and floors can be valued solely from knowledge of the term structure of interest rates, CMS swaps and CMS caps/floors require an interest rate model for valuation.

In this chapter we will mostly focus our attention on a single CMS-linked cash flow. The value of the CMS-linked cash flow is defined by:

$$V_{\text{gCMS}}(0) = \mathbb{E}^{T_p}[g(S(T_0))|\,\mathcal{F}_0], \tag{3.5}$$

where

$$g(s) = \begin{cases} s, & \text{in case of a swaplet,} \\ (s-K)^+, & \text{in case of a caplet,} \\ (K-s)^+, & \text{in case of a floorlet.} \end{cases} \tag{3.6}$$

Note that a CMS swap is a collection of CMS-linked cash flows with $g = s$. Similarly, CMS caps/floors are a collection of CMS-linked cash flows with $g = (s-K)^+$ or $g = (K-s)^+$.

Let us be clear on the notations, consider the case of a swaplet, then:

- $t = 0$ denotes the present date.
- T_0 denotes the start date of the reference swap (e.g., 1 year from now). This will also be the start of the accrual period of the swaplet.
- T_p denotes the payment date of the swaplet (e.g., 6 months from T_0). This will also be the end of the accrual period of the swaplet.
- T_N denotes the maturity date of the reference swap (say, 10 years from T_0).

Remark 3.2.2. *There is another date that plays a role here, namely the date on which the swap rate is fixed. This date is called the reset date of the swap. The reset date is used to correctly incorporate the volatility effect that is present in the market. Usually the reset date is two days before the start date of the swap. For clarity we will ignore this date in the rest of the thesis, although when we deal with real market data we take the reset dates into account.*

We know that for a continuous random variable Y and any function $h(\cdot)$ the expectation of $h(Y)$ is defined as,

$$\mathbb{E}[h(Y)] = \int_{-\infty}^{\infty} h(y)\psi(y)dy, \tag{3.7}$$

where $\psi(\cdot)$ is the *probability density function (PDF)* of Y. So, for the expectation in (3.5) we can write

$$\mathbb{E}^{T_p}\left[g(S(T_0)) \middle| \mathcal{F}_0 \right] = \int_{-\infty}^{\infty} g(s)\psi^{T_p}(s)ds, \tag{3.8}$$

where $\psi_p^T(\cdot)$ is the PDF of a swap rate in the T_p-forward measure. However, the PDF $\psi_p^T(\cdot)$ is not directly available. The PDF $\psi^A(\cdot)$ of a swap rate in the annuity measure on the other hand can be obtained from the market prices of swaptions ([35, pp. 737-739] and [9, pp. 448-449]) via

$$\psi^A(x) = \begin{cases} \frac{\partial^2 c(0,S(0),T_0,x)}{\partial x^2}, & \text{if } x \geq S(0), \\ \frac{\partial^2 p(0,S(0),T_0,x)}{\partial x^2}, & \text{if } x < S(0), \end{cases} \tag{3.9}$$

where

$$c(0, S(0), T_0, x) = \mathbb{E}^A\left[(S(T_0) - x)^+ \right], \tag{3.10}$$

$$p(0, S(0), T_0, x) = \mathbb{E}^A\left[(x - S(T_0))^+ \right]. \tag{3.11}$$

From the market we can imply the dynamics of $S(T_0)$ in the annuity measure so we are going to transform (3.5) to the annuity measure by once again applying Theorem 2.2.6,

$$V_{\text{gCMS}}(0) = \mathbb{E}^A\left[g(S(T_0)) \frac{P(T_0,T_p)A(0)}{P(0,T_p)A(T_0)} \middle| \mathcal{F}_0 \right] \tag{3.12}$$

$$= \frac{A(0)}{P(0,T_p)}\mathbb{E}^A\left[\frac{P(T_0,T_p)}{A(T_0)} g(S(T_0)) \middle| \mathcal{F}_0 \right]. \tag{3.13}$$

The CMS convexity adjustment is given by:

$$\Lambda_{\text{gCMS}}(0) \triangleq \mathbb{E}^{T_p}\left[g(S(T_0)) \middle| \mathcal{F}_0 \right] - \mathbb{E}^A\left[g(S(T_0)) \middle| \mathcal{F}_0 \right]. \tag{3.14}$$

The difficulty in calculating the expectation in (3.13) stems from the term

$$\frac{P(T_0,T_p)}{A(T_0)}, \tag{3.15}$$

since it depends on the joint distribution of a whole set of interest rates. In order to compute this expectation generally a term-structure model is used.

However, in this thesis we will use a TSR model that approximates the term $P(T_0,T_p)/A(T_0)$ with a so-called *annuity mapping function*, denoted by $\alpha(S(T_0))$. The way to obtain such an annuity mapping function will be discussed in Section 3.4, for the moment we will simply assume that such a function can be found. Hence, expression (3.13) can be written as

$$V_{\text{gCMS}}(0) = \frac{A(0)}{P(0,T_p)}\mathbb{E}^A\left[\alpha(S(T_0)) g(S(T_0)) \middle| \mathcal{F}_0 \right], \tag{3.16}$$

where $\alpha(S(T_0))$ is an annuity mapping function.

In order to calculate (3.16) we will make use of the *replication method* which we will present in the next section.

3.3 Replication Method

The replication method is used to replicate the CMS payout by means of European swaptions of various strikes. This method is very popular in practice (e.g. [19] and [31]) because it takes the volatility smile effects into account. Therefore, it is sometimes referred to as the street-standard model for CMS convexity correction. For more information about the replication method we refer to [10].

Let us write (3.16) as follows:

$$V_{\text{gCMS}}(0) = \frac{A(0)}{P(0,T_p)}\mathbb{E}^A\left[f(S(T_0))|\,\mathcal{F}_0\right], \tag{3.17}$$

where $f(S(T_0)) = \alpha(S(T_0))g(S(T_0))$. We can write the expectation as an integral over the density function

$$\mathbb{E}^A\left[f(S(T_0))|\,\mathcal{F}_0\right] = \int_{-\infty}^{\infty} f(x)\psi^A(x)dx, \tag{3.18}$$

where $\psi^A(\cdot)$ is given in (3.9).

The way to calculate expression (3.18) is formulated in Lemma 3.3.1.

Lemma 3.3.1 (**Replication Method for CMS Options**). *Let $f(\cdot)$ be defined on the interval $[a,b]$[1]. The calculation of $\mathbb{E}^A\left[f(S(T_0))|\,\mathcal{F}_0\right]$ in (3.17) is subdivided into three different cases depending on the value of the swap rate, $S(0)$, compared to the boundary conditions, a and b; namely:*

Case 1: If $S(0) < a$,

$$\begin{aligned}\mathbb{E}^A\left[f(S(T_0))|\,\mathcal{F}_0\right] = f(b)&\frac{\partial c(0,S(0),T_0,b)}{\partial x} - f(a)\frac{\partial c(0,S(0),T_0,a)}{\partial x}\\ &- f'(b)c(0,S(0),T_0,b) + f'(a)c(0,S(0),T_0,a)\\ &+ \int_a^b f''(x)c(0,S(0),T_0,x)dx.\end{aligned} \tag{3.19}$$

Case 2: If $S(0) > b$,

$$\begin{aligned}\mathbb{E}^A\left[f(S(T_0))|\,\mathcal{F}_0\right] = f(b)&\frac{\partial p(0,S(0),T_0,b)}{\partial x} - f(a)\frac{\partial p(0,S(0),T_0,a)}{\partial x}\\ &- f'(b)p(0,S(0),T_0,b) + f'(a)p(0,S(0),T_0,a)\\ &+ \int_a^b f''(x)p(0,S(0),T_0,x)dx.\end{aligned} \tag{3.20}$$

Case 3: If $a \leq S(0) \leq b$,

$$\begin{aligned}\mathbb{E}^A\left[f(S(T_0))|\,\mathcal{F}_0\right] = f(S(0)) &- f(a)\frac{\partial p(0,S(0),T_0,a)}{\partial x} + f(b)\frac{\partial c(0,S(0),T_0,b)}{\partial x}\\ &+ f'(a)p(0,S(0),T_0,a) - f'(b)c(0,S(0),T_0,b)\\ &+ \int_a^{S(0)} f''(x)p(0,S(0),T_0,x)dx + \int_{S(0)}^b f''(x)c(0,S(0),T_0,x)dx.\end{aligned} \tag{3.21}$$

[1]This is for numerical reasons (among others).

Here $p(0, S(0), T_0, x)$ and $c(0, S(0), T_0, x)$ are defined by (3.10) and (3.11).

Proof. The proof is given in Appendix A.5. □

Remark 3.3.2. *The minimum strike, K_{min}, or maximum strike, K_{max}, are chosen based on numerical considerations. The values of the boundaries a, b differ depending on the type of CMS option.*

- *In case of a CMS caplet: $a = K$, where K is the given strike, and $b = K_{max}$.*
- *In case of a CMS floorlet: $b = K$, where K is the given strike, and $a = K_{min}$.*
- *In case of a CMS swaplet: $a = K$, where $K_{min} = K = 0$ and $b = K_{max}$.*

Remark 3.3.3. *The values of the call and put options can be obtained by making use of market data and Black's formula. To incorporate the volatility smile we make use of Hagan's formula, (2.46).*

Note that in order to evaluate formula (3.18) we still need to specify the functional form of $f(\cdot)$ and calculate its first and second derivatives. Function $f(\cdot)$ can be specified in different ways depending on the chosen approach. As stated previously we will use the TSR approach.

3.4 TSR Approach

In this section we present the Terminal Swap Rate approach, which we will use to price CMS derivatives. This section is mainly based on [37, pp. 709-739] and [23, pp. 263-273].

It is well-known that European swaptions are relatively easy to price and this is due to the fact that only knowledge about the terminal distribution of a single swap rate, $S(T)$, in the annuity measure is necessary. In fact, all securities whose payoff can be expressed as deterministic functions of the swap rate are relatively easy to price. Unfortunately these kinds of payoffs are rare in the fixed income market; it is much more common that relatively simple payoffs depend not only on the swap rate but also mildly on certain discount bonds. Usually these bonds are observed on the same date. When multiple discount bonds are involved and knowledge of the distribution of the swap rate is not sufficient for valuation of the product one could choose to make use of a term structure model. The downside to this is that a term structure model has high computational costs. An alternative to avoid these high computational costs is the so-called TSR approach. The TSR approach can be used when the dependence on the additional discount bonds is sufficiently mild, so the swap rate is the rate that primarily determines the payoff. In the TSR approach the values of discount bonds on a date T are linked functionally to the driving swap rate $S(T)$.

A critical part of the TSR approach is that the developed models, so-called TSR models, capture precisely those properties of the market which are relevant to the derivative product being priced. The main advantage of this approach over other techniques is that it is guaranteed to price the new product accurately relative to existing products. Following this approach the developed model will have realistic properties and is built upon a solid theoretical basis. The characteristics of the model will usually be highly transparent, so it should be relatively easy to understand the model's strengths and weaknesses. So, the TSR approach is extremely useful in handling a range of liquid European derivatives that are not, strictly speaking, functions of a single rate, but can still be approximated as such. An example of these kinds of liquid European swap derivatives are CMS derivatives. Before we can present the TSR models which we can use to price CMS derivatives, we first have to discuss the basics of the TSR approach and the concept of an annuity mapping function.

3.4.1 Basics of the TSR Approach

In the TSR approach the swap rate $S(T)$ is the single fundamental state variable for the yield curve at time T. Let $\{P(T,M)\}_{M\geq T}$ be the discount bonds of various maturities, all observed at time T. A key feature of the TSR model is that it specifies a map

$$P(T,M) = \pi(S(T),M), \qquad M \geq T, \tag{3.22}$$

where $\pi(\cdot,M)_{M\geq T}$ is a collection of functions such that each discount factor is a known function of the swap rate.

In term structure models the relationship between the market rate $S(T)$ and the discount factors $\{P(T,M)\}_{M\geq T}$ follows directly from the model, as it is derived from no-arbitrage conditions. In order for a TSR model to have the same type of relationship, the TSR model must satisfy the following three conditions:

1. *No-arbitrage* condition;
2. *Consistency* condition;
3. *Realism* condition.

In order to satisfy the *no-arbitrage* condition a restriction must be imposed on the mapping functions $\pi(\cdot,M)_{M\geq T}$. The fundamental pricing formula (2.3) must reproduce the initial discount bond prices. Thus the following must hold for all $M \geq T$,

$$P(0,M) = A(0)\mathbb{E}^A\left[\frac{\pi(S(T),M)}{\sum_{n=0}^{N-1}\tau_n\pi(S(T),T_{n+1})}\right]. \tag{3.23}$$

We will refer to equation (3.23) as the *no-arbitrage* condition.

The *consistency* condition is obtained by observing that the swap rate $S(T)$ itself is a function of discount factors, which follows directly from expression (2.12). Therefore, the following expression must be satisfied for all[2] s:

$$s = \frac{1-\pi(s,T_N)}{\sum_{n=0}^{N-1}\tau_n\pi(s,T_{n+1})}. \tag{3.24}$$

We will refer to equation (3.24) as the *consistency* condition. The *consistency* condition ensures that all relevant relationships which hold in the market will also hold for the model.

The last condition, the *realism condition* has to do with monotonicity and limit properties. We call a set of mapping functions $\{\pi(\cdot,M)\}_{M\geq T}$ *reasonable* if it satisfies the following restrictions:

- For all s and $M \geq T$,
$$0 < \pi(s,M) \leq 1.$$
- For all s, $\pi(s,\cdot)$ is monotonic in M,
$$M_1 < M_2 \Leftrightarrow \pi(s,M_1) \geq \pi(s,M_2).$$
- The function $\pi(s,M)$ is continuous in (s,M).

[2]Here s is a state variable.

Not all of these restrictions are equally important. As an example it is possible to allow negative interest rates, which means that $\pi(s, M) > 1$ for some s, M. However, we cannot allow negative prices of bonds, i.e. having $\pi(s, M) < 0$ for some s, M is not possible.

The three requirements related to the *realism condition* do not define the mapping functions $\{\pi(\cdot, M)\}_{M \geq T}$ uniquely. However, they specify the functions uniquely within a particular parametric class. So, to obtain a concrete TSR model we can first select a particular parametric class for the functions $\{\pi(\cdot, M)\}_{M \geq T}$ and then choose functions within this class such that the model has the *no-arbitrage* and the *consistency* properties.

There are many different types of TSR models that can be used to price CMS derivatives. But first we look more closely to the concept of an annuity mapping function, since it plays a crucial role in CMS valuations.

3.4.2 Annuity Mapping Function

In Section 3.2 we introduced the concept of an annuity mapping function, but stopped just before we developed a method to determine it. The annuity mapping function, denoted by $\alpha(S(T_0))$, in (3.16) is defined to be the function that maps the term $P(T_0, T_p)/A(T_0)$ to a function of the swap rate $q(S(T_0))$. In general, the function $q(\cdot)$ is taken to be a payoff function. In our case $q(\cdot) = g(\cdot)$, where $g(\cdot)$ is given by (3.6). By making use of the *tower property of expectations* we can write for the expectation in the expression of the CMS-linked cash flow given by (3.16),

$$\begin{aligned}\mathbb{E}^A\left[\frac{P(T_0,T_p)}{A(T_0)}g(S(T_0))\middle|\mathcal{F}_0\right] &= \mathbb{E}^A\left[\mathbb{E}^A\left[\frac{P(T_0,T_p)}{A(T_0)}g(S(T_0))\middle|S(T_0)=s,\mathcal{F}_0\right]\middle|\mathcal{F}_0\right]\\ &= \mathbb{E}^A\left[\mathbb{E}^A\left[\frac{P(T_0,T_p)}{A(T_0)}\middle|S(T_0)=s,\mathcal{F}_0\right]g(S(T_0))\middle|\mathcal{F}_0\right]\\ &= \mathbb{E}^A\left[\alpha(S(T_0))g(S(T_0))|\,\mathcal{F}_0\right], \qquad (3.25)\end{aligned}$$

where function $\alpha(s)$ is given by

$$\alpha(s) = \mathbb{E}^A\left[\frac{P(T_0,T_p)}{A(T_0)}\middle|S(T_0)=s,\mathcal{F}_0\right]. \qquad (3.26)$$

This result gives rise to the following useful lemma, presented in [37, pp. 726-727]:

Lemma 3.4.1 (**Annuity Mapping Function for CMS-Linked Cash-Flow**). *The annuity mapping function $\alpha(s)$ in (3.16) may be written as a conditional expectation,*

$$\alpha(s) = \mathbb{E}^{\mathbb{A}}\left[\frac{P(T_0,T_p)}{A(T_0)}\middle|S(T_0)=s,\mathcal{F}_0\right]. \qquad (3.27)$$

This result is model-independent.

Lemma 3.4.1 clarifies the role of various methods of linking discount bond values to rates in order to value approximately single-rate derivatives. The methods that can be used are for instance TSR models and/or approximations inspired by term structure models. So, these methods can be seen as approximations to the true annuity mapping function defined by the conditional expected value in (3.27).

3.5 TSR Models

In this section we present four different TSR models. Two of these TSR models are established in the literature, namely the linear TSR model and the swap-yield TSR model. The other two TSR models we developed ourselves, and they are both based on interpolation. We will refer to these TSR models as the linear interpolation and log-linear interpolation TSR models.

3.5.1 Linear TSR Model

As stated previously in Section 3.4.1 a TSR model must satisfy the *no-arbitrage*, *consistency* and *realism* conditions. The important specifications of the linear TSR model are given by Lemma 3.5.1.

Lemma 3.5.1 (Specifications Linear TSR Model). *The linear TSR model is obtained by specifying*

$$\frac{\pi(s, M)}{\sum_{n=0}^{N-1} \tau_n \pi(s, T_{n+1})} = a(M)s + b(M), \quad M \geq T, \tag{3.28}$$

where $a(\cdot)$ and $b(\cdot)$ are deterministic functions. The three necessary conditions, the no-arbitrage, consistency and realism condition impose the following requirements:

- *Requirement 1:*
$$b(M) = \frac{P(0, M)}{A(0)} - a(M)S(0). \tag{3.29}$$

- *Requirement 2:*
$$b(T_0) = b(T_N), \tag{3.30}$$
$$a(T_0) = 1 + a(T_N). \tag{3.31}$$

- *Requirement 3:*
$$\sum_{n=0}^{N-1} \tau_n a(T_{n+1}) = 0, \tag{3.32}$$
$$\sum_{n=0}^{N-1} \tau_n b(T_{n+1}) = 1. \tag{3.33}$$

To complete the model specification we must proceed as follows:

1. *First choose coefficients $\{a(T_1), \ldots, a(T_N)\}$ subject to condition (3.32).*
2. *Next calculate $a(T) = a(T_0)$ from (3.31) and the remaining $a(M)$'s by linear interpolation of $\{a(T_1), \ldots, a(T_N)\}$.*
3. *Finally calculate all $b(M)$'s via (3.29).*

Proof. The proofs of identities (3.29), (3.30), (3.31), (3.32) and (3.33) are given in Appendix A.6. □

Due to the linear relationship between the market rate and annuity-discounted bonds one of the features of the linear TSR model is that it has good numerical tractability. The linear TSR model is rather popular in financial applications because of this nice feature, although the assumed linear relationship is not completely realistic. So, the decision in favor of the linear TSR model should be made on a case-by-case basis.

The linear TSR model (3.28) is considered to be a rather flexible model, since the coefficients

$$\{a(T_1), \ldots, a(T_N)\},$$

can essentially be selected independently, subject only to condition (3.32).

Remark 3.5.2. *Note that it is still not possible to obtain the values of the respective coefficients, since the obtained system of equations cannot be solved.*

Even if the system of equations could be solved we would not want to deal with all these coefficients individually. It is not clear what the financial implications would be for the various choices of the coefficients. The transparency, one of the nice features, of the linear TSR model would be lost. Therefore, we will parameterize $a(\cdot)$ by a single parameter that also has a financial interpretation.

Simple Version

Note that the coefficients $a(\cdot)$ essentially define the shape of the yield curve at time T for different levels of the state variable $S(T)$. We can estimate $a(\cdot)$ by making assumptions about the yield curve. In the *simple version* we assume that the time T yield curve is very low, which means that $S(T)$ is close to zero. We must have

$$\frac{P(T,M)}{A(T)} \approx \frac{1}{\sum_{n=0}^{N-1} \tau_n}. \tag{3.34}$$

Since we may write

$$\begin{aligned} \frac{P(T,M)}{A(T)} &= \mathbb{E}^A \left[a(M)S(T) + b(M) \middle| S(T) = 0 \right], &(3.35)\\ &= b(M), &(3.36) \end{aligned}$$

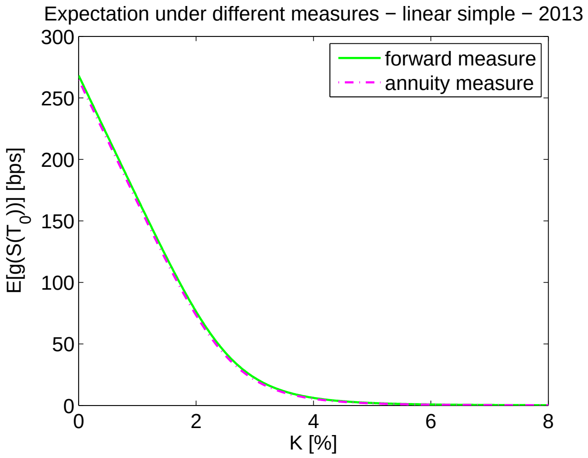
this suggests setting

$$\begin{aligned} a(M) &= \frac{1}{S(0)} \left(\frac{P(0,M)}{A(0)} - b(M) \right), &(3.37)\\ b(M) &= \frac{1}{\sum_{n=0}^{N-1} \tau_n}, &(3.38) \end{aligned}$$

where the equation for $a(M)$ follows from the no-arbitrage condition (3.29). However, this is a very crude approximation and in the next section we will consider a better way to estimate $a(\cdot)$, which we will call the *mean reversion version*.

Before we look into the mean reversion linear TSR model we will perform our first numerical experiment concerning TSR models. We will calculate the price and convexity adjustment of a CMS caplet for various strikes using the simple version of the linear TSR model. The CMS caplet we consider has as underlying the 10 year (10Y) CMS rate (swap rate) with 12 months (12M) frequency. The prices of CMS derivatives will generally be given in basis points (bps), because this is market standard. A basis point is 1/100 of one percent, $1\text{bp} = 10^{-4}$. The market data that was used is from 2013 and is given by Table B.1 and Table B.2; which can be found in Appendix B. The results are given in Figure 3.1.

Figure 3.1 shows that we can already observe that there is a difference between the expectation of $(S(T_0) - K)^+$ under the forward measure and the expectation of $(S(T_0) - K)^+$ under the annuity measure when the simple version of the linear TSR model is used. In the next paragraph we will compare the performance of the simple version of the linear TSR model to the performance of the mean reversion TSR model.



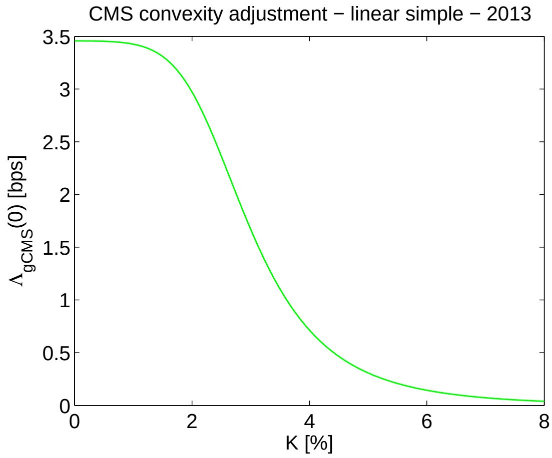


Figure 3.1: Expectation under different measures and CMS convexity adjustment of a CMS caplet on 10Y CMS rate with 12M frequency using the simple version of the linear TSR model.

Mean Reversion Version

In the mean reversion version the coefficients of the linear TSR model are connected to a mean reversion parameter. This has two major advantages, first of all it reduces the number of parameters that need to be specified and secondly the new single parameter has strong financial implications. Calibrating this mean reversion parameter is not straight-forward. The mean reversion parameter could be derived from prices of traded derivatives. The precise connection of $a(\cdot)$ to a mean reversion parameter is given by Lemma 3.5.3.

Lemma 3.5.3 (Mean Reversion Linear TSR Model). *In the mean reversion linear TSR model, the coefficients $a(\cdot)$ in (3.28) are connected to a mean reversion parameter, denoted by $\varkappa$, by the following relation*

$$a(M) = \frac{P(0,M)(\gamma - G(T,M))}{P(0,T_N)G(T,T_N) + A(0)S(0)\gamma}, \quad \text{for all } t \geq T, \tag{3.39}$$

where

$$\gamma = \frac{\sum_{n=0}^{N-1} \tau_n P(0,T_{n+1})G(T,T_{n+1})}{A(0)}, \tag{3.40}$$

and $G(\cdot,\cdot)$ is the function of mean reversion given by

$$G(t,T) = \begin{cases} \frac{1-e^{-\varkappa(T-t)}}{\varkappa}, & \text{for } \varkappa > 0, \\ T-t, & \text{for } \varkappa = 0. \end{cases} \tag{3.41}$$

The coefficients $b(\cdot)$ can be obtained directly by substituting $a(\cdot)$ in (3.29).

Proof. The proof is given in Appendix A.7. □

Linking $a(\cdot)$ to mean reversion leads to a more intuitive parametrization and also ensures better risk management.

As the second numerical experiment concerning TSR models we will compute the CMS convexity adjustment of a CMS caplet on 10Y CMS rate with 12M frequency, we will use the same market data as before [3]. We use different values for the mean reversion parameter. The results are shown in Figure 3.2.

[3] The market data from 2013 will be used for all TSR models in the upcoming subsections.

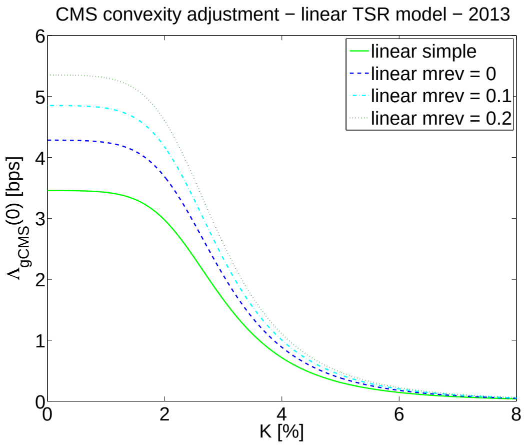


Figure 3.2: CMS convexity adjustment of a CMS caplet on 10Y CMS rate with 12M frequency using the linear TSR model.

In Figure 3.2 also the results of the simple version of the linear TSR model are given. We first note that the values of the CMS convexity adjustment computed with the different TSR models are relatively close. However, usually a CMS cap/floor is a product of a long-term CMS rate (>10Y) with frequency 6M or 3M. Therefore, already a small difference in the CMS option price and CMS convexity adjustment is significant. Especially, since the notional values for these kinds of derivatives are usually quite large. The computed CMS convexity adjustment is the smallest for the simple version of the linear TSR model. The fact that the simple version of the linear TSR model performs satisfactory is probably due to the fact that interest rates were low in 2013, the yield curve was relatively flat. If the yield curve becomes less flat the performance of the simple version of the linear TSR model is expected to decrease. Therefore, we will not consider the simple version of the linear TSR model for valuation of CMS options. For the mean reversion linear TSR model the computed CMS convexity adjustment increases as the value of the mean reversion parameter $\varkappa$ increases.

To see that there is also a timing effect, we investigate the effect of moving the start date further into the future on the value of the CMS convexity adjustment of a CMS swaplet. We will do this by means of a simple example, but first we present the following useful lemma:

Lemma 3.5.4 (CMS Price and CMS Convexity Adjustment under Linear TSR Model). *Using the linear TSR model for a CMS swaplet we can write the CMS price and CMS convexity adjustment in the following form,*

$$V_{gCMS}(0) = S(0) + \frac{A(0)}{P(0,T_p)} a\, Var^A\left(S(T_n)\right), \tag{3.42}$$

$$\Lambda_{gCMS}(0) = \frac{A(0)}{P(0,T_p)} a\, Var^A\left(S(T_n)\right). \tag{3.43}$$

Proof. The proof is given in Appendix A.8. □

Example 3.5.5 (Timing Effect). *We consider the problem of pricing a CMS swaplet on a 10Y CMS rate with 6M frequency. Today's date is 20-nov-13. We will consider different start dates, namely: 20-nov-14, 20-may-15, ..., 20-nov-23. The payment dates are equal to the start dates. Since we are pricing a swaplet, we have $K = 0$. Furthermore, we assume that interest rates are flat at 5%. We can then obtain our bond prices by making use of Lemma 2.4.1,*

$$P(0,T_n) = e^{-0.05T_n}, \tag{3.44}$$

where T_n denotes the tenor of the swaplet. For simplification reasons we will keep the value for the swap rate constant. The swap rate is assumed to have a log-normal distribution with a constant volatility of $\sigma = 17\%$ for all fixing dates. Since the swap rate is assumed to have a log-normal distribution with a constant volatility parameter we have

$$\begin{aligned} Var^A S(T_n) &= \mathbb{E}^A \left[S(T_n)^2 \middle| \mathcal{F}_n \right] - \left(\mathbb{E}^A \left[S(T_n) \middle| \mathcal{F}_0 \right] \right)^2 \\ &= S(0)^2 e^{\sigma^2 T_n} - S(0)^2 \\ &= S(0)^2 (e^{\sigma^2 T_n} - 1). \end{aligned} \tag{3.45}$$

From (3.45) and Lemma 3.5.4 it follows that

$$V_{gCMS}(0) = S(0) + \frac{A(0)}{P(0,T_p)} a S(0)^2 \left(e^{\sigma^2 T_n} - 1 \right), \tag{3.46}$$

$$\Lambda_{gCMS}(0) = \frac{A(0)}{P(0,T_p)} a S(0)^2 \left(e^{\sigma^2 T_n} - 1 \right). \tag{3.47}$$

Figure 3.3 shows the CMS convexity adjustment for the CMS swaplet on a 10Y CMS rate with 6M frequency at different times of maturity for different versions of the linear TSR model.

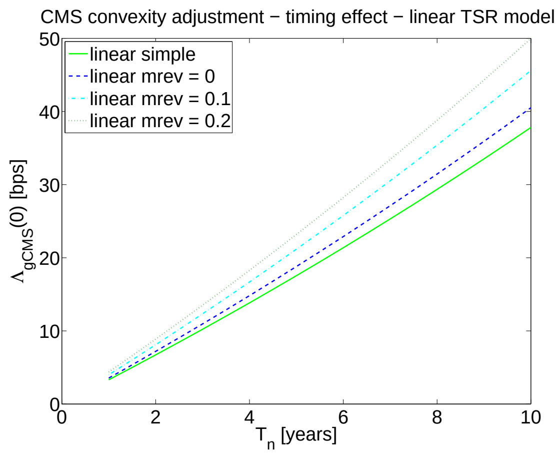


Figure 3.3: Simple example for the timing effect. The convexity adjustment of a CMS swaplet on 10Y CMS rate with 6M frequency is computed using the linear TSR model and simplifying assumptions.

From Figure 3.3 in Example 3.5.5 we can clearly see that as the start date moves further into the future the convexity adjustment for the CMS swaplet increases, indicating that there may be a timing effect.

3.5.2 Swap-Yield TSR Model

Arguably the most popular TSR model in the financial industry is the swap-yield TSR model. Its popularity stems from the fact that only a single assumption is necessary to derive the annuity mapping function. The assumption that is made, as we we will show later, is that all underlying swap rates are approximated by a single swap rate. The derivation of the Swap Market Model (SMM), a sophisticated model (a term structure model), is actually based on all underlying swap rates, [13].

We will now proceed with the actual derivation of the annuity mapping function. Remember that the annuity mapping function is defined to be the function that maps the term

$P(T_0, T_p)/A(T_0)$ to a function of the swap rate. To derive an expression for (3.27) based solely on swap rates, we start by writing the annuity as follows,

$$\begin{aligned} A_{0,N}(T_0) &= \sum_{n=0}^{N-1} \tau_n P(T_0, T_{n+1}) \\ &= \sum_{n=0}^{N-2} \tau_n P(T_0, T_{n+1}) + \tau_{N-1} P(T_0, T_N) \\ &= A_{0,N-1}(T_0) + \tau_{N-1} P(T_0, T_N). \end{aligned} \tag{3.48}$$

The annuity can also be written in the following form:

$$\begin{aligned} A_{0,N}(T_0) &= \sum_{n=0}^{N-1} \tau_n P(T_0, T_{n+1}) \\ &= \frac{\sum_{n=0}^{N-1} \tau_n P(T_0, T_{n+1})}{1 - P(T_0, T_N)} (1 - P(T_0, T_N)) \\ &= \frac{\sum_{n=0}^{N-1} \tau_n P(T_0, T_{n+1})}{P(T_0, T) - P(T_0, T_N)} (1 - P(T_0, T_N)) \\ &= \frac{1}{S_{0,N}(T_0)} (1 - P(T_0, T_N)). \end{aligned} \tag{3.49}$$

From (3.49) we see that,

$$P(T_0, T_N) = 1 - S_{0,N}(T_0) A_{0,N}(T_0). \tag{3.50}$$

Now, substituting (3.50) in (3.48) we get:

$$A_{0,N}(T_0) = A_{0,N-1}(T_0) + \tau_{N-1}(1 - S_{0,N}(T_0) A_{0,N}(T_0)).$$

Rewriting gives us the following useful recursive expression:

$$A_{0,N}(T_0) = \frac{1}{1 + \tau_{N-1} S_{0,N}(T_0)} (\tau_{N-1} + A_{0,N-1}(T_0)). \tag{3.51}$$

Unwrapping the recursion in (3.51) we obtain:

$$A_{0,N}(T_0) = \sum_{n=0}^{N-1} \tau_n \prod_{i=0}^{n} \frac{1}{1 + \tau_i S_{0,i+1}(T_0)}. \tag{3.52}$$

From (3.52) we see that the ZCBs, $P(T, T_n)$ with $n \geq 1$ are given by:

$$P(T_0, T_n) = \prod_{i=0}^{n-1} \frac{1}{1 + \tau_i S_{0,i+1}(T_0)}. \tag{3.53}$$

Equation (3.53) is an expression depending solely on the swap rates. We now make another simplifying assumption. We approximate all underlying swap rates $S_{0,i+1}(T_0)$ in (3.53) by a single swap rate $S_{0,N}(T_0)$. It follows that the mapping functions for the swap-yield TSR model are defined by

$$\pi(s, M) = \prod_{i=0}^{n-1} \frac{1}{1 + \tau_i s}. \tag{3.54}$$

Formula (3.54) essentially tells us to discount all cash flows after T_0 at the same rate, namely a rate given by the realized swap rate $S_{0,N}(T_0)$. Another useful observation to make is that we can write the annuity as follows,

$$A_{0,N}(T_0) = \frac{1}{S_{0,N}(T_0)} \left(1 - \prod_{i=0}^{N-1} \frac{1}{1 + \tau_i S_{0,N}(T_0)} \right). \tag{3.55}$$

Assuming the payment date is equal to the start date we have,

$$P(T_0, T_p) = P(T_0, T_0) = 1. \tag{3.56}$$

Thus, the annuity function is given by:

$$\alpha(s) = \frac{s}{1 - \prod_{i=0}^{N-1} \frac{1}{1+\tau_i s}}. \tag{3.57}$$

We highlight this result by listing it as a lemma, Lemma 3.5.6. Additionally, Lemma 3.5.6 gives expressions for the first and second derivatives of the annuity mapping function[4].

Lemma 3.5.6 (Annuity Mapping Function for Swap-Yield TSR Model). *The annuity function and its first and second derivatives for the swap-yield TSR model are given by:*

$$\alpha(s) = \frac{y}{z}, \tag{3.58}$$

$$\frac{d\alpha}{ds} = \frac{z\frac{dy}{ds} - y\frac{dz}{ds}}{z^2}, \tag{3.59}$$

$$\frac{d^2\alpha}{ds^2} = \frac{z\left(z\frac{d^2y}{ds^2} - y\frac{d^2z}{ds^2}\right) - 2\frac{dz}{ds}\left(z\frac{dy}{ds} - y\frac{dz}{ds}\right)}{z^3}, \tag{3.60}$$

where

$$y = s, \quad \frac{dy}{ds} = 1, \quad \frac{d^2y}{ds^2} = 0, \tag{3.61}$$

$$z = 1 - \prod_{i=0}^{N-1} \frac{1}{1+\tau_i s}, \tag{3.62}$$

$$\frac{dz}{ds} = \prod_{i=0}^{N-1} \frac{1}{1+\tau_i s} \sum_{i=0}^{N-1} \frac{\tau_i}{1+\tau_i s}, \tag{3.63}$$

$$\frac{d^2z}{ds^2} = -\left(\prod_{i=0}^{N-1} \frac{1}{1+\tau_i s} \left(\sum_{i=0}^{N-1} \frac{-\tau_i}{1+\tau_i s} \right)^2 + \prod_{i=0}^{N-1} \frac{1}{1+\tau_i s} \sum_{i=0}^{N-1} \left(\frac{\tau_i}{1+\tau_i s} \right)^2 \right). \tag{3.64}$$

Proof. The proof is given in Appendix A.9. □

To be considered as a proper TSR model the swap-yield model must satisfy the *no-arbitrage*, *consistency* and *realism* conditions. The realism condition is satisfied, as follows directly from (3.54). The consistency condition (3.24) is satisfied automatically as the following identity holds,

$$\frac{1 - \pi(s, T_N)}{\sum_{n=0}^{N-1} \tau_n \pi(s, T_{n+1})} = s.$$

However, the current form of the swap-yield TSR model is not arbitrage-free, as (3.23) is not satisfied.

[4] The first and second derivatives are needed in the replication method, Lemma 3.3.1. They are not trivial from the annuity mapping function as is the case for the linear annuity mapping function.

Arbitrage-Free Swap-Yield TSR Model

Clearly, the swap-yield TSR model violates the no-arbitrage condition, since

$$\mathbb{E}^A\left[\alpha(S(T_0))\right|\mathcal{F}_0] \neq \frac{P(0,T_p)}{A(0)}. \tag{3.65}$$

However, this problem can be fixed by re-scaling the original annuity mapping function. We obtain the new annuity mapping function $\tilde{\alpha}(s)$ in the following way:

$$\tilde{\alpha}(s) \triangleq \frac{P(0,T_p)}{A(0)}\frac{\alpha(s)}{\overline{\alpha}}, \tag{3.66}$$

where

$$\overline{\alpha} = \mathbb{E}^A\left[\alpha(S(T_0))\right|\mathcal{F}_0]. \tag{3.67}$$

Now, we can check that indeed the no-arbitrage condition is satisfied, as we have

$$\begin{aligned}
\mathbb{E}^A\left[\tilde{\alpha}(S(T_0))\right|\mathcal{F}_0] &= \mathbb{E}^A\left[\frac{P(0,T_p)}{A(0)}\frac{\alpha(S(T_0))}{\overline{\alpha}}\middle|\mathcal{F}_0\right] \\
&= \frac{P(0,T_p)}{A(0)}\mathbb{E}^A\left[\frac{\alpha(S(T_0))}{\mathbb{E}^A\left[\alpha(S(T_0))\right|\mathcal{F}_0]}\middle|\mathcal{F}_0\right] \\
&= \frac{P(0,T_p)}{A(0)}.
\end{aligned}$$

We also obtain a new valuation formula which can be written in the following convenient form,

$$\begin{aligned}
V_{gCMS}(0) &= \frac{A(0)}{P(0,T_p)}\mathbb{E}^A\left[\tilde{\alpha}(S(T_0))g(S(T_0))\right|\mathcal{F}_0] \\
&= \frac{A(0)}{P(0,T_p)}\mathbb{E}^A\left[\frac{P(0,T_p)}{A(0)}\frac{\alpha(S(T_0))}{\overline{\alpha}}g(S(T_0))\middle|\mathcal{F}_0\right] \\
&= \mathbb{E}^A\left[\frac{\alpha(S(T_0))g(S(T_0))}{\mathbb{E}^A\left[\alpha(S(T_0))\right|\mathcal{F}_0]}\middle|\mathcal{F}_0\right] \\
&= \frac{\mathbb{E}^A\left[f(S(T_0))\right|\mathcal{F}_0]}{\mathbb{E}^A\left[\alpha(S(T_0))\right|\mathcal{F}_0]}.
\end{aligned} \tag{3.68}$$

Remark 3.5.7. *Two important observations are:*

- *The correction (3.66) is useful even for arbitrage-free models, where the no-arbitrage condition (3.65) holds in theory. This follows from the fact that in practice the no-arbitrage condition can be violated by the used numerical scheme. Therefore, the valuation formula (3.68) can also be useful for other types of TSR models.*
- *The use of valuation formula (3.68) doubles the computation time.*

We again compute the CMS convexity adjustment of the CMS caplet on 10Y CMS rate with 12M frequency, and the result is given in Figure 3.4.

If we compare Figure 3.4 to Figure 3.2 we see that the result of the swap-yield TSR model is closest to the result of the mean reversion linear TSR model with $\varkappa = 0$.

As a final note, the main downside of the swap-yield TSR model is its lack of explicit control over the shape of the yield curve at time T. In the linear TSR model we have explicit control over the yield curve at time T, which was done by imposing a link between parameters of these models to a mean reversion parameter.

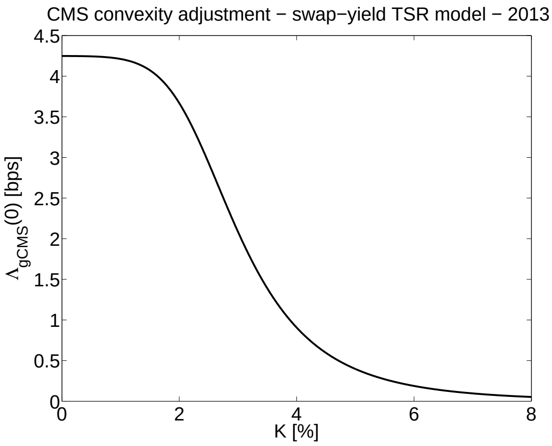


Figure 3.4: CMS convexity adjustment of a CMS caplet on 10Y CMS rate with 12M frequency using the swap-yield TSR model.

3.5.3 Interpolation TSR models

The linear and especially the swap-yield TSR model are relatively well-established in the literature. As we specified earlier, the annuity mapping functions of these TSR models can be seen as approximations of the true annuity mapping function defined by the conditional expected value in (3.27). We will propose two new TSR models that are based on interpolation. The value of ZCB $P(T_0, T_0)$ is known. Furthermore, it can also be assumed that we know the value of the swap rate $S_{0,N}(T_0)$. Therefore, using the definition of the swap rate we can obtain an expression for $P(T_0, T_N)$.

Linear Interpolation TSR model

The *first interpolation TSR model* we will develop is based on linear interpolation of the ZCBs, we will therefore call it the *linear interpolation TSR model*. We make use of the following type of interpolation:

$$P(T_0, T_n) = \theta_n P(T_0, T_0) + (1 - \theta_n) P(T_0, T_N), \tag{3.69}$$

for $T_0 \leq T_n \leq T_N$. Since $P(T_0, T_0) = 1$, we obtain

$$P(T_0, T_n) = \theta_n + (1 - \theta_n) P(T_0, T_N). \tag{3.70}$$

Substituting (3.70) in the definition of the annuity we obtain:

$$\begin{aligned} A_{0,N}(T_0) &= \sum_{n=0}^{N-1} \tau_n P(T_0, T_{n+1}) \\ &= \sum_{n=0}^{N-1} \tau_n \left(\theta_{n+1} + (1 - \theta_{n+1}) P(T_0, T_N)\right) \\ &= \sum_{n=0}^{N-1} \tau_n \theta_{n+1} + P(T_0, T_N) \left(\sum_{n=0}^{N-1} \tau_n - \sum_{n=0}^{N-1} \tau_n \theta_{n+1} \right). \end{aligned} \tag{3.71}$$

For notational convenience we write,

$$\Gamma_1 = \sum_{n=0}^{N-1} \tau_n, \tag{3.72}$$

$$\Gamma_2 = \sum_{n=0}^{N-1} \tau_n \theta_{n+1}. \tag{3.73}$$

So we have:

$$A_{0,N}(T_0) = \Gamma_2 + P(T_0, T_N)\left(\Gamma_1 - \Gamma_2\right). \tag{3.74}$$

From (3.74) it follows that for the swap rate we obtain,

$$\begin{aligned} S_{0,N}(T_0) &= \frac{P(T_0,T_0) - P(T_0,T_N)}{A_{0,N}(T_0)} \\ &= \frac{1 - P(T_0,T_N)}{\Gamma_2 + P(T_0,T_N)\left(\Gamma_1 - \Gamma_2\right)}. \end{aligned} \tag{3.75}$$

From (3.75) we find the following expression for $P(T_0, T_N)$:

$$P(T_0, T_N) = \frac{1 - \Gamma_2 S_{0,N}(T_0)}{1 + (\Gamma_1 - \Gamma_2) S_{0,N}(T_0)}. \tag{3.76}$$

Substituting (3.76) in (3.74) gives us the following expression for the annuity:

$$\begin{aligned} A_{0,N}(T_0) &= \Gamma_2 + (\Gamma_1 - \Gamma_2)\frac{1 - \Gamma_2 S_{0,N}(T_0)}{1 + (\Gamma_1 - \Gamma_2) S_{0,N}(T_0)} \\ &= \frac{\Gamma_2\left(1 + (\Gamma_1 - \Gamma_2) S_{0,N}(T_0)\right)}{1 + (\Gamma_1 - \Gamma_2) S_{0,N}(T_0)} + \frac{(\Gamma_1 - \Gamma_2)\left(1 - \Gamma_2 S_{0,N}(T_0)\right)}{1 + (\Gamma_1 - \Gamma_2) S_{0,N}(T_0)} \\ &= \frac{\Gamma_2 + \Gamma_1\Gamma_2 S_{0,N}(T_0) - \Gamma_2^2 S_{0,N}(T_0) + \Gamma_1 - \Gamma_1\Gamma_2 S_{0,N}(T_0) - \Gamma_2 + -\Gamma_2^2 S_{0,N}(T_0)}{1 + (\Gamma_1 - \Gamma_2) S_{0,N}(T_0)} \\ &= \frac{\Gamma_1}{1 + (\Gamma_1 - \Gamma_2) S_{0,N}(T_0)}. \end{aligned} \tag{3.77}$$

Thus, we obtain the following annuity mapping function,

$$\alpha(s) = \frac{\Gamma_1 - \Gamma_2}{\Gamma_1} s + \frac{1}{\Gamma_1}, \tag{3.78}$$

where Γ_1, Γ_2 are given by (3.72) and (3.73), respectively. The consistency and reasonability conditions are satisfied automatically, as was the case for the swap-yield TSR model. Coefficient Γ_2 can be chosen such that the linear interpolation TSR model is arbitrage-free. Remember that for a TSR model to be arbitrage-free,

$$\mathbb{E}^A\left[\alpha(S(T_0))\middle|\,\mathcal{F}_0\right] = \frac{P(0,T_p)}{A(0)}, \tag{3.79}$$

must hold. We have,

$$\mathbb{E}^A\left[\alpha(S(T_0))\middle|\,\mathcal{F}_0\right] = \frac{1 + (\Gamma_1 - \Gamma_2)\mathbb{E}^A\left[S(T_0)\right]}{\Gamma_1} = \frac{1 + (\Gamma_1 - \Gamma_2) S(0)}{\Gamma_1}. \tag{3.80}$$

If we set

$$\frac{1 + (\Gamma_1 - \Gamma_2) S(0)}{\Gamma_1} = \frac{P(0,T_p)}{A(0)}, \tag{3.81}$$

it follows that coefficient Γ_2 must be chosen such that

$$\Gamma_2 = \Gamma_1 - \frac{1}{S(0)}\left(\frac{P(0,T_p)\Gamma_1}{A(0)} - 1\right). \tag{3.82}$$

The obtained result is summarized by Lemma 3.5.8.

Lemma 3.5.8 (**Annuity Mapping Function for Linear Interpolation TSR Model**). *The annuity function for the linear interpolation TSR model is given by:*

$$\alpha(s) = \frac{\Gamma_1 - \Gamma_2}{\Gamma_1} s + \frac{1}{\Gamma_1}, \tag{3.83}$$

where Γ_1, Γ_2 *are given by*

$$\Gamma_1 = \sum_{n=0}^{N-1} \tau_n, \tag{3.84}$$

$$\Gamma_2 = \Gamma_1 - \frac{1}{S(0)}\left(\frac{P(0,T_p)\Gamma_1}{A(0)} - 1\right). \tag{3.85}$$

We will again compute the CMS convexity adjustment of the CMS caplet on 10Y CMS rate with 12M frequency. Since the market standard is the swap-yield TSR model, we will compare the results of both models. The results are given in Figure 3.5.

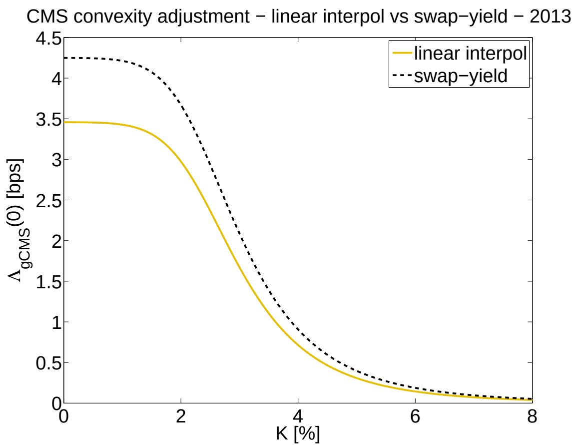


Figure 3.5: CMS convexity adjustment of a CMS caplet on 10Y CMS rate with 12M frequency using the linear interpolation TSR model.

Figure 3.5 shows that the results of the linear interpolation TSR model are close to the results of the swap-yield TSR model. In fact, the difference between the calculated convexity adjustment of the two models is smaller than 1bp.

Log-Linear Interpolation TSR model

The second TSR model we will develop is based on linear interpolation of the logarithm of ZCBs, which can be a better way to describe the future yield curve movement. Therefore, we will call it the *log-linear interpolation TSR model.* We make use of the following type of interpolation:

$$\log(P(T_0,T_n)) = \frac{T_N - T_n}{T_N - T_0}\log(P(T_0,T_0)) + \frac{T_n - T_0}{T_N - T_0}\log(P(T_0,T_N)), \tag{3.86}$$

for $T_0 \leq T_n \leq T_N$. Using the fact that $P(T_0, T_0) = 1$, we obtain

$$\log(P(T_0, T_n)) = \frac{T_n - T_0}{T_N - T_0} \log(P(T_0, T_N)). \tag{3.87}$$

Rewriting (3.87) we get the following expression for $P(T_0, T_n)$,

$$P(T_0, T_n) = P(T_0, T_N)^{\frac{T_n - T_0}{T_N - T_0}}. \tag{3.88}$$

Proceeding in the same fashion as we did with the linear interpolation TSR model, making use of the definition of the swap rate, we find

$$S_{0,N}(T_0) = \frac{1 - P(T_0, T_N)}{\sum_{n=0}^{N-1} \tau_n P(T_0, T_N)^{\frac{T_{n+1} - T_0}{T_N - T_0}}}. \tag{3.89}$$

Unfortunately, we can not obtain an expression for $P(T_0, T_N)$ analytically, but we can solve it by a few iterations of a numerical root finding algorithm. We can rewrite (3.89) as follows,

$$S_{0,N}(T_0) \sum_{n=0}^{N-1} \tau_n P(T_0, T_N)^{\frac{T_{n+1} - T_0}{T_N - T_0}} + P(T_0, T_N) - 1 = 0. \tag{3.90}$$

Next, we set $P(T_0, T_N)$ to be the solution, $z(S(T_0))$, of the equation

$$S_{0,N}(T_0) \sum_{n=0}^{N-1} \tau_n z(S(T_0))^{\frac{T_{n+1} - T_0}{T_N - T_0}} + z(S(T_0)) - 1 = 0. \tag{3.91}$$

We get the following expression for the annuity,

$$A_{0,N}(T_0) = \sum_{n=0}^{N-1} \tau_n (z(S(T_0)))^{\frac{T_{n+1} - T_0}{T_N - T_0}}. \tag{3.92}$$

Thus we obtain the following annuity mapping function,

$$\alpha(s) = \frac{1}{\sum_{n=0}^{N-1} \tau_n (z(s))^{\frac{T_{n+1} - T_0}{T_N - T_0}}}. \tag{3.93}$$

As was the case for the linear interpolation TSR model, the consistency and reasonability conditions are again satisfied automatically. However, this TSR model is not arbitrage-free. To fix this problem we can make use of valuation formula (3.68). Besides the use of valuation formula (3.68), also a numerical root finding algorithm has to be used to calculate the values for the annuity mapping function. Meaning that the log-linear TSR model has the highest computational cost of the considered TSR models, although the computational costs are still very low when compared to sophisticated models.

The annuity function and its derivatives are given by Lemma 3.5.9.

Lemma 3.5.9 (**Annuity Mapping Function for Log-Linear Interpolation TSR Model**). *The annuity function and its first and second derivatives for the log-linear interpolation TSR model are given by:*

$$\alpha(s) = \frac{1}{\sum_{n=0}^{N-1} \tau_n z(s)^{\vartheta_{n+1}}}, \tag{3.94}$$

$$\frac{d\alpha}{ds} = \frac{\Upsilon_1(s)}{\Upsilon_2(s)}, \tag{3.95}$$

$$\frac{d^2}{ds^2}\alpha(s) = \frac{\Upsilon_2(s)\frac{d\Upsilon_1}{ds} - \Upsilon_1(s)\frac{d\Upsilon_2}{ds}}{(\Upsilon_2(s))^2}, \tag{3.96}$$

where $\vartheta_n = \frac{T_n - T_0}{T_N - T_0}$ *and*

$$\Upsilon_1(s) = -\frac{dz}{ds} \sum_{n=0}^{N-1} \tau_n \vartheta_{n+1} z(s)^{\vartheta_{n+1}-1}, \tag{3.97}$$

$$\Upsilon_2(s) = \left(\sum_{n=0}^{N-1} \tau_n z(s)^{\vartheta_{n+1}} \right)^2, \tag{3.98}$$

$$\frac{d\Upsilon_1}{ds} = -\frac{d^2 z}{ds^2} \sum_{n=0}^{N-1} \tau_n \vartheta_{n+1} z(s)^{\vartheta_{n+1}-1} - \left(\frac{dz}{ds}\right)^2 \sum_{n=0}^{N-1} \tau_n \vartheta_{n+1}(\vartheta_{n+1}-1) z(s)^{\vartheta_{n+1}-2}, \tag{3.99}$$

$$\frac{d\Upsilon_2}{ds} = 2 \left(\sum_{n=0}^{N-1} \tau_n z(s)^{\vartheta_{n+1}} \right) \left(\frac{dz}{ds} \sum_{n=0}^{N-1} \tau_n \vartheta_{n+1} z(s)^{\vartheta_{n+1}-1} \right). \tag{3.100}$$

The first and second derivatives of $z(s)$ *with respect to* s *are given by,*

$$\frac{dz}{ds} = \frac{-\sum_{n=0}^{N-1} \tau_n z(s)^{\vartheta_{n+1}}}{1 + s \sum_{n=0}^{N-1} \tau_n \vartheta_{n+1} z(s)^{\vartheta_{n+1}-1}}, \tag{3.101}$$

$$\begin{aligned} \frac{d^2 z}{ds^2} = & \frac{-\frac{dz}{ds} \sum_{n=0}^{N-1} \tau_n \vartheta_{n+1} z(s)^{\vartheta_{n+1}-1}}{1 + s \sum_{n=0}^{N-1} \tau_n \vartheta_{n+1} z(s)^{\vartheta_{n+1}-1}} \\ & + \frac{\left(\sum_{n=0}^{N-1} \tau_n z(s)^{\vartheta_{n+1}}\right)\left(\sum_{n=0}^{N-1} \tau_n \vartheta_{n+1} z(s)^{\vartheta_{n+1}-1} + s\frac{dz}{ds}\sum_{n=0}^{N-1} \tau_n \vartheta_{n+1}(\vartheta_{n+1}-1) z(s)^{\vartheta_{n+1}-2}\right)}{\left(1 + s \sum_{n=0}^{N-1} \tau_n \vartheta_{n+1} z(s)^{\vartheta_{n+1}-1}\right)^2}. \end{aligned} \tag{3.102}$$

Proof. The proof is given in Appendix A.10 □

For this model we will also compute the CMS convexity adjustment of the CMS caplet on 10Y CMS rate with 12M frequency. We will again compare the results to the results obtained with the swap-yield TSR model. The results are given in Figure 3.6.

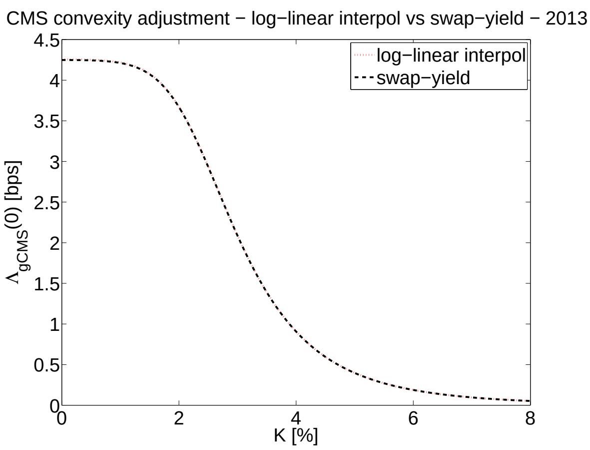


Figure 3.6: CMS convexity adjustment of a CMS caplet on 10Y CMS rate with 12M frequency using the log-linear interpolation TSR model.

Figure 3.6 shows that the results of the log-linear interpolation TSR model and the swap-yield TSR model are almost identical.

To evaluate the performance of the respective TSR models additional numerical experiments are necessary, which we will perform in the next section.

3.6 Numerical Experiments

In this section we will perform more numerical experiments to get better insight in the performance of the different TSR models. Besides the market data from 2013 we will now also make use of market data from 2007. The market data from 2007 is given by Table B.8 and Table B.9; which can be found in Appendix B. We will compare the difference in prices computed with the different TSR models and a reference model. The swap-yield TSR is chosen as the benchmark/reference model, since currently it is the most popular TSR model that is used in the market.

Remark 3.6.1. *The swap-yield TSR model is not necessarily better than the other TSR models that will be considered. It is simply used as a benchmark, since there are no other references available.*

The price that is calculated with the reference model will be denoted by V_{ref}. The difference between the price computed with a chosen TSR model and the reference model will be denoted by ζ, where ζ is defined as:

$$\zeta = V_{\text{gCMS}}(0) - V_{\text{ref}}.$$

Besides looking at the difference in prices computed by the different TSR models we also study the volatility and timing effects.

3.6.1 CMS Caplet Price: 2007 vs 2013

To gain insight into the performance of the different TSR models we will price a CMS caplet[5] on a 10Y CMS rate with 12M frequency for market data from 2007 and 2013. The computed CMS caplet prices and ζ for 2007 and 2013 are given in Figure 3.7.

If we compare the results of 2007 with 2013, we can make a number of observations. First, we note that the performance of the respective TSR models is similar for both 2007 and 2013. The computed CMS caplet prices with the respective TSR models only differ slightly from the price computed with the reference model. For both years the results of the log-linear TSR model are closest to the results of the reference model, although the results of the mean reversion linear TSR model with $\varkappa = 0$ are also quite close. We also observe that the differences for 2013 are larger than the differences for 2007, which is probably due to the fact that the volatilities observed in 2013 are more extreme. The fact that the differences for 2013 are bigger than for 2007 is an indication that it is now of even more importance to choose a correct TSR model for the pricing of CMS derivatives.

[5]The reason that we price a CMS caplet/floorlet instead of a CMS cap/floor is that to price a CMS cap/floor a lot more market data has to be used. We can already draw our conclusions about the model performance from the pricing of a single CMS caplet/floorlet, since a CMS cap/floor is just a sum of CMS caplets/floorlets.

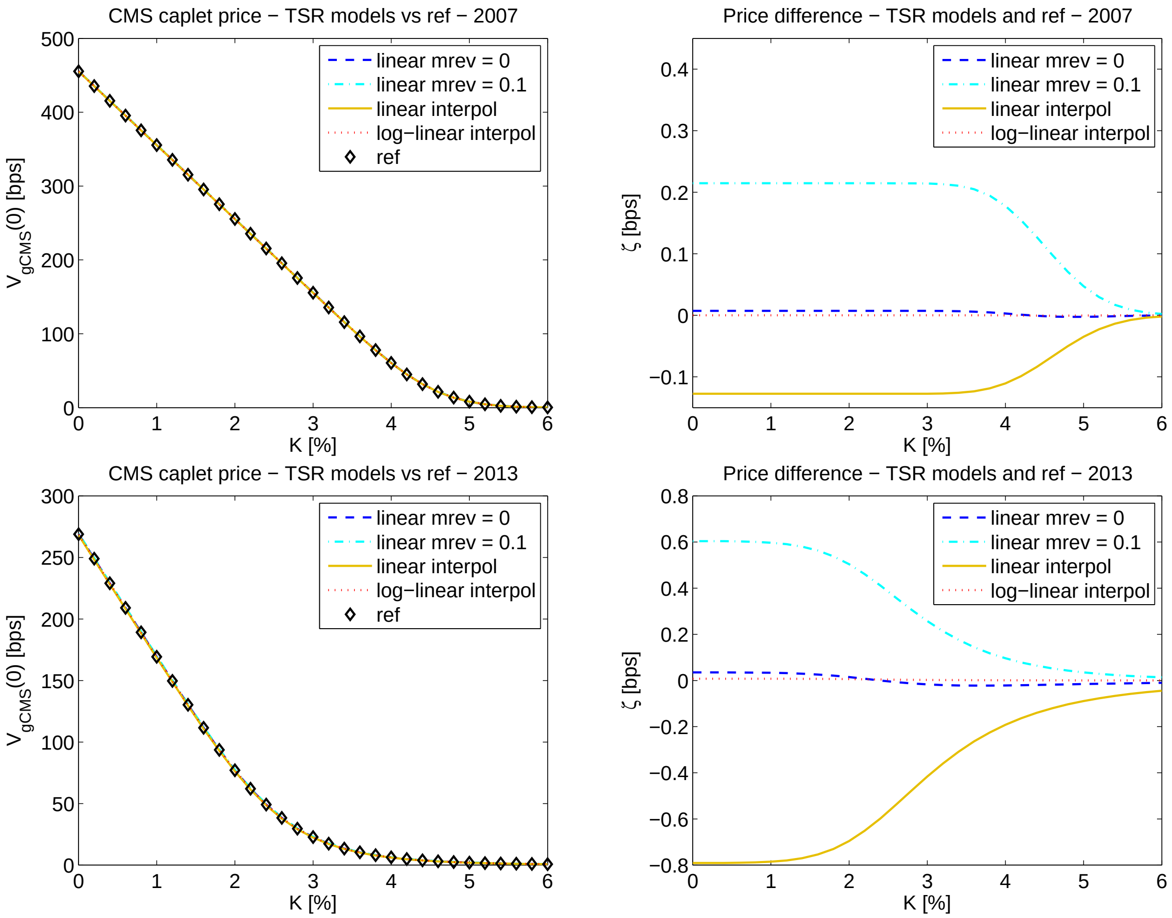


Figure 3.7: Prices of a CMS caplet on 10Y CMS rate with 12M frequency for 2007 and 2013 - mean reversion linear TSR model, linear interpolation TSR model and log-linear interpolation TSR model vs reference model.

3.6.2 Investigate the Timing Effect

Next, we wish to investigate the timing effect. We already observed the timing effect to some extent in Example 3.5.5, but this time we will make use of the market data of 2007 and 2013. We will again make use of different start dates to study the timing effect. Originally the start date of the CMS caplet we considered was 1 year from today, Figure 3.7. We will now consider start dates up to 10 years from today. In Figure 3.8 ζ is given for 2007 and 2013 for two different starting dates.

From Figure 3.7 and Figure 3.8 it is clear that moving the start date further into the future leads to an increase in ζ for both market data from 2007 and 2013. From this we can infer that there is indeed a timing effect. We also note as before that the differences in prices for 2013 are larger than the differences in prices for 2007.

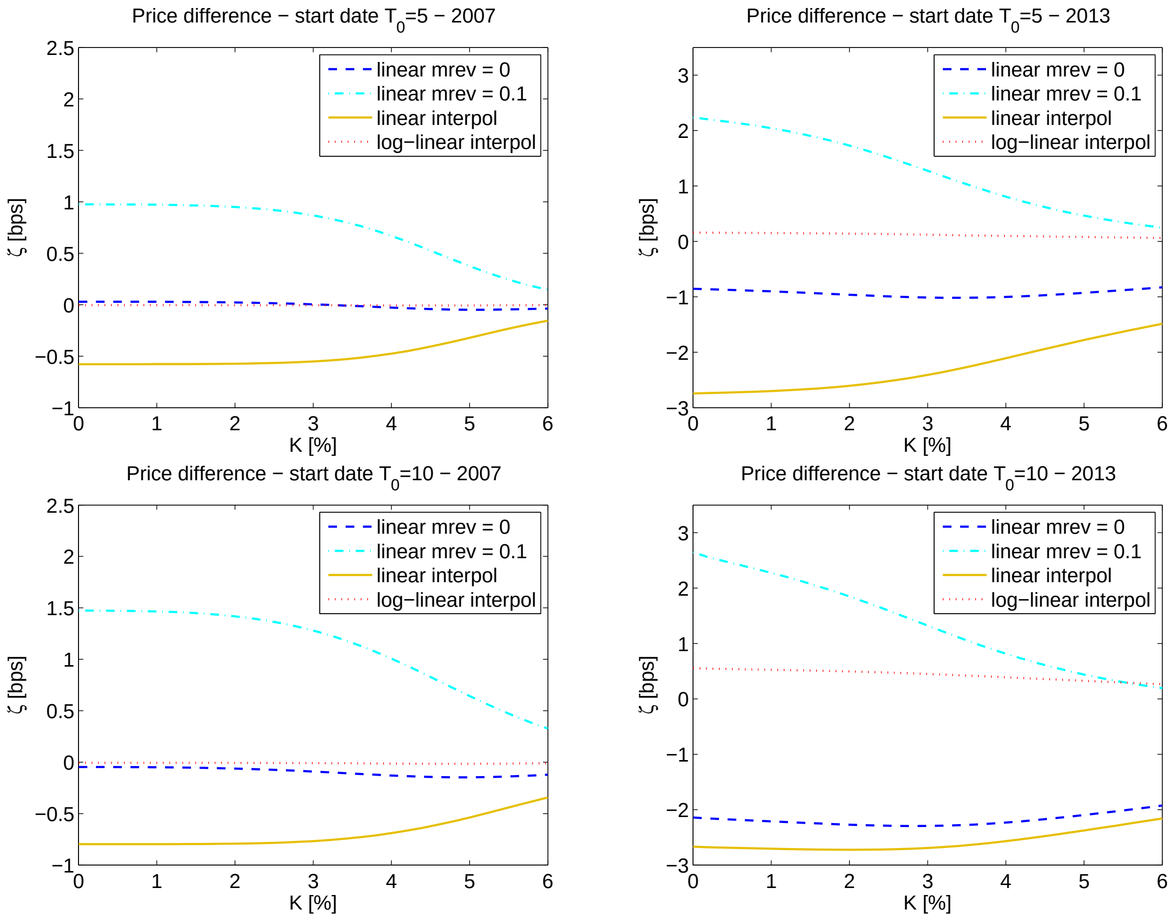


Figure 3.8: Price differences for a CMS caplet on 10Y CMS rate with 12M frequency for 2007 and 2013 - mean reversion linear TSR model, linear interpolation TSR model and log-linear interpolation TSR model vs reference model. Using different start dates, $T_0 = 5$ and $T_0 = 10$.

3.6.3 Investigate the Volatility Effect

To investigate the volatility effect we will again price a CMS caplet on a 10Y CMS rate with 12M frequency for low and high volatilities. We will partly use the market data from 2013, only now we will assume that the volatility is constant. We consider the case where the start date is 10 years from today, $T_0 = 10$. The calculated CMS caplet prices for low and high constant volatility are given in Figure 3.9.

From Figure 3.9 it is obvious that for high volatility the computed CMS caplet prices with the different TSR models differ more than for low volatility, indicating that there is a volatility effect.

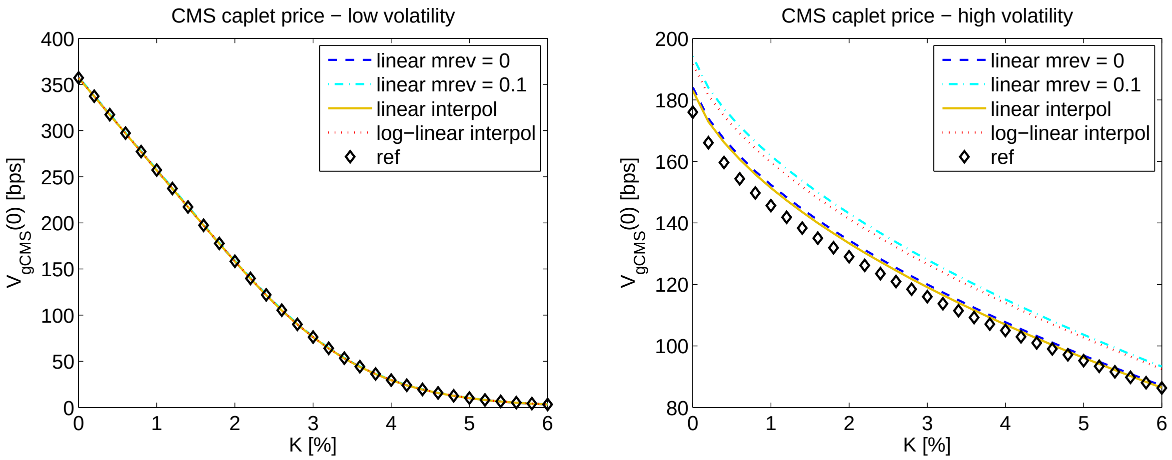


Figure 3.9: Prices of a CMS caplet on 10Y CMS rate with 12M frequency using different TSR models for low and high volatilities. The volatility is assumed to be constant, $\sigma_{\text{low}} = 0.1$ and $\sigma_{\text{high}} = 0.9$.

3.6.4 No-Arbitrage Condition

Two of the TSR models we considered are actually not arbitrage-free and we had to make use of rescaling by using valuation formula (3.68) instead of the theoretical valuation formula (3.17). In order to show the necessity of the rescaling, we will check if the no-arbitrage condition (3.79) is satisfied, we compute the difference,

$$\mathbb{E}^A \left[\alpha(S(T_0)) | \mathcal{F}_0 \right] - \frac{P(0, T_p)}{A(0)}, \tag{3.103}$$

for the swap-yield, linear interpolation and log-linear interpolation TSR models. The results are given in Figure 3.10.

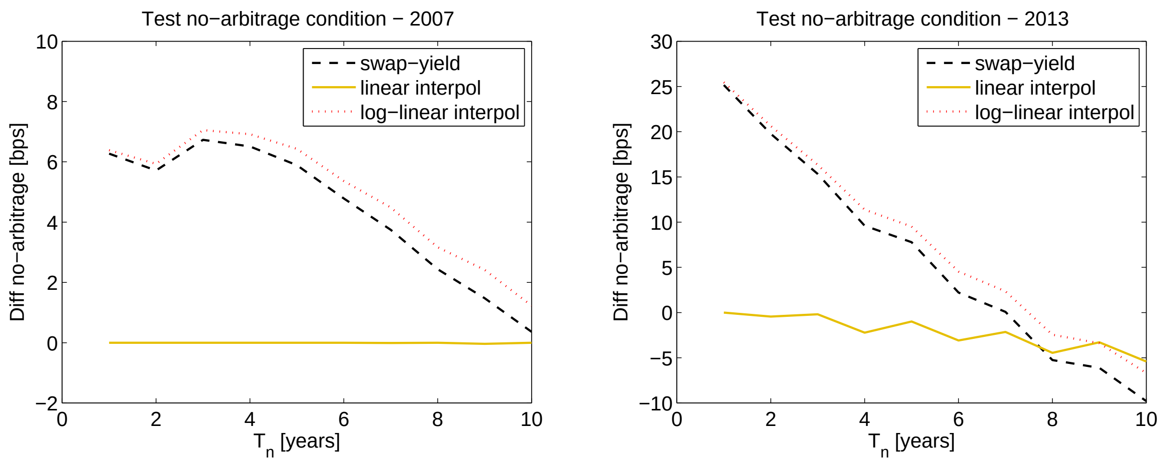


Figure 3.10: Test for the no-arbitrage condition. The difference given by (3.103), is computed with the swap-yield TSR model, the linear interpolation TSR model and the log-linear interpolation TSR model.

The mean reversion linear TSR model is arbitrage-free by definition, but additional work needs to be done to obtain a correct value for the mean reversion parameter $\varkappa$. As of yet we do not have a method to properly calibrate this mean reversion parameter. The linear interpolation

TSR model on the other hand requires no additional calibration and the model is arbitrage-free by construction. From Figure 3.10 we can conclude that indeed the linear interpolation TSR model is arbitrage-free by construction, while the swap-yield TSR model and the log-linear interpolation TSR model are certainly not arbitrage-free.

Since, not every TSR model is arbitrage-free a fairer way to compare the TSR models is by not making use of the rescaling in the annuity mapping function. We will again compute the CMS caplet on a 10Y CMS rate with 12M frequency, but this time we will use valuation formula (3.17) for all TSR models. The reference model is still the same as before, the swap-yield TSR model where we make use of rescaling. The results are given in Figure 3.11.

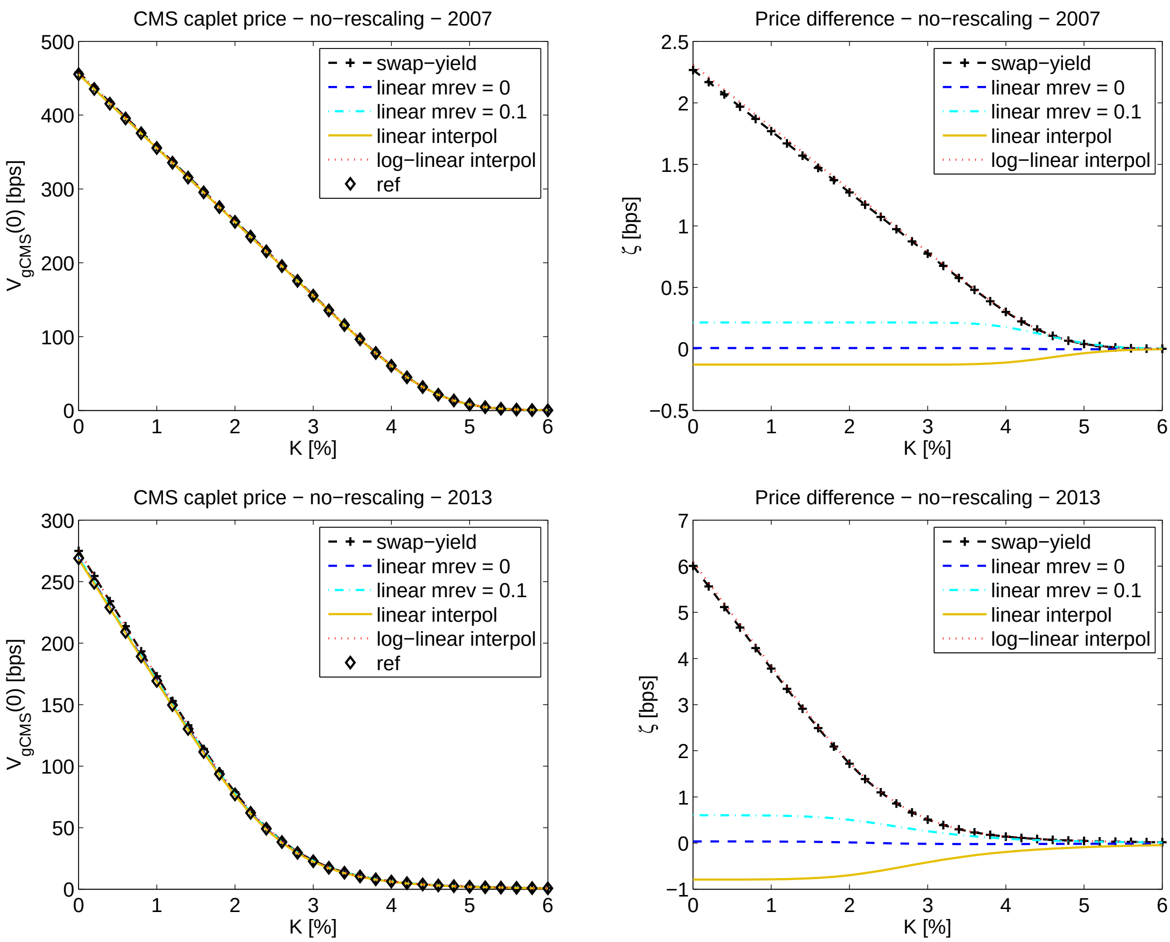


Figure 3.11: Prices of a CMS caplet on 10Y CMS rate with 12M frequency for 2007 and 2013 - no-rescaling - swap-yield TSR model, mean reversion linear TSR model, linear interpolation TSR model and log-linear interpolation TSR model vs reference model.

From Figure 3.11 it is clear that for the swap-yield TSR model and the log-linear TSR model the rescaling is absolutely necessary to obtain the correct price. For the remaining TSR models there is no notable difference when we use either (3.68) or (3.17). So, from this point of view the mean reversion linear and the linear interpolation TSR model are superior to the swap-yield and log-linear TSR models.

3.7 Conclusions

CMS-based products are widely used by insurance companies and pension funds in their Asset & Liability Management, because these institutions are very vulnerable to movements in the interest rates. CMS caps and floors are collections of options on CMS rates. The pricing of these products has to be efficient and accurate. However, the use of sophisticated models is not always desirable due to too time-consuming calculations. Therefore, different approaches are used in practice.

One of these approaches is the use of a Terminal Swap Rate model. TSR models are obtained by using the TSR approach. The TSR approach can be used when the dependence on the additional discount bonds is sufficiently small, so that primarily the swap rate determines the payoff.

We have demonstrated that it is convenient to change to the annuity measure when pricing CMS derivatives. We used the replication method to price a single CMS-linked cash flow. To compute the implied volatilities for different strikes we made use of Hagan's formula.

We have considered two types of TSR models described in the literature, namely the linear TSR model and the swap-yield TSR model. We also developed two new TSR models both based on interpolation, the linear interpolation TSR model and the log-linear interpolation TSR model.

Many numerical experiments were performed to study the performance of the respective TSR models. We considered market data from 2007 and 2013. The results for both sets of market data were similar, but we did observe that the differences for the year 2013 were bigger than for the year 2007, which is probably due to the fact that the volatilities observed in 2013 are more extreme. Therefore, nowadays correct valuation of CMS derivatives is of even more importance.

We have seen that depending on the chosen TSR model the computed price of the CMS option can differ. We also showed that there is a timing and a volatility effect. The further the start date is moved into the future the bigger the differences will be between the computed prices of the CMS derivative with the respective TSR models, indicating that there is a timing effect. We also demonstrated the volatility effect, by showing that for higher volatilities the price differences between the respective TSR models are larger.

From the numerical experiments we have seen that all TSR models have their pros and cons. The swap-yield TSR model is most widely used in the financial industry. Its popularity stems from the fact that only a single assumption is necessary to derive the annuity mapping function. The assumption that is made, is that all underlying swap rates are approximated by a single swap rate. A downside of the swap-yield model is that it is not arbitrage-free. A rescaling has to be used to correctly calculate the price of the CMS option price, which doubles the computation time. The mean reversion linear TSR model is arbitrage-free by definition. Of the four considered TSR models the mean reversion linear TSR model is the only TSR model that incorporates a mean reversion parameter, making it the most flexible TSR model. However, calibrating this mean reversion parameter is not straight-forward and is an issue that should be further researched. The linear and log-linear interpolation TSR models on the other hand require no additional calibration. The linear interpolation TSR model is based on a linear interpolation of the zero-coupon bonds. Another advantage of this model compared to the swap-yield TSR model is that it is arbitrage-free by construction. The log-linear interpolation TSR model is based on a linear interpolation of the logarithm of zero-coupon bonds, which can be a better way to describe the future yield curve movement. For the log-linear interpolation TSR model the same rescaling as for the swap-yield model has to be used. Besides the necessary rescaling, also a numerical root finding algorithm has to be used to calculate the values for the annuity mapping function. Meaning that the log-linear TSR model has the highest computational cost of the considered TSR models, although the computational costs are still very low when compared to sophisticated models.

So we would recommend the use of the log-linear interpolation TSR model to price CMS options, depending on the view of the movement of the yield curve. When it is important to reduce the calculation time, we recommend the use of the linear interpolation TSR model. If more flexibility needs to be added we would recommend the use of the mean reversion linear TSR model.

Usually a CMS cap/floor is a product of a long-term CMS rate (>10Y) with frequency 6M or 3M, so already a small difference when pricing a CMS caplet/floorlet is significant. Especially, since the notional values for these kind of derivatives are usually quite large. So even these small differences can lead to substantial losses. Thus obtaining a fast, efficient and accurate model to price CMS derivatives is of vital importance.

Chapter 4

Copula Approach for Pricing CMS Spread Derivatives

4.1 Introduction

In this chapter we look into the pricing of CMS spread derivatives by making use of the copula approach. Specifically, we will focus on the pricing of CMS spread options. This chapter is organized as follows.

In Section 4.2 we introduce CMS spread derivatives. Section 4.3 shortly describes the pricing approach we are going to apply. In Section 4.4 copulas are discussed, in particular the Gaussian copula. In Section 4.5 an efficient pricing formula for CMS spread options is derived, additionally a Monte Carlo method is presented by which CMS spread options can be priced. In Section 4.6 several numerical experiments are performed. Finally, Section 4.7 concludes.

This chapter is mainly based on [37, pp. 765-815].

4.2 CMS Spread Derivatives

A CMS spread derivative is a financial instrument whose payoff is a function of the spread between two swap rates of different maturity. This can for example be the 10-year swap rate minus the 2-year swap rate. These type of derivatives are traded by parties that wish to take advantage of, or hedge against, future changes in the slopes of specific parts of the yield curve. Therefore, this type of derivative has become quite popular among insurance companies and pension funds. The most common CMS spread derivatives are CMS spread notes/bonds (steepener or flattener), CMS spread range accrual notes/bonds, and CMS spread caps and floors. There are also other CMS spread derivatives that are not commonly traded in the market, but they are embedded in other financial instruments. Examples of such CMS derivatives are CMS spread digital options and CMS spread swaptions, [45].

The valuation of these CMS spread derivatives is an important subject of research for both practitioners and academics. Our focus will be on the pricing of CMS spread caps and floors, which are also referred to as a CMS spread options (CMSSOs). Typically banks use CMSSOs to hedge the CMS spread swaps that they have entered into with customers. A CMS spread cap/floor consists of a series of options which are also known as caplets/floorlets. Each caplet/floorlet ensures the buyer protection for a single payment period. A cap protects against an increase in the spread, whereas a floor protects against an inversion or reduction in the swap rate spread. On each fixing date, if the underlying is above the strike (for a cap) or below the strike (for a floor) the buyer receives a payout. As mentioned earlier, a CMSSO is an efficient way to exercise a view on the shape of the yield curve. In environments where the yield curve is very flat, the forecasted spread is typically around zero. However from historical analysis, yield curves tend to be characterized by low short term rates relative to longer term rates. In recent times for example, the 10Y2Y spread in USD has been as high as 200bps. Therefore, in markets where yield curves are currently flat, an investor could purchase a CMS spread cap based on a 10Y-2Y spread for a relatively low price. Subsequently, if over the tenor of the option the curve normalizes, the investor will be ITM and generate a significant gain.

The payoff of a CMS spread option[1] is given by

$$V_{\text{CMSSO}}(T_p) = (\varpi_1 S_1(T) - \varpi_2 S_2(T) - K)^+, \tag{4.1}$$

where $T_p \geq T$ is the time of payment, $S_1(T)$ and $S_2(T)$ are two swap rates of different tenors fixing at time T, K is the strike and ϖ_1 and ϖ_2 are the *gearing factors* of the respective swap rates. The gearing factor defines which percentage of a swap rate will be used in the payout calculation. The default value is 1, which represents 100%. Alternatively, the gearing factor can for example be chosen to be 0.5 which represents 50% or 1.5 which represents 150%. We will assume the default values for the gearing factors, so $\varpi_1 = \varpi_2 = 1$. In this case the undiscounted value of a CMS spread option is given by[2]

$$V_{\text{CMSSO}}(0) = \mathbb{E}^{T_p}\left[(S_1(T) - S_2(T) - K)^+ \middle| \mathcal{F}_0 \right]. \tag{4.2}$$

The difficulty in pricing CMS spread derivatives arises from the fact that unlike a single interest rate, a CMS spread rate can take both positive and negative values. The yield curve can move in a way such that any part can be either flat, upward or downward sloping. It is this feature that adds an extra complication in the pricing of derivative instruments for which a CMS spread rate is the underlying.

4.3 Pricing Approach

Various approaches have been developed to value financial derivatives on spread rates. In Chapter 3 we have seen that TSR models can be used to price CMS derivatives, which are single-rate derivatives. The TSR models are much more convenient and easier from a mathematical point of view compared to full term-structure models. It is generally more difficult to fit the market-implied distribution of one particular rate with a model that simultaneously specifies the dynamics of the whole yield curve. In addition, the TSR models are usually faster then a full term-structure model. Given the highly traded volumes in many derivatives markets, there is often not much room for pricing errors due to not being able to fit market-observable prices. The same holds true for pricing CMS spread options, which have become fairly liquid in recent years. Closed-form solutions for CMS spread options can be obtained only in rare cases, such as the case of caplets and floorlets with zero strike in which the *Margrabe exchange option formula* can be used, [29]. Most research regarding the valuation of CMS spread derivatives involves the Libor Market Model (LMM) or Swap Market Model (SMM) (see [1], [3], [27], and [25]). It is commonly assumed that each rate used to calculate the spread is log-normally distributed and there may be a nonzero correlation between them. A downside to this approach is that it has limited analytical tractability, as the linear combination of log-normal variables has an unknown distribution. Our aim is to have a fast, analytical tractable and flexible (two-rate) model that specifies the joint dynamics of only the two underlying swap rates.

[1]From here on when we talk about a CMS spread option we mean a CMS spread caplet unless specified otherwise.

[2]The discounted value can be obtained by multiplying $P(0, T_p)$ to the RHS of equation (4.2).

We recall that we can obtain the PDF $\psi^{A_i}(\cdot)$ of each swap rate in the annuity measure from the market prices of swaptions across strikes. By specifying annuity mapping functions $\alpha_i(\cdot)$ given by,

$$\alpha_i(s) = \mathbb{E}^{A_i}\left[\left.\frac{P(T,T_p)}{A_i(T)}\right| S_i(T) = s\right], \tag{4.3}$$

we can obtain the PDF of each swap rate in the T_p-forward measure (see Section 4.5). The formula (4.3) is exact. In numerical calculations, (4.3) will be approximated by the annuity function of a chosen TSR model (see Chapter 3).

An important observation to make is that this approach will lead to some inconsistencies, since each quantity $P(T,T_p)/A_i(T)$, $i = 1,2$ will generally depend on both swap rates $S_1(T)$ and $S_2(T)$. Therefore, the calculation of $\alpha_i(s)$ should incorporate the dependence structure of both rates. However, for tractability reasons the measure change related calculations are done independently from the dependence structure modeling.

An important observation to make is that this approach will lead to some inconsistencies, since each quantity $P(T,T_p)/A_i(T)$, $i = 1,2$ will generally depend on both swap rates $S_1(T)$ and $S_2(T)$. Therefore, the calculation of $\alpha_i(s)$ should incorporate the dependence structure of both rates. However, for tractability reasons the measure change related calculations are done independently from the dependence structure modeling.

4.4 Copula

When the marginal distributions of $S_i(T)$ under the T_p-forward measure are determined, the joint distribution of $(S_1(T), S_2(T))$ can be obtained by linking the margins with a so-called *copula*. The word copula originates from Latin, meaning 'tie, connection or link'. So by using a copula we are able to connect or couple marginal distributions into a multivariate distribution. Since we will be dealing with CMS spread options, where we have two underlying swap rates we will from here on focus our attention on two-dimensional copulas, also called *bivariate copulas*. For a thorough treatment of copulas including the multivariate case we refer the reader to [32] and [24]. The formal definition of a bivariate copula is given as follows, [32]:

Definition 4.4.1 (Bivariate Copula). *A bivariate copula is a function $C : [0,1]^2 \to [0,1]$ that has the following three properties:*

1. $\forall u, v \in [0,1]$,
$$C(u,0) = C(0,v) = 0, \tag{4.4}$$
2. $\forall u, v \in [0,1]$,
$$C(u,1) = u, \quad C(1,v) = v, \tag{4.5}$$
3. $\forall u_1, u_2, v_1, v_2 \in [0,1]$ *with* $u_2 \geq u_1, v_2 \geq v_1$,
$$C(u_2,v_2) - C(u_1,v_2) - C(u_2,v_1) + C(u_1,v_1) \geq 0. \tag{4.6}$$

Property 1 is called the *groundness* property of a function. A function satisfying Property 3 is called a *2-increasing* function, this can be seen as the two-dimensional analogue of a nondecreasing one-dimensional function. As a consequence of property 1 and 3 additional properties follow for a copula function, which we present in Lemma 4.4.2.

Lemma 4.4.2 (Additional Properties Bivariate Copula). *Let C be a bivariate copula. As a consequence of the groundedness and 2-increasing property for copulas, the following additional properties follow:*

1. *C is nondecreasing in each variable.*
2. *C satisfies the following Lipschitz condition $\forall u_1, u_2, v_1, v_2 \in [0,1]$,*

$$|C(u_2, v_2) - C(u_1, v_1)| \leq |u_2 - u_1| + |v_2 - v_1|, \tag{4.7}$$

   it follows that every copula C is uniformly continuous on its domain.

3. *$\forall u \in [0,1]$, the partial derivative $\frac{\partial C(u,v)}{\partial v}$ exist for almost every $v \in [0,1]$. For such u and v it follows that,*

$$0 \leq \frac{\partial C(u,v)}{\partial v} \leq 1. \tag{4.8}$$

   The analogous statement is true for the partial derivative $\frac{\partial C(u,v)}{\partial u}$.

4. *The functions*

$$u \to \frac{\partial C(u,v)}{\partial v}, \quad v \to \frac{\partial C(u,v)}{\partial u} \tag{4.9}$$

   are well-defined and nondecreasing a.e. on $[0,1]$.

Definition 4.4.1 and its additional properties given by Lemma 4.4.2 are rather technical. Therefore, we will formulate an alternative definition of a bivariate copula with which we will work, [37].

Definition 4.4.3 (Bivariate Copula (Alternative)). *Consider a function $C : [0,1]^2 \to [0,1]$. $C(u_1, u_2)$ is said to be a bivariate copula function if it defines a valid joint distribution function for a 2-dimensional vector of random variables, with each variable being uniformly distributed on $[0,1]$.*

The main attraction of the copula approach is that a copula separates the dependence structure of a multivariate distribution from its marginal distributions. The fundamental result of copula theory is Sklar's Theorem. But before we present Sklar's Theorem we first state the following useful result which is needed for the proof of Sklar's Theorem.

Lemma 4.4.4. *Let Ψ be a continuous distribution function. Then,*

$$X \sim \Psi \;\; \text{if and only if} \;\; \Psi(X) \sim U[0,1]. \tag{4.10}$$

Proof. The proof is given in Appendix A.12 □

The two-dimensional version of Sklar's Theorem is given by Theorem 4.4.5.

Theorem 4.4.5 (Sklar's Theorem (2-D version)). *Let Ψ_C be a two-dimensional joint distribution function with marginal distribution functions Ψ_1, Ψ_2. Then there exists a two-dimensional copula C such that for all[3] $x \in \bar{\mathbb{R}}^2$,*

$$\Psi_C(x_1, x_2) = C(\Psi_1(x_1), \Psi_2(x_2)). \tag{4.11}$$

The bivariate copula C is uniquely determined in (4.11) if the marginals Ψ_1, Ψ_2 are continuous. Otherwise, C is only uniquely determined on $ran(\Psi_1) \times ran(\Psi_2)$, where $ran(\Psi_i)$ is the range of the function Ψ_i. Conversely, if C is a bivariate copula and Ψ_1, Ψ_2 are univariate distribution functions, then the function Ψ_C defined in (4.11) is a bivariate distribution function with marginals Ψ_1, Ψ_2.

[3]Here $\bar{\mathbb{R}}$ is the extended real line.

Proof. The proof is given in Appendix A.13. □

So with Sklar's Theorem it is possible to uniquely define a copula for a given bivariate distribution function Ψ_C if the marginals are continuous. That is, we have to know the joint distribution function of x_1, x_2 to determine the copula function. However, we are actually interested in the converse statement of Sklar's Theorem. Specifically, given a bivariate copula function and two marginal CDFs we can construct a two-dimensional joint distribution function. Next, we state the following useful lemma:

Lemma 4.4.6 (Joint PDF of a Bivariate Copula). *The joint PDF, ψ_C, associated with the joint CDF, Ψ_C, in (4.11) of a bivariate copula is defined as*

$$\psi_C(x_1, x_2) = c(\Psi_1(x_1), \Psi_2(x_2))\psi_1(x_1)\psi_2(x_2), \tag{4.12}$$

where

$$c(u_1, u_2) = \frac{\partial^2}{\partial u_1 \partial u_2} C(u_1, u_2). \tag{4.13}$$

Proof. The proof is given in Appendix A.14. □

There are several famous families of copulas, such as the Gaussian copula, Student-t copula and the Archimedean copula. As introductory examples we will first consider for the two-dimensional case the *independence copula*, C_{ID}, the *perfect dependence copula*, C_{D}, and the *perfect anti-dependence copula* C_{AD}. The independence copula is perhaps the simplest example of a copula. We introduce the uniform random variables U_1, U_2. The two uniform random variables underlying the copula are assumed to be independent. The perfect dependence copula can then be formally defined as follows:

Definition 4.4.7 (Independence Copula). *The copula function for the independence copula is given by,*

$$C_{ID} = u_1 \cdot u_2. \tag{4.14}$$

To obtain the copula that defines perfect dependence we set $U_1 = U_2$. We get:

$$\begin{aligned} C_D(u_1, u_2) &= \mathbb{P}(U_1 \leq u_1, U_2 \leq u_2) \\ &= \mathbb{P}(U_1 \leq u_1, U_1 \leq u_2) \\ &= \mathbb{P}(U_1 \leq \min\{u_1, u_2\}) \\ &= \min\{u_1, u_2\}. \end{aligned}$$

The perfect dependence copula can be formally defined as follows:

Definition 4.4.8 (Perfect Dependence Copula). *The copula function for the perfect dependence copula is given by,*

$$C_D(u_1, u_2) = \min\{u_1, u_2\}. \tag{4.15}$$

The perfect anti-dependence copula is obtained by choosing $U_2 = 1 - U_1$, we obtain:

$$\begin{aligned} C_{AD}(u_1, u_2) &= \mathbb{P}(U_1 \leq u_1, U_2 \leq u_2) \\ &= \mathbb{P}(U_1 \leq u_1, 1 - U_1 \leq u_2) \\ &= \mathbb{P}(1 - u_2 \leq U_1 \leq u_1) \\ &= (u_1 + u_2 - 1)^+. \end{aligned}$$

So the perfect anti-dependence copula can be formally defined as follows:

Definition 4.4.9 (Perfect Anti-Dependence Copula). *The copula function for the perfect anti-dependence copula is given by,*

$$C_{AD}(u_1, u_2) = (u_1 + u_2 - 1)^+. \tag{4.16}$$

Figure 4.1 shows plots with the three described copulas. From Figure 4.1 it is clear that the perfect anti-dependence copula bounds the independence copula from below, while the perfect dependence copula bounds the dependence copula from above. In fact, these results hold for all copulas, see Theorem 4.4.10.

Theorem 4.4.10 (Fréchet-Hoeffding Bounds). *Any valid two-dimensional copula function C must satisfy the Fréchet-Hoeffding bounds,*

$$C_{AD}(u_1, u_2) \leq C(u_1, u_2) \leq C_D(u_1, u_2), \tag{4.17}$$

where $C_{AD}(u_1, u_2)$ and $C_D(u_1, u_2)$ are given by (4.16) and (4.15).

Proof. The proof can be found in [32]. □

So, the perfect dependence and anti-dependence copulas can be used to bound any copula. The perfect anti-dependence and dependence copulas are also referred to as the lower and upper Fréchet-Hoeffding bounds.

Next, we will consider the copula that we will apply in the pricing of our CMSSOs, the Gaussian copula. The Gaussian copula is the copula that is most widely known and used in finance. In particular, we will consider the two-dimensional Gaussian copula, also called the *bivariate Gaussian copula*. The bivariate Gaussian copula is constructed from the bivariate normal distribution via Sklar's Theorem. But before we present the definition of the bivariate Gaussian copula, we first give the definition of Pearson's correlation coefficient[4] ϱ:

Definition 4.4.11 (Pearson's Correlation). *Let $(X_1, X_2)'$ be a random vector with both $\mathbb{E}[X_1^2], \mathbb{E}[X_2^2] < \infty$, then Pearson's correlation ϱ, also called the linear correlation coefficient, is defined by*

$$\varrho(X_1, X_2) \triangleq \frac{Cov(X_1, X_2)}{\sqrt{Var(X_1)\,Var(X_2)}}, \tag{4.18}$$

with

$$Cov(X_1, X_2) = \mathbb{E}[X_1 X_2] - \mathbb{E}[X_1]\mathbb{E}[X_2]. \tag{4.19}$$

The definition of the bivariate Gaussian copula is as follows:

Definition 4.4.12 (Bivariate Gaussian Copula). *A bivariate Gaussian copula is a copula function of the form*

$$C_{gauss}(u_1, u_2; \varrho) = \Phi_2(\Phi^{-1}(u_1), \Phi^{-1}(u_2); \varrho), \tag{4.20}$$

where $\Phi_2(\cdot)$ denotes the joint distribution of two-dimensional standard normal marginal distributions, with correlation coefficient ϱ given by

$$\Phi_2(h, k; \varrho) = \int_{-\infty}^{h} \int_{-\infty}^{k} \phi_2(x, y; \varrho) dy dx, \tag{4.21}$$

with

$$\phi_2(x, y; \varrho) = \frac{1}{2\pi\sqrt{1-\varrho^2}} \exp\left(-\frac{x^2 - 2\varrho xy + y^2}{2(1-\varrho^2)}\right), \tag{4.22}$$

[4] Pearson's correlation only accounts for linear correlation and thus does not measure any higher-order dependence.

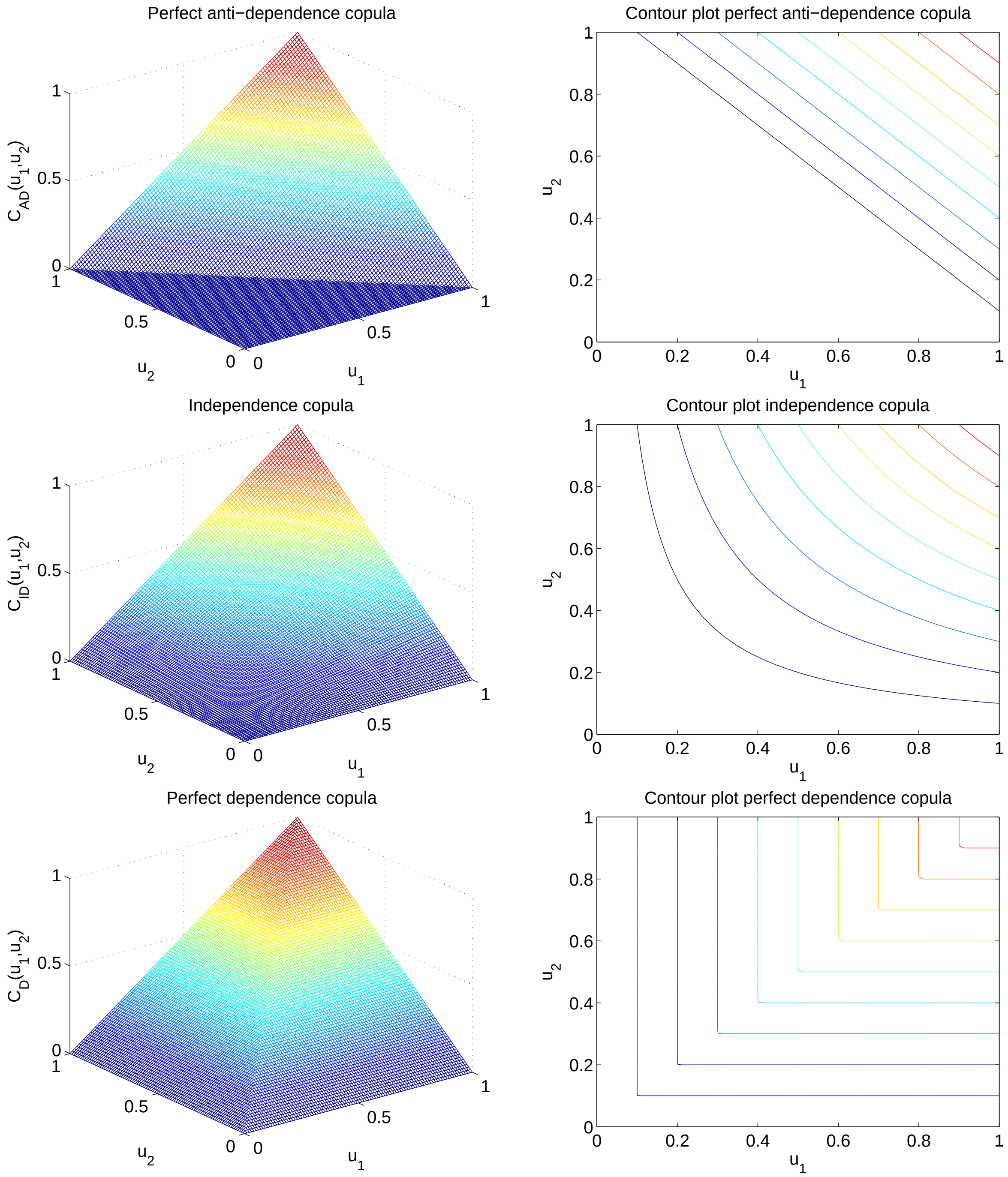


Figure 4.1: The perspective plots and contour plots of the CDFs of the perfect anti-dependence, independence and perfect dependence copulas.

and $\Phi(\cdot)$ *is the standard normal distribution function given by*

$$\Phi(h) = \int_{-\infty}^{h} \phi(x)dx, \tag{4.23}$$

with

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right), \tag{4.24}$$

the standard normal density function.

Figure 4.2 shows 1000 random samples from a Gaussian copula with different values of correlation coefficient ϱ.

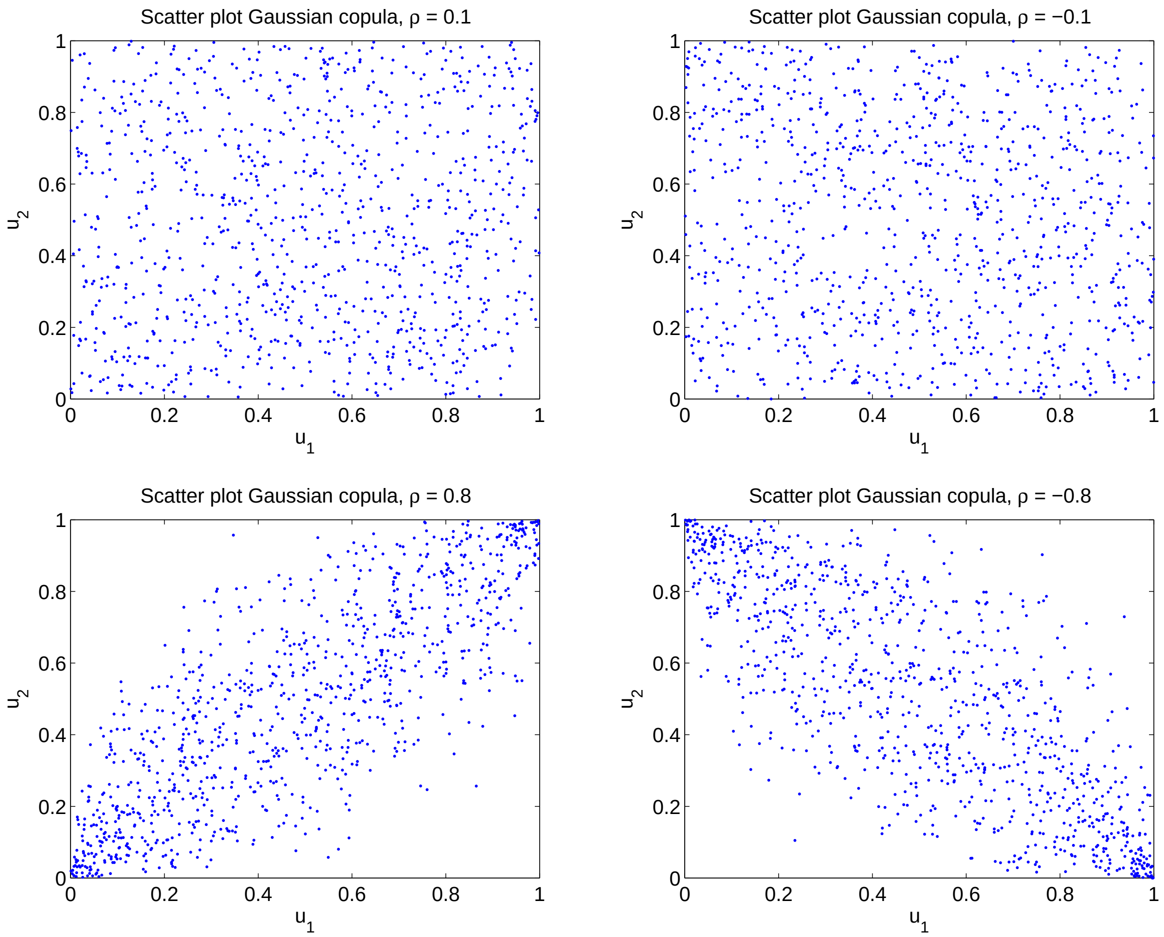


Figure 4.2: Scatter plot Gaussian copula for different values of correlation, 1000 samples.

The definition of the bivariate Gaussian copula density, the PDF of the Gaussian copula is as follows:

Definition 4.4.13 (**Bivariate Gaussian Copula Density**). *The bivariate Gaussian copula density is given by:*

$$c_{gauss}(u_1, u_2; \varrho) = \frac{1}{\sqrt{1-\varrho^2}} \exp\left(\frac{2\varrho\Phi^{-1}(u_1)\Phi^{-1}(u_2) - \varrho^2(\Phi^{-1}(u_1)^2 + \Phi^{-1}(u_2)^2)}{2(1-\varrho^2)}\right). \tag{4.25}$$

In Figure 4.3 the PDF and CDF of the Gaussian Copula are shown for different values of the correlation coefficient ϱ.

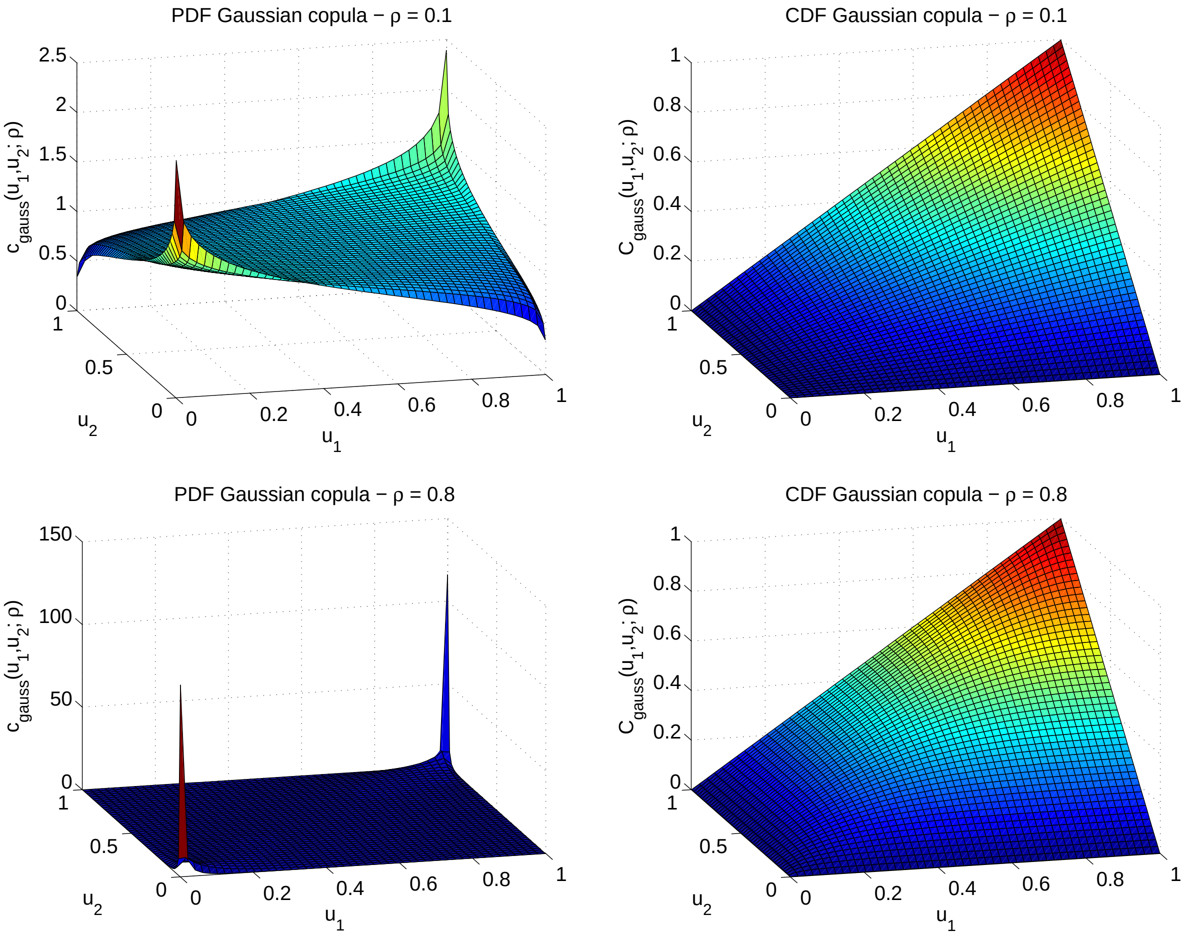


Figure 4.3: PDF and CDF for Gaussian copula for different values of correlation.

4.5 Pricing Formulas for CMSSOs

In order to evaluate expression (4.2) two different pricing formulas can be used, both having their pros and cons. When the joint density $\psi^{T_p}(x_1, x_2)$ of $S_1(T)$ and $S_2(T)$ is known, $V_{\mathrm{CMSSO}}(0)$ can be represented as an integral:

$$V_{\mathrm{CMSSO}}(0) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x_1 - x_2 - K)^+ \psi^{T_p}(x_1, x_2) dx_1 dx_2. \tag{4.26}$$

Now, if the dependence structure between the swap rates is defined by a copula $C(u_1, u_2)$ then we can apply Sklar's Theorem, Theorem 4.4.5, and obtain

$$V_{\mathrm{CMSSO}}(0) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x_1 - x_2 - K)^+ c\left(\Psi_1^{T_p}(x_1), \Psi_2^{T_p}(x_2)\right)\psi_1^{T_p}(x_1)\psi_2^{T_p}(x_2) dx_1 dx_2. \tag{4.27}$$

We note that for implementation the PDF $\psi_i^{T_p}(\cdot)$ and the CDF $\Psi_i^{T_p}(\cdot)$ are necessary. To obtain the PDF $\psi_i^{T_p}(\cdot)$ and CDF $\Psi_i^{T_p}(\cdot)$ under the T_p-forward measure we can make use of Lemma 4.5.1.

Lemma 4.5.1 (PDF/CDF of T_p-Forward Measure). *The PDF* $\psi^{T_p}(\cdot)$[5] *and CDF* $\Psi^{T_p}(\cdot)$

[5] In Lemma 4.5.1 and Lemma 4.5.2 the subscript i was dropped for both $\psi^{T_p}(\cdot)$ and $\Psi^{T_p}(\cdot)$ since the i merely indicates which swap rate is considered.

can be obtained via

$$\psi^{T_p}(x) = \frac{\partial^2 \mathbb{E}^{T_p}\left[(S(T)-x)^+ | \mathcal{F}_0\right]}{\partial x^2}, \tag{4.28}$$

$$\Psi^{T_p}(x) = 1 + \frac{\partial \mathbb{E}^{T_p}\left[(S(T)-x)^+ | \mathcal{F}_0\right]}{\partial x}. \tag{4.29}$$

Proof. The proof is given in [35, pp. 278-279]. □

Alternatively $\psi_i^{T_p}(\cdot)$ and $\Psi_i^{T_p}(\cdot)$ can be obtained by making use of the annuity mapping function $\alpha(\cdot)$ and PDF $\psi^A(\cdot)$ in the annuity measure directly, as is described in Lemma 4.5.2.

Lemma 4.5.2 (Linking PDF/CDF of T_p-Forward and Annuity Measure). *Given an annuity mapping function $\alpha(s)$ defined by (3.27), the PDF $\psi^{T_p}(s)$ and the CDF $\Psi^{T_p}(s)$ of the swap rate in the T_p-forward measure are linked to the PDF $\Psi^A(s)$ and the CDF $\Psi^A(s)$ of the swap rate in the annuity measure by*

$$\psi^{T_p}(s) = \frac{A(0)}{P(0,T_p)}\alpha(s)\psi^A(s), \tag{4.30}$$

$$\Psi^{T_p}(s) = \frac{A(0)}{P(0,T_p)}\int_{-\infty}^{s}\alpha(u)\psi^A(u)du, \tag{4.31}$$

where density $\psi^A(\cdot)$ is known from the market prices of swaptions, given by (3.9).

Proof. The proof is given in Appendix A.11. □

An important observation to make is that the main downside of (4.27) is that the integrals have to be truncated to apply a numerical integration scheme.

Besides valuation formula (4.27) another valuation formula can be obtained by applying a change-of-variables, namely $u_i = \Psi_i^{T_p}(x_i)$. This enables us to rewrite (4.27) as follows:

$$V_{\text{CMSSO}}(0) = \int_0^1\int_0^1 \left(\left[\Psi_2^{T_p}\right]^{-1}(u_2) - \left[\Psi_1^{T_p}\right]^{-1}(u_1) - K\right)^+ c\left(u_1,u_2\right)du_1du_2. \tag{4.32}$$

In this case the inverse CDF $\left[\Psi_i^{T_p}\right]^{-1}$ and the density $\psi_i^{T_p}$ are necessary for the implementation. The main advantage of this approach is that the domain of integration is now a bounded region, $[0,1]\times[0,1]$, which simplifies the discretization. In addition the marginal PDFs $\psi_i^{T_p}(x_i)$ are not necessary to evaluate the integral. A downside to this approach is that an efficient algorithm is needed for calculating the inverses of the marginal CDFs $\left[\Psi_i^{T_p}\right]^{-1}$. The CDFs and inverse CDFs are not available in closed form and must be calculated numerically. For efficiency, these inverse CDFs should always be pre-computed before the integration is performed.

Although both (4.27) and (4.32) can be used to price CMS derivatives, we are making use of a 2-dimensional integral in order to do so. This is generally not very efficient. Therefore, our aim will be to reduce the dimensionality.

4.5.1 Dimensionality Reduction for CMSSOs

To obtain a one-dimensional pricing formula that is based on a copula our starting point will be expression (4.26). We will reduce the dimensionality by making use of partial integration. We can rewrite (4.26) as follows,

$$\begin{aligned} V_{\text{CMSSO}}(0) &= \int_{-\infty}^{\infty} x_1 \left(\int_{-\infty}^{x_1 - K} \psi^{T_p}(x_1, x_2) dx_2 \right) dx_1 \\ &\quad - \int_{-\infty}^{\infty} (x_2 + K) \left(\int_{x_2 + K}^{\infty} \psi^{T_p}(x_1, x_2) dx_1 \right) dx_2. \end{aligned} \tag{4.33}$$

By definition,

$$\psi^{T_p}(x_1, x_2) = \frac{\partial^2}{\partial x_1 \partial x_2} \Psi^{T_p}(x_1, x_2). \tag{4.34}$$

Substituting (4.34) in (4.33) and integrating, we obtain,

$$\begin{aligned} V_{\text{CMSSO}}(0) &= \int_{-\infty}^{\infty} x_1 \left(\frac{\partial}{\partial x_1} \Psi^{T_p}(x_1, x_1 - K) - \frac{\partial}{\partial x_1} \Psi^{T_p}(x_1, -\infty) \right) dx_1 \\ &\quad - \int_{-\infty}^{\infty} (x_2 + K) \left(\frac{\partial}{\partial x_2} \Psi^{T_p}(\infty, x_2) - \frac{\partial}{\partial x_2} \Psi^{T_p}(x_2 + K, x_2) \right) dx_2. \end{aligned} \tag{4.35}$$

We have that,

$$\begin{aligned} \frac{\partial}{\partial x_1} \Psi^{T_p}(x_1, -\infty) &= 0, \\ \frac{\partial}{\partial x_2} \Psi^{T_p}(\infty, x_2) &= \psi_2^{T_p}(x_2). \end{aligned}$$

Therefore,

$$\begin{aligned} V_{\text{CMSSO}}(0) &= \int_{-\infty}^{\infty} x_1 \frac{\partial}{\partial x_1} \Psi^{T_p}(x_1, x_1 - K) dx_1 \\ &\quad - \int_{-\infty}^{\infty} (x_2 + K) \left(\psi_2^{T_p}(x_2) - \frac{\partial}{\partial x_2} \Psi^{T_p}(x_2 + K, x_2) \right) dx_2. \end{aligned} \tag{4.36}$$

We can rewrite (4.36) and obtain,

$$\begin{aligned} V_{\text{CMSSO}}(0) &= \int_{-\infty}^{\infty} x_1 \frac{\partial}{\partial x_1} \Psi^{T_p}(x_1, x_1 - K) dx_1 \\ &\quad + \int_{-\infty}^{\infty} (x_2 + K) \frac{\partial}{\partial x_2} \Psi^{T_p}(x_2 + K, x_2) dx_2 - \int_{-\infty}^{\infty} (x_2 + K) \psi_2^{T_p}(x_2) dx_2 \\ &= \int_{-\infty}^{\infty} x_1 \frac{\partial}{\partial x_1} \Psi^{T_p}(x_1, x_1 - K) dx_1 + \int_{-\infty}^{\infty} (x_2 + K) \frac{\partial}{\partial x_2} \Psi^{T_p}(x_2 + K, x_2) \\ &\quad - \int_{-\infty}^{\infty} x \psi_2^{T_p}(x) dx - K \int_{-\infty}^{\infty} \psi_2^{T_p}(x) dx. \end{aligned} \tag{4.37}$$

By definition of the PDF and expectation, we have,

$$\begin{aligned} V_{\text{CMSSO}}(0) &= \int_{-\infty}^{\infty} x_1 \frac{\partial}{\partial x_1} \Psi^{T_p}(x_1, x_1 - K) dx_1 + \int_{-\infty}^{\infty} (x_2 + K) \frac{\partial}{\partial x_2} \Psi^{T_p}(x_2 + K, x_2) \\ &\quad - \mathbb{E}^{T_p} \left[S_2(T) \middle| \mathcal{F}_0 \right] - K. \end{aligned} \tag{4.38}$$

Next, we define $\gamma(x, K)$ as follows,

$$\gamma(x, K) = \frac{d}{dx} \Psi^{T_p}(x, x - K). \tag{4.39}$$

So we can write

$$\gamma(x, K) = \frac{\partial}{\partial x_1}\Psi^{T_p}(x, x-K) + \frac{\partial}{\partial x_2}\Psi^{T_p}(x, x-K), \tag{4.40}$$

and

$$\int_{-\infty}^{\infty} x\gamma(x, K)dx = \int_{-\infty}^{\infty} x\frac{\partial}{\partial x_1}\Psi^{T_p}(x, x-K)dx + \int_{-\infty}^{\infty} x\frac{\partial}{\partial x_2}\Psi^{T_p}(x, x-K)dx. \tag{4.41}$$

Substituting $x = x_1$ in the first integral and $x = x_2 + K$ in the second integral, we get,

$$\int_{-\infty}^{\infty} x\gamma(x, K)dx = \int_{-\infty}^{\infty} x_1\frac{\partial}{\partial x_1}\Psi^{T_p}(x_1, x_1-K)dx_1 + \int_{-\infty}^{\infty} (x_2+K)\frac{\partial}{\partial x_2}\Psi^{T_p}(x_2+K, x_2)dx_2. \tag{4.42}$$

Thus,

$$V_{\text{CMSSO}}(0) = \int_{-\infty}^{\infty} x\gamma(x, K)dx - \mathbb{E}^{T_p}\left[S_2(T)|\,\mathcal{F}_0\right] - K. \tag{4.43}$$

Remember that we wish to obtain a one-dimensional pricing formula that is based on a copula. We can make use of integration by parts, so that we can write the integral of $\gamma(x, K)$ as follows,

$$\begin{aligned}
\int_{-\infty}^{\infty} x\gamma(x, K)dx &= \int_{-\infty}^{\infty} x\left(\frac{d}{dx}\Psi^{T_p}(x, x-K)\right)dx \\
&= \int_{-\infty}^{0} xd\Psi^{T_p}(x, x-K) - \int_{0}^{\infty} xd(1-\Psi^{T_p}(x, x-K)) \\
&= -\int_{-\infty}^{0} \Psi^{T_p}(x, x-K)dx + \int_{0}^{\infty}(1-\Psi^{T_p}(x, x-K))dx \\
&= \int_{-\infty}^{\infty}\left(\mathbb{1}_{\{x>0\}} - \Psi^{T_p}(x, x-K)\right)dx.
\end{aligned} \tag{4.44}$$

Substituting (4.44) in (4.43) we obtain,

$$V_{\text{CMSSO}}(0) = \int_{-\infty}^{\infty}\left(\mathbb{1}_{\{x>0\}} - \Psi^{T_p}(x, x-K)\right)dx - \mathbb{E}^{T_p}\left[S_2(T)|\,\mathcal{F}_0\right] - K.$$

The final step we have to take to obtain our desired result is an application of Sklar's Theorem, Theorem 4.4.5. From Sklar's Theorem it follows that,

$$\Psi_C^{T_p}(x, x-K) = C(\Psi_1^{T_p}(x), \Psi_2^{T_p}(x-K)). \tag{4.45}$$

Hence,

$$V_{\text{CMSSO}}(0) = \int_{-\infty}^{\infty}\left(\mathbb{1}_{\{x>0\}} - C(\Psi_1^{T_p}(x), \Psi_2^{T_p}(x-K))\right)dx - \mathbb{E}^{T_p}\left[S_2(T)|\,\mathcal{F}_0\right] - K.$$

The obtained result is summarized by Lemma 4.5.3.

Lemma 4.5.3 (CMSSO 1-D Pricing Formula). *Assume that CMS rates $S_1(T)$, $S_2(T)$ have marginal CDFs $\Psi_i^{T_p}$ under the T_p-forward measure and marginal PDFs $\psi_i^{T_p}$, $i = 1, 2$ that can be coupled with a copula C. The undiscounted price of a CMS spread caplet with maturity T, payment date T_p and strike K is then given by*

$$V_{CMSSO}(0) = \int_{-\infty}^{\infty}\left(\mathbb{1}_{\{x>0\}} - C(\Psi_1^{T_p}(x), \Psi_2^{T_p}(x-K))\right)dx - \mathbb{E}^{T_p}\left[S_2(T)|\,\mathcal{F}_0\right] - K. \tag{4.46}$$

The price of the CMSSO can now be calculated relatively easy with expression (4.46). The term $\mathbb{E}^{T_p}\left[S_2(T)|\mathcal{F}_0\right]$ can be calculated directly with a chosen TSR model. For the term that remains we have to select a copula and the CDFs $\Psi_i^{T_p}$ can be calculated making use of Lemma 4.5.1 or Lemma 4.5.2.

Remark 4.5.4. *Two important observations for Lemma 4.46 are:*

- *The one-dimensional pricing formula for the CMSSO given by (4.46) is valid for general copulas.*
- *The downside of choosing copulas other than the Gaussian copula is that extra parameters need to be specified, which are difficult or impossible to calibrate.*

4.5.2 Monte Carlo Method for CMSSOs

Another attractive feature of the Gaussian copula is that the value of the CMSSO given by (4.2) can be calculated relatively easy by a Monte Carlo (MC) method. In case of a Gaussian copula with correlation matrix R given by

$$R = \begin{pmatrix} 1 & \varrho \\ \varrho & 1 \end{pmatrix}, \tag{4.47}$$

the random variable S_i can be specified by:

$$S_i = \left[\Psi_i^{T_p}\right]^{-1}(\Phi(Z_i)), \tag{4.48}$$

where Z_1 and Z_2 are standard normal random variables with correlation ϱ. We can calculate the value of the CMSSO as follows:

$$V_{\text{CMSSO}}(0) \approx \frac{1}{N}\sum_{n=1}^{N}\left(\left[\Psi_1^{T_p}\right]^{-1}(\Phi(Z_{n,1})) - \left[\Psi_2^{T_p}\right]^{-1}(\Phi(Z_{n,2})) - K\right)^+, \tag{4.49}$$

where $\mathbf{Z}_1, \ldots, \mathbf{Z}_N$, with $\mathbf{Z}_n = (Z_{n,1}, Z_{n,2})$, are N independent samples from a two-dimensional Gaussian distribution.

4.6 Numerical Experiments

In this section we will perform numerical experiments to get better insight in the copula approach for pricing CMS spread options. Additionally, more insight is gained in the respective TSR models. The same market data is used as in Chapter 3. The market data for the respective swap rates for both 2007 and 2013 can be found in Appendix B. The correlation parameter ϱ, which specifies the correlation of the two swap rates is given to be $\varrho = 0.8$. We will calculate the price of a CMSSO using pricing formula (4.46). As copula the Gaussian copula is chosen, furthermore all four different TSR models will be considered. As reference we choose the MC method given by (4.49), where in this case the chosen TSR model is the swap-yield TSR model. The price that is calculated with the reference model will be denoted by V_{ref}. The difference between the price computed with pricing formula (4.46) and the reference will be denoted by ς, where ς is defined as:

$$\varsigma = V_{\text{CMSSO}}(0) - V_{\text{ref}}.$$

Besides looking at the difference in prices we will also study the timing, volatility and correlation effects.

4.6.1 CMSSO Price: 2007 vs 2013

The first numerical experiment we will perform is the pricing of a CMSSO on a 10Y-2Y spread with 12M frequency for the market data of 2007 and 2013. The start date is taken to be 1 year from today. The computed CMSSO prices and the difference in prices, ς, for both 2007 and 2013 are given in Figure 4.4.

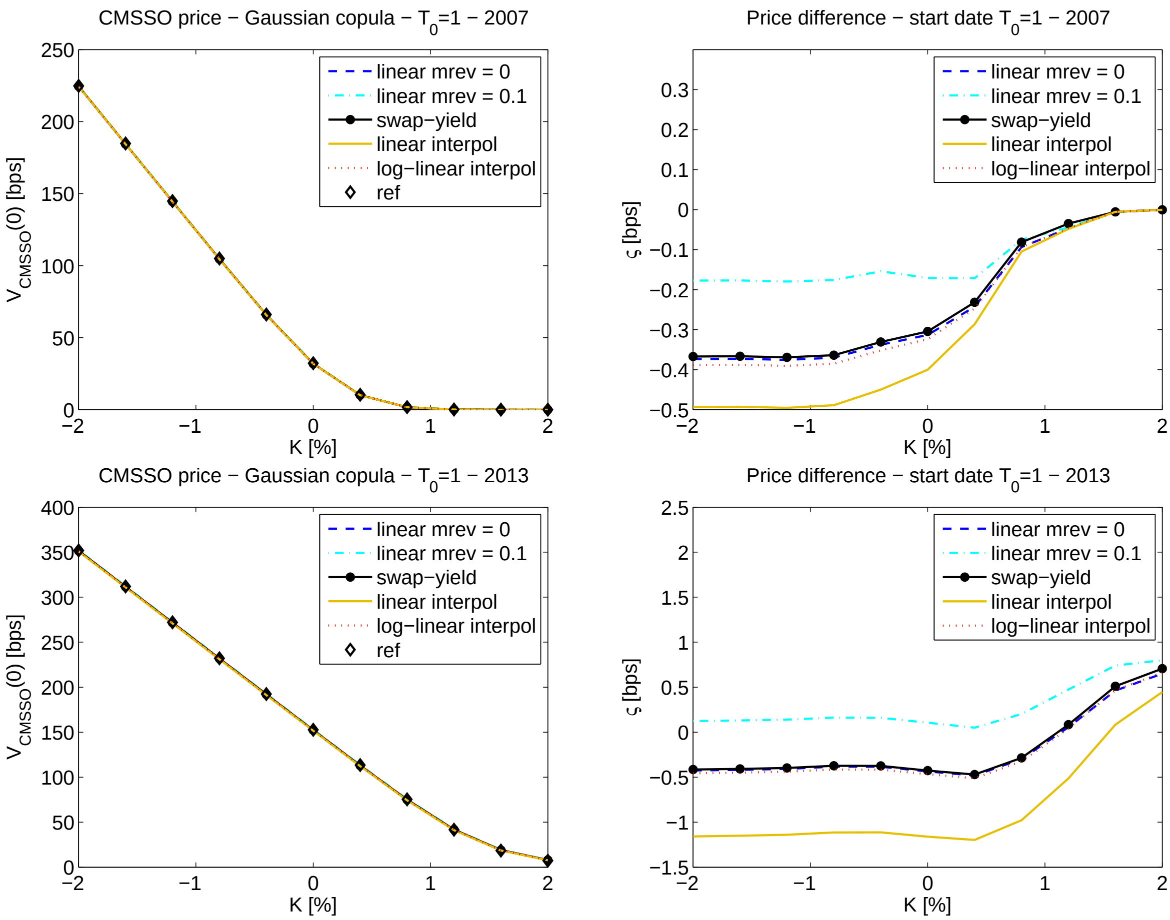


Figure 4.4: Prices of a CMSSO on 10Y-2Y spread with 12M frequency for 2007 and 2013 using the copula approach. Different TSR models in combination with a Gaussian copula are used. As reference the MC method is used with the swap-yield TSR model. The number of independent samples for the MC method is $N = 50000$.

From Figure 4.4 it is clear that for both 2007 and 2013 the CMSSO price computed using pricing formula (4.46) is very close to the chosen reference price. Furthermore, the CMSSO price computed with the different TSR models only differs slightly. The price differences are generally below 1bp. The results of the swap-yield TSR model, the mean reversion linear TSR model with $\varkappa = 0$ and the log-linear TSR model are almost identical for the market data from 2007 and 2013. The results of the linear interpolation TSR model differ the most from the reference model, but this can be attributed to the fact that the swap-yield TSR model is chosen as the reference TSR model in the MC method. Finally, we observe that the differences for 2013 are bigger than the differences for 2007. We also observed this behaviour when we priced CMS options, see Chapter 3. In Chapter 3 we also mentioned that the bigger differences we observe for 2013 probably stem from the fact that the volatilities are more extreme for 2013 when compared to 2007. So, also in the case of pricing CMSSOs it is nowadays even more important to choose the correct TSR model.

4.6.2 Investigate the Timing Effect

Next, we wish to investigate the timing effect. We proceed in the same manner as was done in Section 3.6 and consider start dates up to 10 years from today. The results for start date $T_0 = 1$ were shown in Figure 4.4. We present the results for start date $T_0 = 5$ for 2007 and 2013 in Figure 4.5.

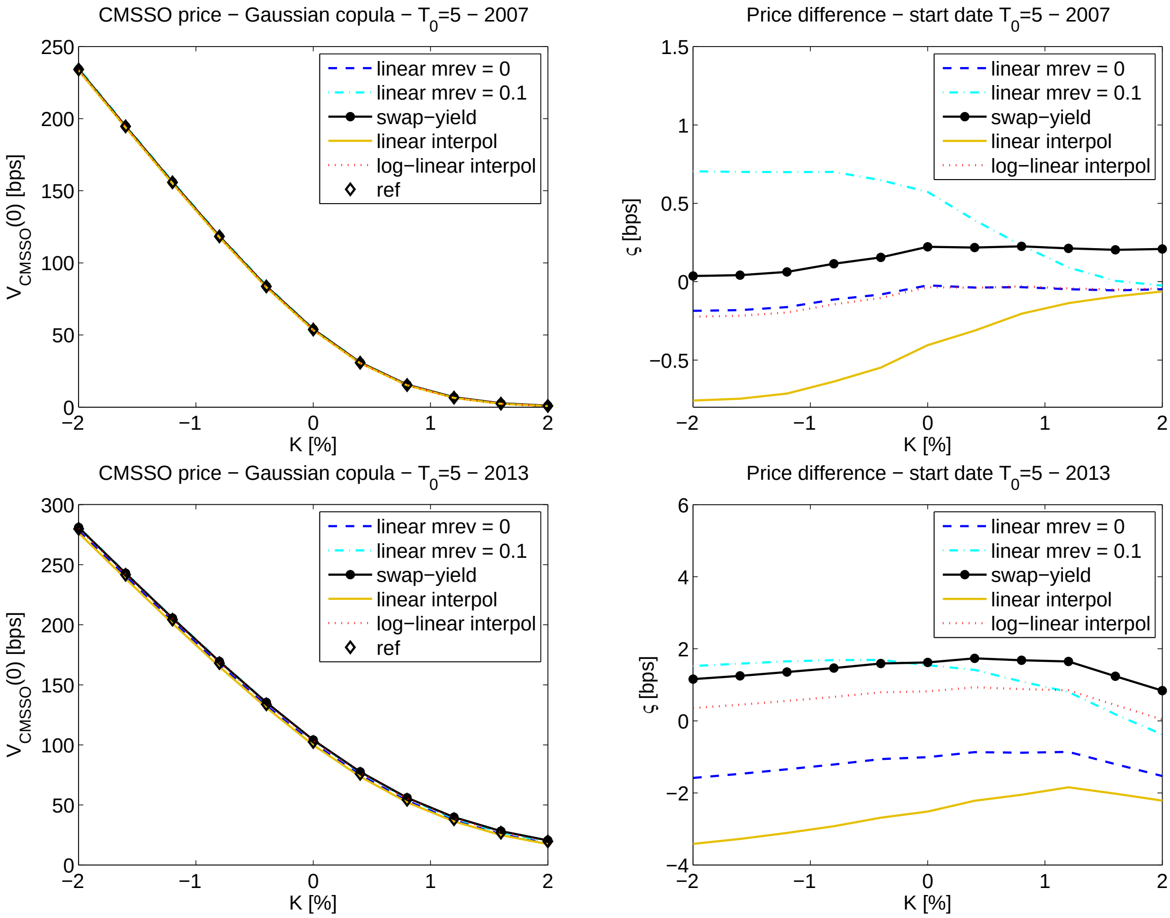


Figure 4.5: Prices of a CMSSO on 10Y-2Y spread with 12M frequency for 2007 and 2013 using the copula approach. Different TSR models in combination with a Gaussian copula are used. As reference the MC method is used with the swap-yield TSR model. The number of independent samples for the MC method is $N = 100000$. The start date is $T_0 = 5$.

Looking at Figure 4.5 we can see that the price differences for 2007 are generally smaller than 1bp and for 2013 they are smaller than 2bps. So for a start date 5 years from today the copula approach seems to give an accurate approximation of the CMSSO price.

Comparing the results presented in Figure 4.4 with the results presented in Figure 4.5 we see that as the start date has moved further into the future the CMSSO price computed with the different TSR models will differ more.

We can also make some other interesting observations. For start date $T_0 = 1$ the results of the swap-yield TSR model, the mean reversion linear TSR model with $\varkappa = 0$ and the log-linear TSR model are almost identical, which is not the case for start date $T_0 = 5$. It is also interesting to note that for 2013 the prices computed using the mean reversion linear TSR model with

$\varkappa = 0.1$ are closer to the reference prices than when $\varkappa = 0$. This is opposite from what we have seen so far. So obtaining a proper value for the mean reversion parameter is important.

Next, we consider a start date 10 years from today. In Figure 4.6 again the CMSSO price and ς are given for both 2007 and 2013.

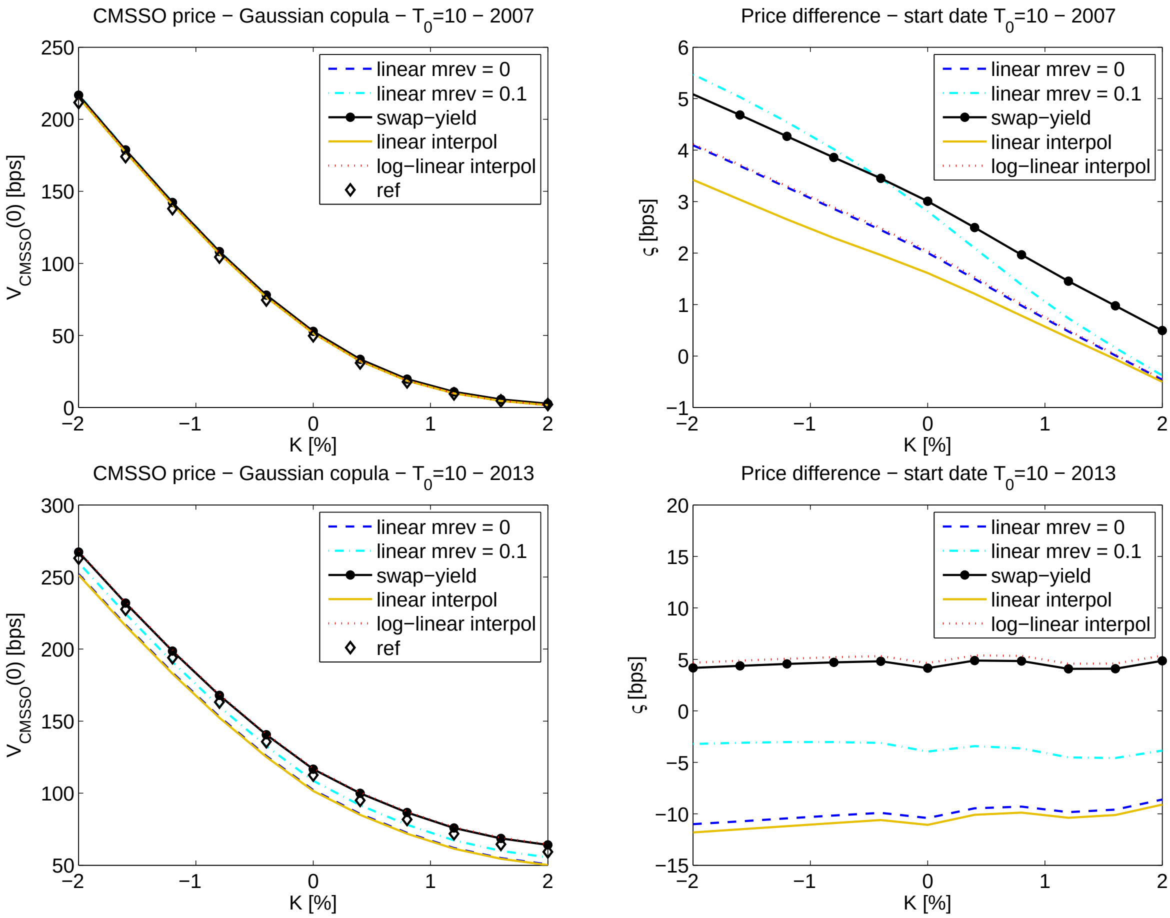


Figure 4.6: Prices of a CMSSO on 10Y-2Y spread with 12M frequency for 2007 and 2013 using the copula approach. Different TSR models in combination with a Gaussian copula are used. As reference the MC method is used with the swap-yield TSR model. The number of independent samples for the MC method is $N = 200000$. The start date is $T_0 = 10$.

Figure 4.6 shows that the price differences for 2007 are generally smaller than 6bps and for 2013 they are smaller than 11bp. These are satisfactory results considering the start date is 10 years from today. In fact the differences in price using the one-dimensional pricing formula and the MC method can be explained mostly by numerical issues. To obtain accurate results, especially for the market data from 2013, the number of MC paths had to be increased. Furthermore, the integration bounds needed in valuation formula (4.5.3) had to be set to larger values as the start date increased. For the three start dates we considered, the behavior of the two new TSR models, the linear interpolation and log-linear interpolation TSR models, is very satisfactory.

We have seen that the further the start date is moved into the future, the bigger the differences are between the computed prices of the CMS spread option with the respective TSR models, indicating that there is a timing effect.

4.6.3 Investigate the Volatility and Correlation Effect

To investigate the volatility and correlation effects, we are again going to price a CMSSO on 10Y-2Y spread. Part of the data from 2007 is used when the start date is 1 year from today, only this time we assume a flat volatility. We will consider combination of both a low and high constant volatility and a low and high correlation. The results are given in Figure 4.7.

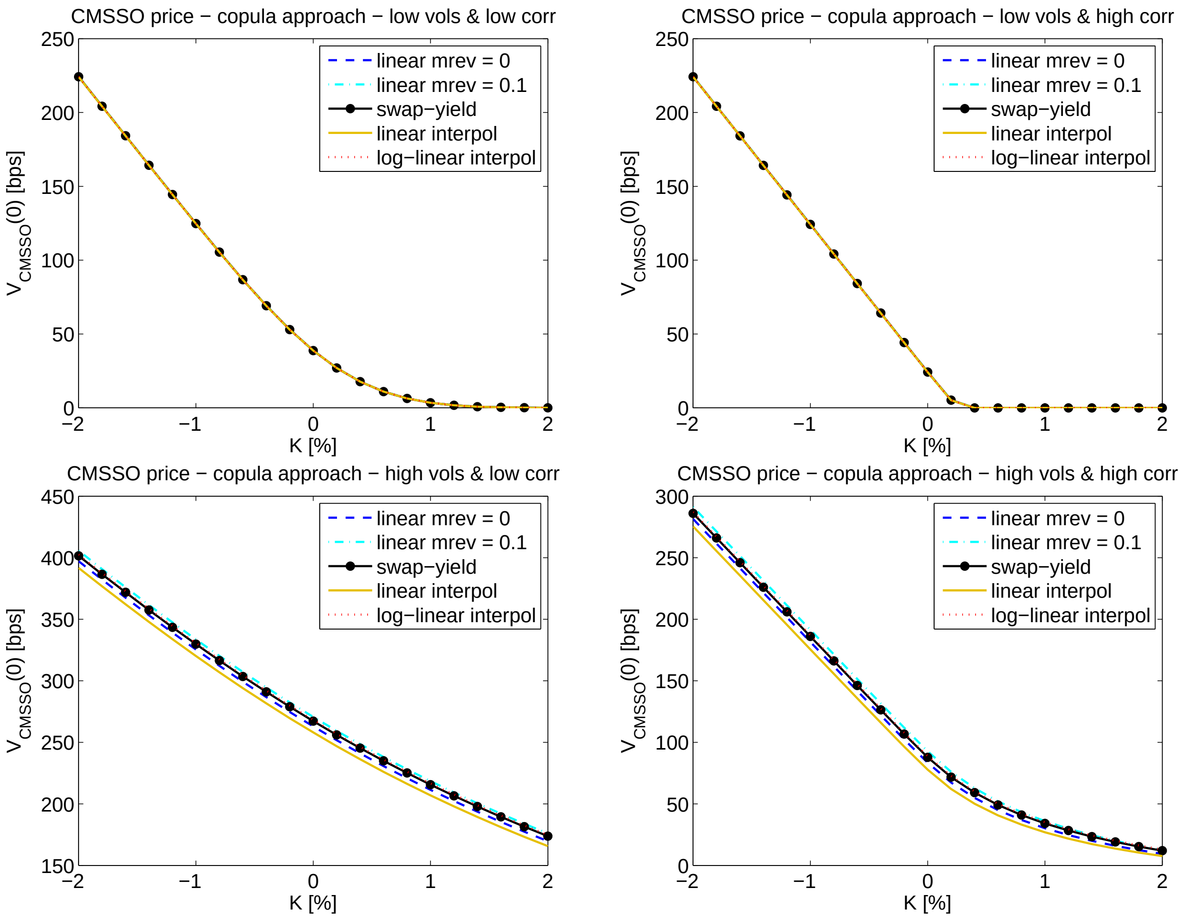


Figure 4.7: Prices of a CMSSO on 10Y-2Y spread with 12M frequency using the copula approach. Different TSR models and a Gaussian copula are used. The volatility is assumed to be flat. Both a low, $\sigma^1_{\text{low}} = \sigma^2_{\text{low}} = 0.1$, and high, $\sigma^1_{\text{high}} = \sigma^2_{\text{high}} = 0.9$, volatility is considered. Additionally, a low, $\varrho_{\text{low}} = 0.01$, and high, $\varrho_{\text{high}} = 0.99$, correlation is considered.

Figure 4.7 shows that as the volatility increases the CMSSO price increases. Furthermore, for higher volatilities the CMSSO price computed with the different TSR models will differ more. Note that we already observed the same behaviour when we priced CMS options, Chapter 3.

Another important observation from Figure 4.7 is that the correlation parameter, ϱ, has influence on the curvature. The higher the value of the correlation parameter is, the more convex the plot of the computed CMSSO price becomes. Therefore, we can conclude that there is indeed a volatility and a correlation effect.

4.7 Conclusions

A CMS spread derivative is a financial instrument whose payoff is a function of the spread between two swap rates of different maturity. The difficulty in pricing CMS spread derivatives arises from the fact that unlike a single interest rate, a CMS spread rate can achieve both positive and negative values. The yield curve can move in a way that any part can be either flat, upward or downward sloping. It is this feature that adds an extra complication in the pricing of derivative instruments for which a CMS spread rate is the underlying.

An important widely used type of CMS spread derivative are CMS spread options. Banks typically use CMS spread options to hedge the CMS spread swaps that they have entered into with customers. The notional values for these kind of derivatives are usually quite large. So even small differences can lead to substantial losses. Therefore, the pricing of these products has to be efficient and accurate.

Most research regarding the valuation of CMS spread options involves the Libor or Swap Market Models. A downside of this approach is that it is rather time consuming and it has limited analytical tractability when applied to CMS spread derivatives. Our aim was to have a fast, analytically tractable and flexible (two-rate) model that specifies the joint dynamics of only the two underlying swap rates. In order to obtain our desired model we made use of the copula approach. We determined the marginal distributions of the two swap rates under the forward measure and the joint distribution can then be obtained by linking the marginal distributions using a copula function.

Different copulas were discussed, but our main focus was on the Gaussian copula, which we used for the pricing of CMS spread options. Furthermore, Sklar's Theorem was discussed which is a key component in the copula approach.

Using the copula approach a two-dimensional pricing formula can be obtained for the pricing of CMS spread options. Using integration by parts we derived a one-dimensional pricing formula from the two-dimensional pricing formula. A copula and a TSR model have to be selected in order to make use of this one-dimensional pricing formula.

We also presented a Monte Carlo method which can be used to price CMS spread options, in case a Gaussian copula is assumed.

Many numerical experiments were performed to study the performance of the copula approach and the respective TSR models. Market data from 2007 and 2013 was considered. In general, the results for both sets of market data were similar. The differences in prices using the one-dimensional pricing formula and the MC method can be explained mostly by numerical issues, i.e. large number of MC paths, chosen values for the integration bounds.

An important observation was that the differences for the year 2013 are bigger than for the year 2007, which is probably due to the fact that the volatilities in 2013 were more extreme. We can conclude that nowadays correct valuation of CMS spread options is of even greater importance.

Furthermore, we have seen that depending on the chosen TSR model the computed prices of the CMS spread option can differ slightly. We also showed that there are timing, volatility and correlation effects. The further the start date is moved into the future, the bigger the differences will be between the computed prices of the CMS spread option with the respective TSR models. The volatility effect follows from the fact that the higher the volatility is, the bigger the differences will be between the computed prices of the CMS spread option with the respective TSR models. The correlation parameter has influence on the curvature. The higher the value of the correlation parameter is, the more convex the plot of the computed CMS spread option price becomes. It was shown that the behavior of the two new TSR models, the linear interpolation and log-linear interpolation TSR model, is highly satisfactory.

Thus with the copula approach we can efficiently and accurately price CMS spread options.

Chapter 5

DD SABR Model for Pricing CMS Spread Derivatives

5.1 Introduction

Stochastic volatility models are often the preferred choice for pricing exotic derivatives. In this chapter our focus is on the pricing of CMS spread options using a stochastic volatility model. A CMS spread option is a European multi-rate option whose payoff is a function of the spread between two swap rates of different maturities, see Chapter 4. The distribution of each rate can be described by a stochastic volatility model. This enables us to define co-dependence between these rates by techniques other than the copula approach. Actually, if each swap rate involved in the payoff of a given multi-rate derivative has its own asset process and its own stochastic variance process, then the co-dependence structure between rates can be controlled by correlating the Brownian motions that drive the asset and stochastic variance process.

Remember that our aim is to obtain a model which can be used to efficiently and accurately price CMS spread options. We start with a two-dimensional version of the SABR model that can be used to price these derivatives. However, to compute the CMS spread option prices with this model we need to apply a MC simulation and this is not very efficient. Using the Markovian projection method we can obtain a model by which we can efficiently and accurately price CMS spread options. This chapter is organized as follows.

In Section 5.2 a two-dimensional version of the SABR model is introduced. In order to use the two-dimensional SABR model for CMS spread options pricing, CMS-adjusted forward rates and the associated adjusted SABR parameters are defined. Section 5.3 presents the Markovian projection method, which can be used to project a given model onto a simpler model. In Section 5.4 we present the displaced diffusion SABR model, by which we calculate the prices of CMS spread options efficiently. In particular, we show in detail how to project the two-dimensional SABR model onto the displaced diffusion SABR model for the spread. In Section 5.5 several numerical experiments are performed. The results of the copula approach and the displaced diffusion SABR model are compared. Finally, Section 5.6 concludes.

This chapter is based on [26, pp. 159-171] and [37, pp. 1129-1156].

5.2 Two-dimensional SABR Model

In this section we discuss a two-dimensional version of the SABR model, which can be used for the pricing of CMS spread options. A multi-dimensional version of the SABR model is described in [26, pp. 141-142].

In Chapter 3 we saw that for the pricing of CMS derivatives it is necessary to compute the expectation of the future CMS rates under the forward measure that is associated with the payment date. However, the natural martingale measure of the CMS rate (swap rate) is the annuity measure. Therefore, we cannot model them as driftless processes under the T_p-forward measure $\mathbb{Q}^{T_p}$. Assuming the drift term is given by μ_i we have:

$$\begin{aligned} dS_i(t) &= \mu_i dt + \alpha_i(t) S_i(t)^{\beta_i} dW_i^{T_p}(t), \\ d\alpha_i(t) &= \nu_i \alpha_i(t) dZ_i^{T_p}(t), \\ S_i(0) &= s_i^0, \\ \alpha_i(0) &= \alpha_i^0, \\ \langle dW_i^{T_p}(t), dW_j^{T_p}(t)\rangle &= \rho_{ij} dt, \\ \langle dW_i^{T_p}(t), dZ_j^{T_p}(t)\rangle &= \gamma_{ij} dt, \\ \langle dZ_i^{T_p}(t), dZ_j^{T_p}(t)\rangle &= \xi_{ij} dt, \quad i,j = 1,2. \end{aligned} \tag{5.1}$$

Here ρ_{ij} is the correlation between the Brownian motions driving the asset price processes, γ_{ij} is the so called cross-skew and ξ_{ij} is the so called de-correlation between the stochastic volatilities.

In order to avoid dealing with the drift terms in (5.1) we will consider an approach that can be seen as a combination of the approaches described in [26, pp. 159-171] and [37, pp. 804-805]. Although, some alterations had to be made to make it applicable for our problem.

We consider so-called CMS-adjusted forward rates instead of the actual CMS rates (swap rates). Drifts of the swap rates under the forward measure are rather complicated, see e.g. [9], while CMS-adjusted forward rates are martingales under the T_p-forward measure. The CMS-adjusted forward rate is formally defined as follows:

$$\tilde{S}_i(t) \triangleq \mathbb{E}^{T_p}[S_i(T_0)|\, \mathcal{F}_t]. \tag{5.2}$$

From (5.2) it follows that at expiry T_0 we have:

$$\tilde{S}_i(T_0) = \mathbb{E}^{T_p}[S_i(T_0)|\, \mathcal{F}_{T_0}] = S_i(T_0). \tag{5.3}$$

In Chapter 4 we have seen that the undiscounted value of a CMS spread option is given by:

$$V_{\text{CMSSO}}(0) = \mathbb{E}^{T_p}\left[(S_1(T_0) - S_2(T_0) - K)^+ \middle|\, \mathcal{F}_0 \right]. \tag{5.4}$$

Substituting (5.3) in (5.4) we obtain the following valuation formula for the CMS spread option:

$$V_{\text{CMSSO}}(0) = \mathbb{E}^{T_p}\left[\left(\tilde{S}_1(T_0) - \tilde{S}_2(T_0) - K\right)^+ \middle|\, \mathcal{F}_0 \right]. \tag{5.5}$$

We can now define a two-dimensional SABR (2D SABR) model that can be used for the pricing of CMSSOs.

Definition 5.2.1 (2D SABR Model for CMS-adjusted Forward Rates). *The stochastic dynamics for CMS-adjusted forward rate $\tilde{S}_i$ and associated stochastic volatility $\tilde{\alpha}_i$, where $i = 1, 2$*

are given by:

$$
\begin{aligned}
d\tilde{S}_i(t) &= \tilde{\alpha}_i(t)\tilde{S}_i(t)^{\tilde{\beta}_i} dW_i^{T_p}(t),\\
d\tilde{\alpha}_i(t) &= \tilde{\nu}_i\tilde{\alpha}_i(t)dZ_i^{T_p}(t),\\
\tilde{S}_i(0) &= \tilde{s}_i^0,\\
\tilde{\alpha}_i(0) &= \tilde{\alpha}_i^0,\\
\langle dW_i^{T_p}(t), dW_j^{T_p}(t)\rangle &= \tilde{\rho}_{ij}dt,\\
\langle dW_i^{T_p}(t), dZ_j^{T_p}(t)\rangle &= \tilde{\gamma}_{ij}dt,\\
\langle dZ_i^{T_p}(t), dZ_j^{T_p}(t)\rangle &= \tilde{\xi}_{ij}dt, \quad i,j = 1,2.
\end{aligned} \tag{5.6}
$$

where $\tilde{\rho}_{ij}$ is the correlation between the Brownian motions driving the CMS-adjusted forward rates, $\tilde{\gamma}_{ij}$ is the cross-skew and $\tilde{\xi}_{ij}$ is the de-correlation between the stochastic volatilities.

Since we are now using CMS-adjusted forward rates we also need to use adjusted SABR parameters. The quadruple of T_p-measure adjusted SABR parameters associated with CMS-adjusted forward rate $\tilde{S}_i(T_0)$ is given by:

$$(\tilde{\alpha}_i^0, \tilde{\beta}_i, \tilde{\gamma}_{ii}, \tilde{\nu}_i). \tag{5.7}$$

The CMS-adjusted rate $\tilde{S}_i(t)$ can be calculated using the replication method described in Lemma 3.3.1 with a chosen TSR model[1]. We also need a method to calibrate the adjusted SABR parameters. In Chapter 3 we have seen that the value of a CMS caplet with underlying swap rate $S_i(T_0)$ is given by:

$$V_{\text{CMScaplet}}(0) = \mathbb{E}^{T_p}\left[(S_i(T_0) - K)^+ \middle| \mathcal{F}_0\right]. \tag{5.8}$$

We can rewrite (5.8) in the following form for $\tilde{S}_i(T_0)$,

$$V_{\text{CMScaplet}}(0) = \mathbb{E}^{T_p}\left[\left(\tilde{S}_i(T_0) - K\right)^+ \middle| \mathcal{F}_0\right]. \tag{5.9}$$

CMS caplets are simply European call options on $\tilde{S}_i(T_0)$. CMS-adjusted forward rates are defined such that each CMS-adjusted rate follows SABR dynamics. Hence, we can obtain $(\tilde{\alpha}_i^0, \tilde{\beta}_i, \tilde{\gamma}_{ii}, \tilde{\nu}_i)$ by calibrating the SABR model as described in Chapter 2 to CMS caplets prices. The CMS caplet prices are computed using the replication method described in Lemma 3.3.1 with a chosen TSR model. On the other hand parameters $\tilde{\gamma}_{12}$, $\tilde{\gamma}_{21}$ and $\tilde{\xi}$ cannot be calibrated using the CMS caplet prices. In order to calibrate these parameters additional market date has to be used. Unfortunately, this type of market data is usually not available.

The correlation matrix of the 2D SABR model, denoted by $\mathbf{R}$, has the form

$$\mathbf{R} = \begin{pmatrix} 1 & R_{12}^{WW} & R_{11}^{WZ} & R_{12}^{WZ} \\ R_{12}^{WW} & 1 & R_{21}^{WZ} & R_{22}^{WZ} \\ R_{11}^{WZ} & R_{21}^{WZ} & 1 & R_{12}^{ZZ} \\ R_{12}^{WZ} & R_{22}^{WZ} & R_{12}^{ZZ} & 1 \end{pmatrix} = \begin{pmatrix} 1 & \tilde{\rho}_{12} & \tilde{\gamma}_{11} & \tilde{\gamma}_{12} \\ \tilde{\rho}_{21} & 1 & \tilde{\gamma}_{21} & \tilde{\gamma}_{22} \\ \tilde{\gamma}_{11} & \tilde{\gamma}_{21} & 1 & \tilde{\xi}_{12} \\ \tilde{\gamma}_{12} & \tilde{\gamma}_{22} & \tilde{\xi}_{21} & 1 \end{pmatrix}, \tag{5.10}$$

where $\tilde{\rho} = \tilde{\rho}_{12} = \tilde{\rho}_{21}$, $\tilde{\xi} = \tilde{\xi}_{12} = \tilde{\xi}_{21}$. Note that unlike in the copula approach, we now consider the full correlation structure including cross-skew (parameters: $\tilde{\gamma}_{12}, \tilde{\gamma}_{21}$) and de-correlation (parameter $\tilde{\xi}$). We can write the system of SDEs in (5.6) in matrix-vector notation:

$$d\mathbf{X}(t) = \mathbf{R}\mathbf{X}(t)d\mathbf{Y}(t). \tag{5.11}$$

[1] We make use of the swap-yield TSR model, since it is market standard.

We cannot efficiently price CMS spread option prices with the 2D SABR model, because a MC simulation has to applied. To apply a MC simulation we have to express the system of SDEs in uncorrelated Brownian motions. To do so we make use of the Cholesky decomposition. We will give the definition of the Cholesky decomposition, Definition 5.2.2, and an algorithm that can be used to obtain a Cholesky decomposition for a given matrix, Lemma 5.2.3.

Definition 5.2.2 (Cholesky Decomposition). *The decomposition*

$$\boldsymbol{M} = \boldsymbol{C}\boldsymbol{C}^{\boldsymbol{T}} \tag{5.12}$$

of any semi positive definite matrix M as a product of a nonsingular lower triangular matrix and its transpose is called a Cholesky decomposition.

Lemma 5.2.3 (Algorithm Cholesky Decomposition). *A Cholesky decomposition for matrix M can be obtained as follows:*

1. *Initialize $C_1 = \sqrt{m_{11}}$.*
2. *For $k = 2, \ldots, n$*
   (a) *Solve $C_{k-1}c_k = m_k$ for c_k (C_{k-1} is $k-1 \times k-1$: for $k = 2$ this is a 1×1 or scalar equation);*
   (b) *$c_{kk} = \sqrt{m_{kk} - c_k^T c_k}$;*
   (c) *$C_k = \begin{pmatrix} C_{k-1} & 0 \\ c_k^T & c_{kk} \end{pmatrix}$.*

Notation:

- *C_{k-1}: the $k-1 \times k-1$ upper left corner of C;*
- *m_k: the first $k-1$ entries in column k of M;*
- *c_k: the first $k-1$ entries in column k of C^T;*
- *m_{kk} and c_{kk}: the entries of M and C.*

Using a Cholesky decomposition our multi-factor system of SDEs can be expressed as:

$$d\mathbf{X}(t) = \mathbf{R}\mathbf{X}(t)\mathbf{C}d\tilde{\mathbf{Y}}(t), \tag{5.13}$$

where $\tilde{\mathbf{Y}}(t)$ is a vector of independent Brownian motions, and $\mathbf{C}$ is the lower triangular Cholesky matrix. We will approximate (5.6) by a first-order Taylor approximation scheme, better known as the Euler scheme. It is a known fact that approximating SABR dynamics with an Euler scheme introduces some bias. More efficient simulation schemes for the SABR model have been developed, [12]. However, applying them to the 2D SABR model is quite involved and can be considered a research topic in itself. Therefore, we will only consider an Euler scheme. Generally, to obtain satisfactory results by a MC simulation a small stepsize and a large number of paths have to be chosen. Next, we will discuss a relatively simple method that reduces the number of paths needed in the MC simulation to obtain accurate results. The method we are referring to is called *antithetic sampling*, which is based on the fact that if we have a random variable Z that has a standard normal distribution, $Z \sim N(0, 1)$, then also $-Z \sim N(0, 1)$. Suppose that $\hat{V}$ is the approximation obtained from MC, and $\tilde{V}$ is the one obtained using $-Z$. Now by taking the average

$$V = \frac{1}{2}\left(\hat{V} + \tilde{V}\right), \tag{5.14}$$

we obtain a new approximation. Since $\hat{V}$ and $\tilde{V}$ are both random variables we aim at:

$$\mathrm{Var}(V) < \mathrm{Var}(\hat{V}). \tag{5.15}$$

We have:

$$\mathrm{Var}(V) = \frac{1}{4}\mathrm{Var}\left(\hat{V} + \tilde{V}\right) = \frac{1}{4}\mathrm{Var}\left(\hat{V}\right) + \frac{1}{4}\mathrm{Var}\left(\tilde{V}\right) + \frac{1}{2}\mathrm{Cov}(\hat{V}\tilde{V})$$

So it is clear that,

$$\mathrm{Var}(V) \leq \frac{1}{2}\mathrm{Var}\left(\hat{V} + \tilde{V}\right). \tag{5.16}$$

If desired other MC methods besides antithetic sampling can be considered, such as variate recycling, control variates, stratified sampling and importance sampling.

Our aim is to obtain a model, which can be used to calculate spread option prices efficiently. As mentioned earlier this cannot be done with the 2D SABR model. However, using the *Markovian projection method* we can obtain a model that enables us to calculate spread option prices efficiently and even analytically. With the Markovian projection method we can project the 2D SABR model onto a so-called *displaced diffusion SABR model*.

5.3 Markovian Projection

In this section we discuss the Markovian projection method, [37, pp. 1129-1156]. The method is based on a fundamental result, [18]. In [34] the fundamental result is presented in a form that is convenient for our problem. We also present this form in Theorem 5.3.1.

Theorem 5.3.1 (Gyöngy). *Consider the stochastic process defined by*

$$dX(t) = \mu(t)dt + \sigma(t)dW(t), \tag{5.17}$$

where $\mu(\cdot)$ and $\sigma(\cdot)$ are adapted stochastic processes such that (5.17) admits a unique solution. Define $a(t,x)$ and $b(t,x)$ by:

$$a(t,x) \triangleq \mathbb{E}[\mu(t)|\, X(t) = x], \tag{5.18}$$

$$b^2(t,x) \triangleq \mathbb{E}[\sigma^2(t)\big|\, X(t) = x]. \tag{5.19}$$

Then the SDE

$$dY(t) = a(t, Y(t))dt + b(t, Y(t))dW(t), \tag{5.20}$$

with $Y(0) = X(0)$ admits a weak solution $Y(t)$ that has the same one-dimensional distribution as $X(t)$.

Proof. The proof is given in Appendix A.15. □

The process $Y(\cdot)$ follows a so called *local volatility* process. The function $b(t,x)$ is often referred to as *Dupire's local volatility*. Since, $X(\cdot)$ and $Y(\cdot)$ have the same one-dimensional distributions, the prices of European options on $X(\cdot)$ and $Y(\cdot)$ for all strikes K and maturities T will be the same. Thus, for the purpose of European option valuation and/or calibration to European options, a very complicated process $X(\cdot)$ can be replaced by a simpler Markov process[2] $Y(\cdot)$, which is called the *Markovian projection of* $X(\cdot)$. From Theorem 5.3.1 we obtain the following useful result:

[2]The stochastic differential equations that are generally considered in financial modeling are already of the Markovian type.

Lemma 5.3.2 (**Connection Processes and Dupire's Local Volatility**). *If two processes have the same Dupire's local volatility, the European option prices on both are identical for all strikes and expiries.*

In fact Theorem 5.3.1 and Lemma 5.3.2 provide us with the means to approximate a given model by essentially any model of choice.

To get more insight in the Markovion projection method we will look at an example.

Example 5.3.3 (**Markovian Projection Application**). *Consider a stochastic volatility model:*

$$dX(t) = b_1(t, X(t))\sqrt{z_1(t)}dW(t), \tag{5.21}$$

where $z_1(t)$ is some variance process. Suppose we would like to match the European option prices on $X(\cdot)$ for all expiries and strikes in a model of the form

$$dY(t) = b_2(t, Y(t))\sqrt{z_2(t)}dW(t), \tag{5.22}$$

where $z_2(t)$ is a different, and potentially simpler, variance process. Then Theorem 5.3.1 and Lemma 5.3.2 imply that $b_2(t, x)$ must be chosen such that

$$\mathbb{E}[\, z_2(t)|\, X_2(t) = x] = \mathbb{E}[\, z_1(t)|\, X_1(t) = x]. \tag{5.23}$$

Rewriting gives us:

$$b^2(t, x) = b^1(t, x)\frac{\mathbb{E}[\, z_1(t)|\, X_1(t) = x]}{\mathbb{E}[\, z_2(t)|\, X_2(t) = x]}. \tag{5.24}$$

The coefficients for the SDE of the Markovian projection are obtained by calculating conditional expected values. From (5.24) it is clear that when applying the Markovian projection method we are limited by the accuracy of the approximations of the conditional expectations. The fact that we calculate the ratio of two expected values, enables us to minimize the error. Note, that even if each individual approximation is inaccurate, they are inaccurate "in the same way" and the overall error diminishes when the ratio is formed. So, in order to maximize the error calculation effect, it is obviously beneficial to choose $z_2(t)$ as close to $z_1(t)$ as possible, while still retaining analytical tractability.

From Example 5.3.3 it is clear that the main difficulty when applying Markovian projection is calculating conditional expectations. Generally, Gaussian approximation is used to obtain these conditional expected values and this is also what we will use.

Lemma 5.3.4 (**Conditioning Formula for Gaussian variables**). *Let X, Y be two normally distributed random variables, $X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma_Y^2)$. Then the conditional formula is given by:*

$$\mathbb{E}[\, Y|\, X = x] = \mathbb{E}[Y] + \frac{\mathrm{Cov}(Y, X)}{\mathrm{Var}(X)}\,(x - \mathbb{E}[X])\,. \tag{5.25}$$

Formula (5.25) can be used as a base for general Gaussian approximation.

5.4 Displaced Diffusion SABR Model

The spread between the CMS-adjusted rates ($\tilde{S_1}(t)$ and $\tilde{S_2}(t)$) is defined by[3]:

$$S(t) = S_1(t) - S_2(t). \tag{5.26}$$

Our aim is to project the 2D SABR model (in the sense of the spread SDE) onto a simpler one-dimensional model. The one-dimensional model we consider is the *displaced diffusion SABR (DD SABR) model*, which is formally defined as follows:

[3]For notational convenience we drop the tildes from hereon out.

Definition 5.4.1 (DD SABR Model). *A displaced diffusion SABR (DD SABR) model is given by the following set of SDEs:*

$$\begin{aligned} dS(t) &= u(t)F(S(t))dW(t), \\ du(t) &= \eta u(t)dZ(t), \\ \langle dW(t), dZ(t)\rangle &= \gamma dt, \\ \text{with } F(S(t)) &= p + q(S(t) - S(0)), \\ p &= F(S(0)), \\ q &= F'(S(0)), \end{aligned} \tag{5.27}$$

where γ denotes the correlation between the forward price and the volatility process.

A displaced model is a reasonable choice, since in case of spread options negative realizations of the spread must have positive probabilities.

The result of projecting the 2D SABR model onto the DD SABR model is presented in Theorem 5.4.2.

Theorem 5.4.2 (Approximated Spread Dynamics). *The dynamics associated with the spread*

$$dS(t) = dS_1(t) - dS_2(t), \tag{5.28}$$

where S_1, S_2 are given by (5.6) can be approximated by the DD SABR model given by (5.27). Here

$$\begin{aligned} p &= \sqrt{p_1^2 + p_2^2 - 2p_1p_2\rho}, \\ q &= \frac{p_1q_1\rho_1^2 - p_2q_2\rho_2^2}{p}, \\ \gamma &= \frac{1}{\eta p^2}\left(p_1^2\nu_1\rho_1\gamma_{11} + p_2^2\nu_2\rho_2\gamma_{22} - p_1p_2\nu_2\rho_2\gamma_{21} - p_1p_2\nu_1\rho_1\gamma_{12}\right), \\ \eta &= \frac{1}{p}\sqrt{(p_1\nu_1\rho_1)^2 + (p_2\nu_2\rho_2)^2 - 2\xi_{12}p_1\nu_1\rho_1p_2\nu_2\rho_2}, \\ p_1 &= \alpha_1(0)S_1(0)^{\beta_1};\;\; p_2 = \alpha_2(0)S_2(0)^{\beta_2}, \\ q_1 &= \alpha_1(0)\beta_1S_1(0)^{\beta_1-1};\;\; q_2 = \alpha_2(0)\beta_2S_2(0)^{\beta_2-1}, \\ \rho_1 &= \frac{p_1 - p_2\rho}{p}; \rho_2 = \frac{p_1\rho - p_2}{p}, \\ S(0) &= S_1(0) - S_2(0), \\ u(0) &= 1. \end{aligned} \tag{5.29}$$

We will thoroughly describe each step that is needed to obtain the approximation. The first step is to ensure the starting values of the process are preserved. From the 2D SABR model given by (5.6) we have [4]:

$$dS_i(t) = \alpha_i(t)S_i(t)^{\beta_i}dW_i(t), \tag{5.30}$$

with $i = 1, 2$. In order to preserve the starting values of the process we have to make use of the following rescaling:

$$u_i(t) = \frac{\alpha_i(t)}{\alpha_i(0)}. \tag{5.31}$$

[4] From hereon out for notational convenience we drop the superscript T_p in dW^{T_p}.

Substituting (5.31) in (5.30) we obtain:

$$dS_i(t) = u_i(t)\varphi(S_i(t))dW_i(t), \tag{5.32}$$

where

$$\varphi(S_i(t)) = \alpha_i(0)S_i(t)^{\beta_i}. \tag{5.33}$$

For the dynamics of the spread we have:

$$\begin{aligned} dS(t) &= dS_1(t) - dS_2(t), \\ &= u_1(t)\varphi(S_1(t))dW_1(t) - u_2(t)\varphi(S_2(t))dW_2(t). \end{aligned} \tag{5.34}$$

Next, we define

$$\sigma^2(t) \triangleq u_1^2(t)\varphi^2(S_1(t)) + u_2^2(t)\varphi^2(S_2(t)) - 2\rho u_1(t)u_2(t)\varphi(S_1(t))\varphi(S_2(t)), \tag{5.35}$$

$$dW(t) \triangleq \frac{1}{\sigma(t)}\left(u_1(t)\varphi(S_1(t))dW_1(t) - u_2(t)\varphi(S_2(t))dW_2(t)\right). \tag{5.36}$$

Then, we can rewrite (5.34) as follows:

$$dS(t) = \sigma(t)dW(t). \tag{5.37}$$

To apply the result of Gyöngy we need to compute the variance of the displaced diffusion SABR model. For notational convenience we define, p_i and q_i as follows:

$$p_i \triangleq \varphi(S_i(0)) = \alpha_i(0)S_i(0)^{\beta_i}, \tag{5.38}$$

$$q_i \triangleq \varphi'(S_i(0)) = \alpha_i(0)\beta_i S_i(0)^{\beta_i-1}. \tag{5.39}$$

Using (5.38) and (5.39) we obtain:

$$u^2(t) = \frac{1}{p^2}\left(p_1^2u_1^2(t) + p_2^2u_2^2(t) - 2\rho p_1 p_2 u_1(t)u_2(t)\right), \tag{5.40}$$

where

$$p = \sigma(0) = \sqrt{p_1^2 + p_2^2 - 2\rho p_1 p_2}. \tag{5.41}$$

The division by $p^2 = \sigma^2(0)$ is necessary to preserve the scaling $u(0) = 1$.

Now, we can apply the result of Gyöngy, Theorem 5.3.1 and Lemma 5.3.2. We set

$$b(t, x) \triangleq \mathbb{E}[\sigma^2(t)|\, S(t) = x]. \tag{5.42}$$

We also have

$$b(t, x) = \mathbb{E}[u^2(t)|\, S(t) = x] \cdot F^2(x). \tag{5.43}$$

Therefore,

$$F^2(x) = \frac{\mathbb{E}[\sigma^2(t)|\, S(t) = x]}{\mathbb{E}[u^2(t)|\, S(t) = x]}. \tag{5.44}$$

Next, we wish to compute the two conditional expectations in (5.44). We notice that $\sigma^2(t)$ and $u(t)$ are linear combinations of the form:

$$f_{ij}(t) = f(S_i(t), S_j(t), u_i(t), u_j(t)), \tag{5.45}$$

$$g_{ij}(t) = g(u_i(t), u_j(t)), \tag{5.46}$$

where

$$f(S_i(t), S_j(t), u_i(t), u_j(t)) = \varphi(S_i(t))\varphi(S_j(t))u_i(t)u_j(t), \tag{5.47}$$
$$g(u_i(t), u_j(t)) = \frac{1}{p^2}p_ip_ju_i(t)u_j(t). \tag{5.48}$$

So, for $\sigma^2(t)$ and $u^2(t)$ we can write:

$$\sigma^2(t) = f_{11}(t) + f_{22}(t) - 2f_{12}(t)\rho, \tag{5.49}$$
$$u^2(t) = g_{11}(t) + g_{22}(t) - 2g_{12}(t)\rho. \tag{5.50}$$

Using a first-order Taylor expansion we can obtain expressions for $f_{ij}(t)$ and $g_{ij}(t)$.

Lemma 5.4.3 (First-Order Taylor Expansions $f_{ij}(t)$ and $g_{ij}(t)$). *The first-order Taylor expansion for $f_{ij}(t)$ and $g_{ij}(t)$ read:*

$$\begin{aligned} f_{ij}(t) \approx\ & \varphi(S_i(0))\varphi(S_j(0)) + (S_i(t) - S_i(0))\varphi'(S_i(0))\varphi(S_j(0)) \\ & + (S_j(t) - S_j(0))\varphi(S_i(0))\varphi'(S_j(0)) + (u_i(t) - 1)\varphi(S_i(0))\varphi(S_j(0)) \\ & + (u_j(t) - 1)\varphi(S_i(0))\varphi(S_j(0)), \end{aligned} \tag{5.51}$$
$$g_{ij}(t) \approx \frac{p_ip_j}{p^2}\left(1 + (u_i(t) - 1) + (u_j(t) - 1)\right). \tag{5.52}$$

Proof. The proof is given in Appendix A.16. □

In order to calculate the expectations in (5.44) we need simple expressions for the conditional expectations $\mathbb{E}[S_i(t) - S_i(0)|\, S(t) = x]$ and $\mathbb{E}[u_i(t) - 1|\, S(t) = x]$. The conditional expected values $\mathbb{E}[S_i(t) - S_i(0)|\, S(t) = x]$ and $\mathbb{E}[u_i(t) - 1|\, S(t) = x]$ can be computed using Gaussian approximation. In particular,

$$\mathbb{E}[S_i(t) - S_i(0)|\, S(t) = x] \approx \mathbb{E}[\bar{S}_i(t) - \bar{S}_i(0)\big|\, \bar{S}(t) = x], \tag{5.53}$$
$$\mathbb{E}[u_i(t) - 1|\, S(t) = x] \approx \mathbb{E}[\bar{u}_i(t) - 1|\, \bar{S}(t) = x]. \tag{5.54}$$

where

$$d\bar{S}(t) = pdW(t), \tag{5.55}$$
$$d\bar{S}_i(t) = p_idW_i(t), \tag{5.56}$$
$$d\bar{u}_i(t) = \nu_idZ_i(t), \tag{5.57}$$
$$d\bar{W}(t) = \frac{1}{p}\left(p_1dW_1(t) - p_2dW_2(t)\right). \tag{5.58}$$

Using these approximations we find the correlation structure of W, W_i and W, Z_i:

$$\langle d\bar{W}(t), dW_i(t)\rangle = \frac{1}{p}\left(p_1\rho_{1i}dt - p_2\rho_{2i}dt\right) = \rho_idt, \tag{5.59}$$
$$\langle d\bar{W}(t), dZ_i(t)\rangle = \frac{1}{p}\left(p_1\gamma_{1i}\rho_3dt - p_2\gamma_{2i}\rho_4dt\right) = \rho_{i+2}dt. \tag{5.60}$$

Finally, applying formula (5.25) to $\mathbb{E}[\bar{S}_i(t) - \bar{S}_i(0)\big|\, \bar{S}(t) = x]$ and $\mathbb{E}[\bar{u}_i(t) - 1|\, \bar{S}(t) = x]$ we can obtain the earlier mentioned simple expressions.

Lemma 5.4.4 (Simple Expressions for Conditional Expectations (5.53) and (5.54)). *Applying Gaussian approximation, formula (5.25), to (5.53) and (5.54); we obtain the following simple expressions:*

$$\mathbb{E}[\,S_i(t) - S_i(0)|\, S(t) = x] \approx \frac{p_i \rho_i}{p}(x - S(0)), \tag{5.61}$$

$$\mathbb{E}[\,u_i(t) - 1|\, S(t) = x] \approx \frac{\nu_i \rho_{i+2}}{p}(x - S(0)). \tag{5.62}$$

Proof. The proof is given in Appendix A.17. □

Now, we can calculate $\mathbb{E}\left[\sigma^2(t)\middle|\, S(t) = x\right]$ and $\mathbb{E}\left[u^2(t)\middle|\, S(t) = x\right]$. The results are given by the following lemma:

Lemma 5.4.5 (Approximations for Conditional Expectations in (5.44)). *The conditional expectations in (5.44) can be approximated as follows:*

$$\mathbb{E}\left[\sigma^2(t)\middle|\, S(t) = x\right] \approx p^2 + (x - S(0))\Theta_1, \tag{5.63}$$

$$\mathbb{E}\left[u^2(t)\middle|\, S(t) = x\right] \approx 1 + (x - S(0))\Theta_2, \tag{5.64}$$

where p is given by (5.41) and

$$\Theta_1 = \frac{2}{p}\left(p_1^2(q_1\rho_1 + \nu_1\rho_3) + p_2^2(q_2\rho_2 + \nu_2\rho_4) - p_1p_2\rho(q_1\rho_1 + q_2\rho_2 + \nu_1\rho_3 + \nu_2\rho_4)\right), \tag{5.65}$$

$$\Theta_2 = \frac{2}{p^3}\left(\nu_1p_1(p_1 - p_2\rho)\rho_3 + \nu_2p_2(p_2 - p_1\rho)\rho_4\right). \tag{5.66}$$

Proof. The proof is given in Appendix A.18. □

This gives us the following results for $F(x)$, given by Lemma 5.4.6.

Lemma 5.4.6 (Results for $F(x)$). *The function $F(x)$ can be approximated by*

$$F(x) \approx \sqrt{\frac{p^2 + (x - S(0))\Theta_1}{1 + (x - S(0))\Theta_2}}. \tag{5.67}$$

Furthermore, for $F(S(0))$ and $F'(S(0))$ we have:

$$F(S(0)) = p, \tag{5.68}$$

$$F'(S(0)) = q, \tag{5.69}$$

where

$$q = \frac{p_1q_1\rho_1^2 - p_2q_2\rho_2^2}{p}. \tag{5.70}$$

Proof. The proof is given in Appendix A.19. □

Finally, we need to derive a SABR-like diffusion for the stochastic volatility. We wish to obtain the coefficients for the SDE:

$$du(t) = \eta u(t)dZ(t). \tag{5.71}$$

To do this we need to apply Itô's lemma (see [44]) to $u(t)$. Remember that for $u(t)$ we have:

$$u^2(t) = \frac{1}{p^2}\left(p_1^2u_1^2(t) + p_2^2u_2^2(t) - 2p_1p_2\rho u_1(t)u_2(t)\right). \tag{5.72}$$

Applying Itô's lemma gives us:

$$
\begin{aligned}
du(t) = {} & \frac{1}{p^2}\left(p_1^2\nu_1\frac{u_1^2}{u^2} - p_1p_2\rho\nu_1\frac{u_1u_2}{u^2}\right)u(t)dZ_1(t) \\
& + \frac{1}{p^2}\left(p_2^2\nu_2\frac{u_2^2}{u^2} - p_1p_2\rho\nu_2\frac{u_1u_2}{u^2}\right)u(t)dZ_2(t).
\end{aligned}
\tag{5.73}
$$

We can replace the quotients

$$
\frac{u_i(t)u_j(t)}{u^2(t)},
\tag{5.74}
$$

by the expected value and set it equal to 1,

$$
\mathbb{E}\left[\frac{u_i^2(t)}{u^2(t)}\right] = \mathbb{E}\left[\frac{u_i(t)u_j(t)}{u^2(t)}\right] = 1.
\tag{5.75}
$$

We get:

$$
\begin{aligned}
du(t) &= \frac{1}{p^2}\left(p_1^2\nu_1\mathbb{E}\left[\frac{u_1^2}{u^2}\right] - p_1p_2\rho\nu_1\mathbb{E}\left[\frac{u_1u_2}{u^2}\right]\right)u(t)dZ_1(t) \\
&\quad + \frac{1}{p^2}\left(p_2^2\nu_2\mathbb{E}\left[\frac{u_2^2}{u^2}\right] - p_1p_2\rho\nu_2\mathbb{E}\left[\frac{u_1u_2}{u^2}\right]\right)u(t)dZ_2(t) \\
&= \frac{1}{p^2}\left(p_1^2\nu_1 - p_1p_2\rho\nu_1\right)u(t)dZ_1(t) + \frac{1}{p^2}\left(p_2^2\nu_2 - p_1p_2\rho\nu_2\right)u(t)dZ_2(t) \\
&= \frac{p_1\nu_1(p_1-p_2\rho)}{p^2}u(t)dZ_1(t) + \frac{p_2\nu_2(p_2-p_1\rho)}{p^2}u(t)dZ_2(t) \\
&= u(t)\left(\frac{p_1\nu_1\rho_1}{p}dZ_1(t) - \frac{p_2\nu_2\rho_2}{p}dZ_2(t)\right).
\end{aligned}
\tag{5.76}
$$

To obtain the SDE (5.71) we have to set:

$$
Z(t) = \frac{1}{\eta p}\left(p_1\nu_1\rho_1 dZ_1 - \rho_2\nu_2\rho_2 dZ_2\right),
\tag{5.77}
$$

$$
\eta = \frac{1}{p}\sqrt{(p_1\nu_1\rho_1)^2 + (p_2\nu_2\rho_2)^2 - 2\xi_{12}p_1\nu_1\rho_1p_2\nu_2\rho_2},
\tag{5.78}
$$

where η is chosen such that $Z(t)$ scales to $\langle Z(t)\rangle = t$. We determine the correlation between the dynamics of the forward price process and the stochastic volatility as:

$$
\begin{aligned}
\gamma &= \frac{\langle dW(t), dZ(t)\rangle}{dt} \\
&= \frac{\langle d\bar{W}(t), dZ(t)\rangle}{dt} \\
&= \frac{1}{\eta p^2}\left(p_1^2\nu_1\rho_1\gamma_{11} + p_2^2\nu_2\rho_2\gamma_{22} - p_1p_2\nu_2\rho_2\gamma_{21} - p_1p_2\nu_1\rho_1\gamma_{12}\right).
\end{aligned}
\tag{5.79}
$$

For more details regarding the projection of the 2D SABR model onto the DD SABR model we refer the reader to [26, pp. 159-171] and [37, pp. 1129-1156].

Remark 5.4.7. *When applying the Markovian projection method we are limited by the accuracy of the approximations of the conditional expectations. In [34] and [37, pp. 1129-1156] Piterbarg claims that the use of first-order Taylor expansions and Gaussian approximations to obtain approximations for the conditional expectations are reasonable. However, quantifying the error of these approximations is not straight-forward and deserves further research.*

So, using the Markovian Projection method we have obtained the DD SABR model, which allows for more efficient simulation. Note that using Hagan's formula the prices of CMS spread options can be obtained even analytically.

The main advantage of the DD SABR model compared to the copula approach is that, unlike in the copula approach, now the full correlation structure is incorporated into the pricing.

5.5 Numerical Experiments

In this section we will perform numerical experiments to gain insight in the performance of the DD SABR model. The 2D SABR model will be used as the reference model. We start by investigating for which time to maturity the DD SABR model can be reasonably used, we do this by pricing a European call spread option for different times to maturity. Next, we will consider the same market data of 2007 and 2013 that was used in Chapter 4 to price a 10Y-2Y CMSSO by the DD SABR model and the reference model. We will compare the results of the DD SABR model with the results of the copula approach in Chapter 4. The chosen TSR model in the copula approach, Lemma 4.5.3, is the swap-yield TSR model, because it is the market standard TSR model. After that, we perform another experiment which gives us insight in the performance of the DD SABR model and the copula approach. For the year 2013 market prices are available which can be used to calibrate the copula approach and the DD SABR model. Finally, we will look at the cross-skew and de-correlation effects that are not present in the copula approach.

5.5.1 Pricing a European Call Spread Option

By the first numerical experiment we will perform related to the DD SABR model we gain insight in up to which maturity time the DD SABR model can accurately be used. In order to do this we will price a European call spread option with both the DD SABR model and the 2D SABR model for different times to maturity. The parameters that are chosen can be found in [26, pp. 169-171].

Chosen parameters[5]: $S_1(0) = 0.030$, $S_2(0) = 0.026$, $\alpha_1(0) = 0.23$, $\alpha_2(0) = 0.20$, $\rho = -0.4$, $\gamma_{11} = -0.2$, $\gamma_{12} = -0.2$, $\gamma_{21} = -0.2$, $\gamma_{22} = -0.3$, $\xi = 0.3$, $\beta_1 = 0.75, \beta_2 = 0.85$, $\nu_1 = 0.20$, $\nu_2 = 0.25$.

Three different maturity times are considered, namely $T = 1$, $T = 5$ and $T = 10$. The results are given in Figure 5.1.

From Figure 5.1 it can be seen that the fit of the DD SABR model is quite good up to five years to maturity. For ten years to maturity the prices calculated by the DD SABR model deviate from the reference prices calculated by the 2D SABR model.

[5] Note that in this case S_1,S_2 are not swap rates

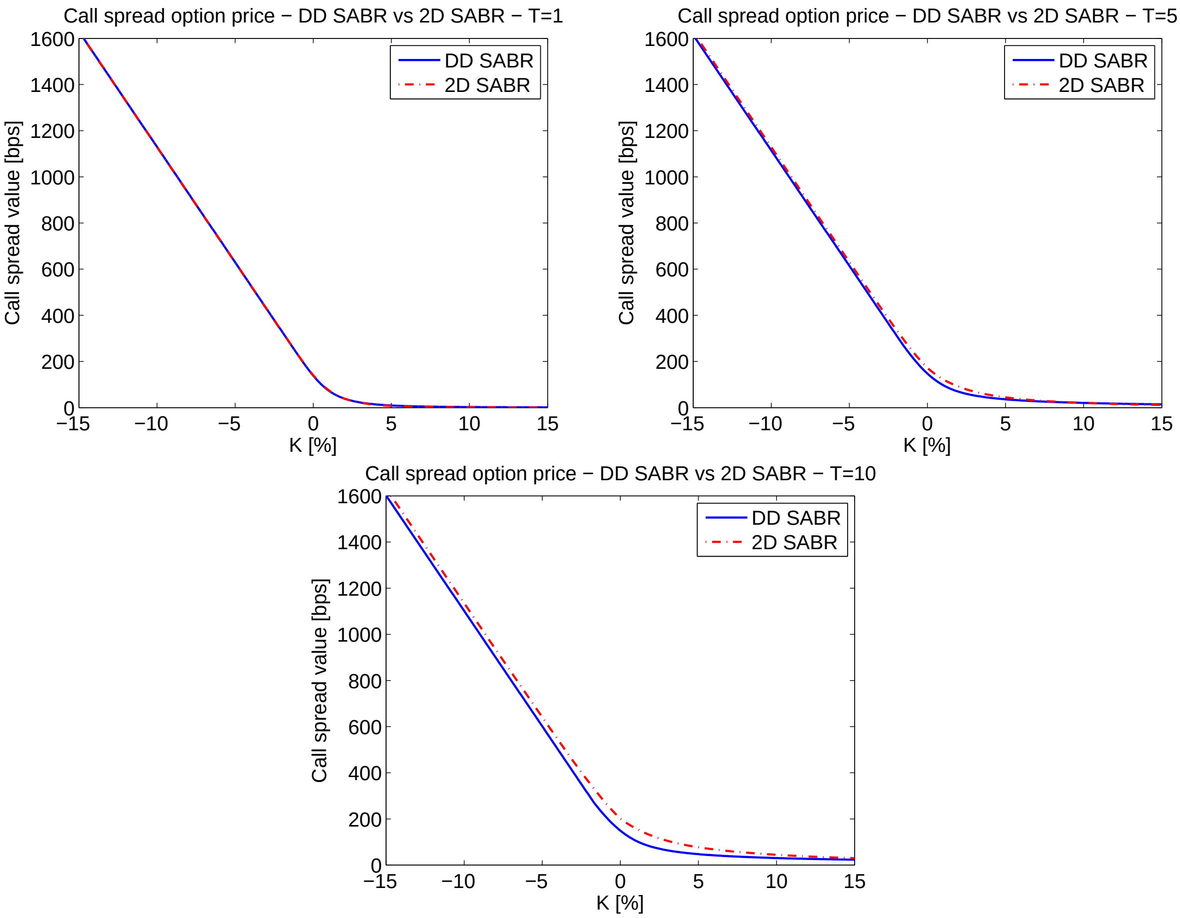


Figure 5.1: European call spread option price using the DD SABR model and the 2D SABR model. Three maturity times are considered: $T = 1$, $T = 5$ and $T = 10$. Number of MC paths is 100000.

5.5.2 DD SABR Model vs Copula Approach - 2007 and 2013

In Chapter 4 we calculated the price of a CMSSO on a 10Y-2Y spread with 12M frequency for market data of 2007 and 2013 using the copula approach. We will now calculate the CMSSO price using the DD SABR model. The 2D SABR model is chosen as the reference model. We have seen that the fit of the DD SABR model is highly satisfactory up to five years to maturity for European call spread options. Therefore, we will consider start dates 1 year and 5 years from today. The difference between the price computed with the DD SABR model (or the copula approach) and the reference model will be denoted by ς, where ς is defined as:

$$\varsigma = V_{\text{CMSSO}}(0) - V_{\text{ref}}.$$

To calculate the CMSSO prices by the 2D and DD SABR model we need to make use of CMS-adjusted forward rates and the associated adjusted SABR paramters as described in Section 5.2. The calibration results regarding the adjusted SABR parameters for 2013 and 2007 are given by Table B.4 and Table B.11, which can be found in Appendix B. The correlation parameter is set to the same value that was used in the copula approach[6], $\rho = 0.8$. The cross-skew parameters, γ_{12} and γ_{21}, and de-correlation parameter, ξ, are set equal to 1 in order to compare the results of the DD SABR model with the copula approach.

[6]In Chapter 4 the correlation parameter was denoted by ϱ.

The computed CMSSO prices and the difference in prices for 2007 and 2013 for start date $T_0 = 1$ are given in Figure 5.2.

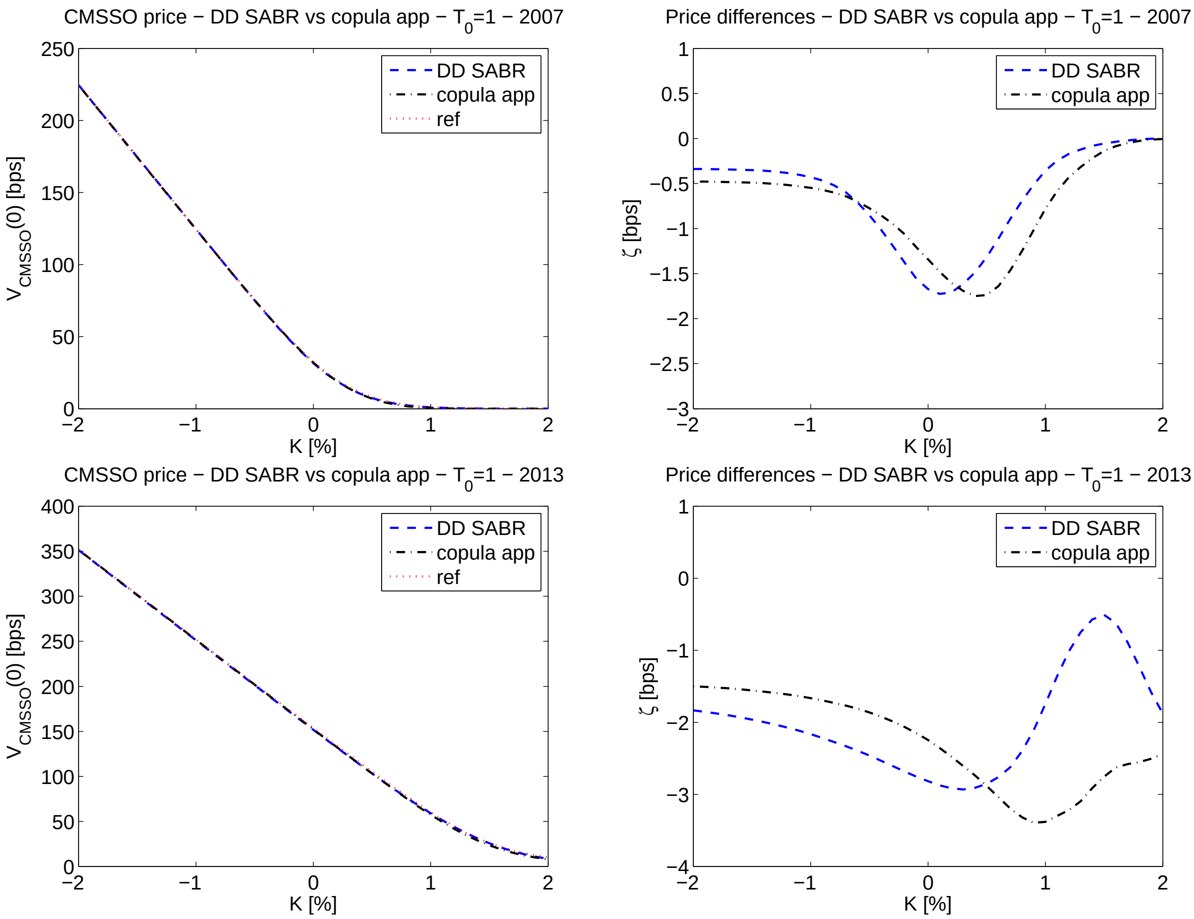


Figure 5.2: Copula approach vs DD SABR model 2007 and 2013. The start date is taken to be 1 year from today. The swap-yield TSR model was used in the copula approach. The reference model is 2D SABR model, number of MC paths is 100000.

From Figure 5.2 it is clear that the fit of both the copula approach and the DD SABR model are good for start date 1 year from today. For 2007 we see that the price differences are smaller than 2bps, while for 2013 they are smaller than 4bps. For the majority of strikes we considered, the prices obtained by the DD SABR model are closer to the reference prices, than the prices obtained by the copula approach. The price differences for 2013 are bigger than for 2007. We already observed this in Chaper 4. As we already mentioned, this is probably due to the fact that the implied volatilities for the year 2013 are more extreme.

Now, we consider a start date 5 years from today. In Figure 5.3 again the CMSSO prices and ς are given for 2007 and 2013.

Figure 5.3 again shows that the results from the DD SABR model are slightly better than the results of the copula approach. Comparing Figure 5.2 with Figure 5.3 we see that the price differences increase as the start date is moved further into the future. We now see that for 2007 the price differences are smaller than 3bps, while for 2013 they are smaller than 8bps. So it seems that both the copula approach and the DD SABR model give accurate approximations of the CMSSO price. Furthermore, also in this example the DD SABR model outperforms the copula approach.

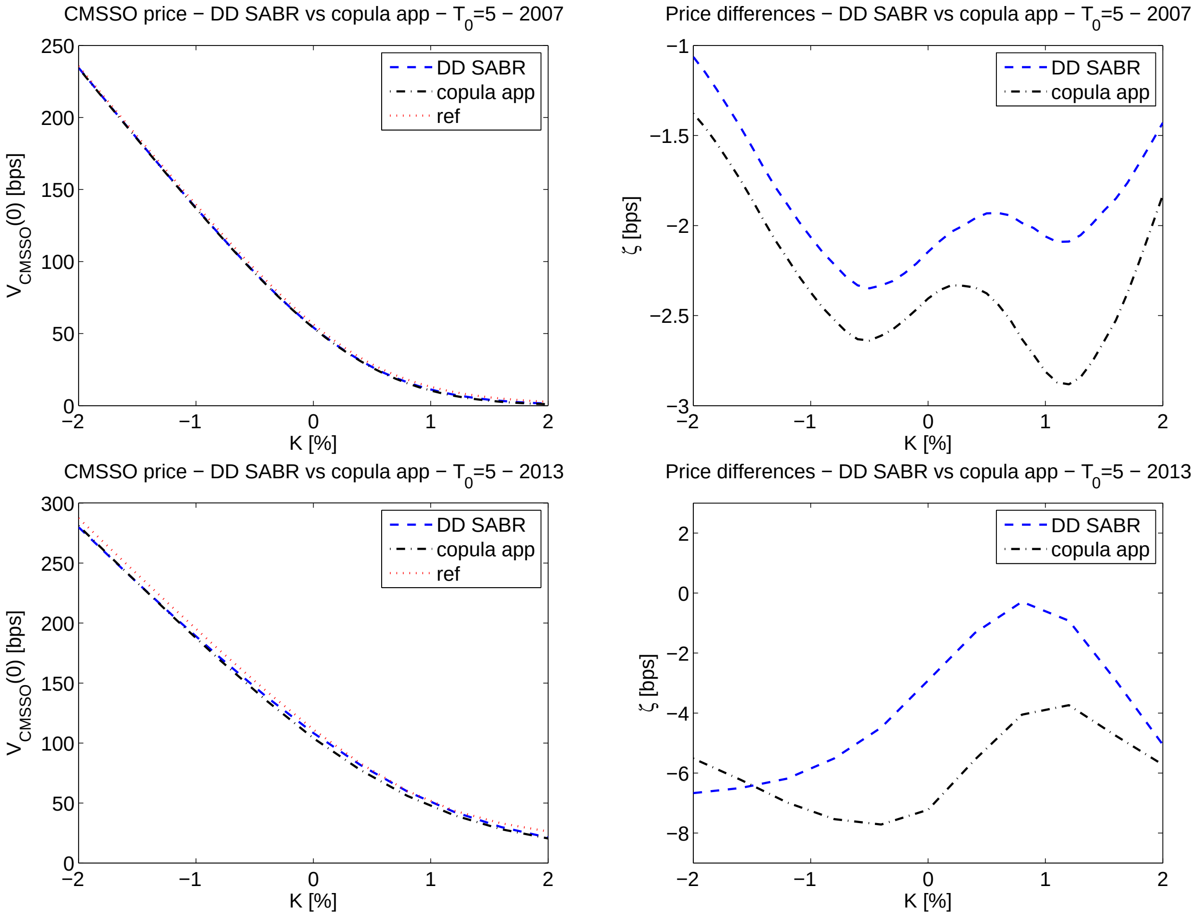


Figure 5.3: Copula approach vs DD SABR model 2007 and 2013. The start date is taken to be 5 years from today. The swap-yield TSR model was used in the copula approach. The reference model is 2D SABR model, number of MC paths is 100000.

5.5.3 Comparing to Market Prices

For the year 2013 we have market prices available for a 10Y-2Y CMSSO with start dates 1 year and 5 years from today[7]. We will compare these market prices to the CMSSO prices calculated with both the DD SABR model and the copula approach. Since we now have market prices available we can calibrate the DD SABR model and the copula approach. Calibrating the copula approach gives us a value for the single correlation parameter ρ. We can then calculate the CMSSO prices by the copula approach using this value for the correlation parameter. We also calibrate the DD SABR model using the market prices. In this case we obtain values for the correlation parameter, the cross-skew parameters and the de-correlation parameter. Remember that for the DD SABR model CMS-adjusted rates and adjusted SABR parameters have to be used. The calibration results regarding the correlation parameters for start date $T_0 = 1$ and $T_0 = 5$ are given by Table B.6 and Table B.7, which can be found in Appendix B. With the obtained parameters we calculate the CMSSO prices. The CMSSO prices, market prices and price differences are given by Table 5.1 and Table 5.2 for their respecting start dates. In order to compare the results of the DD SABR model with the copula approach the sum of squared errors (SSE) is computed for the price differences obtained with both the DD SABR model and

[7]For the year 2007 we do not have this market prices available.

the copula approach. The SSE is defined as follows:

$$\text{SSE} \triangleq \sum_{i=1}^{N} \left(\text{Model Price}(i) - \text{Market Price}(i)\right)^2, \tag{5.80}$$

where N denotes the number of market/model prices.

Strike [%]	DD SABR	Copula	Market	DD SABR diff	Copula Diff
0.25	176.724	177.190	177.124	-0.399	0.065
0	152.158	152.520	152.520	-0.361	0.001
0.25	127.903	128.039	128.191	-0.287	-0.151
0.5	104.184	103.937	104.334	-0.149	-0.396
0.75	81.373	80.645	81.285	0.087	-0.640
1	60.061	59.008	59.830	0.231	-0.821
			SSE	0.456	1.270

Table 5.1: DD SABR model vs copula approach for start date 1 year from today. The CMSSO prices (in bps) calculated with the DD SABR model and the copula approach are compared with the CMSSO market prices. The prices, price differences and SSE for both the DD SABR model and the copula approach are reported.

Strike [%]	DD SABR	Copula	Market	DD SABR diff	Copula Diff
0.25	112.463	113.115	112.387	0.076	0.727
0	92.894	92.823	92.822	0.072	0.001
0.25	75.103	74.887	75.224	-0.121	-0.336
0.5	59.496	59.041	59.598	-0.101	-0.556
0.75	46.360	45.528	46.335	0.025	-0.807
1	35.751	34.631	35.678	0.073	-1.047
			SSE	0.0421	2.703

Table 5.2: DD SABR model vs copula approach for start date 5 years from today. The CMSSO prices (in bps) calculated with the DD SABR model and the copula approach are compared with the CMSSO market prices. The prices, price differences and SSE for both the DD SABR model and the copula approach are reported.

From Table 5.1 and Table 5.2 we see that the SSE is smaller for the DD SABR model than for the copula approach. The CMSSO prices calculated with the DD SABR model are closer to the market prices, than the CMSSO prices calculated with the copula approach. Once again the DD SABR model outperforms the copula approach.

5.5.4 The Cross-Skew and De-Correlation Effect

The results of the previous numerical experiments have shown that the DD SABR model outperforms the copula approach. The main advantage of the DD SABR model in comparison to the copula approach is that the cross-skew and the de-correlation are incorporated into the pricing. The next experiment we perform will show the influence of these parameters on the arbitrage-free prices.

We consider the following base scenario[8]: $S_1(0) = 0.045$, $S_2(0) = 0.032$, $\alpha_1(0) = 0.25$, $\alpha_2(0) = 0.2$, $\rho = 0.9$, $\gamma_{11} = -0.2$, $\gamma_{12} = -0.3$, $\gamma_{21} = -0.3$, $\gamma_{22} = -0.3$ $\xi = 0.75$, $\beta_1 = 0.7$, $\beta_2 = 0.7$, $\nu_1 = 0.4$ and $\nu_2 = 0.4$.

[8]Note that the tildes were dropped from the notation.

First we look into the effect of the cross-skew parameters, γ_{12} and γ_{21}. Two maturity times will be considered, namely $T = 1$ and $T = 5$. Besides the base scenario case we will vary the value for the cross-skew parameters both up and down (e.g. $\gamma_{12} \pm 0.5$). The results for parameter γ_{12} are given in Figure 5.4.

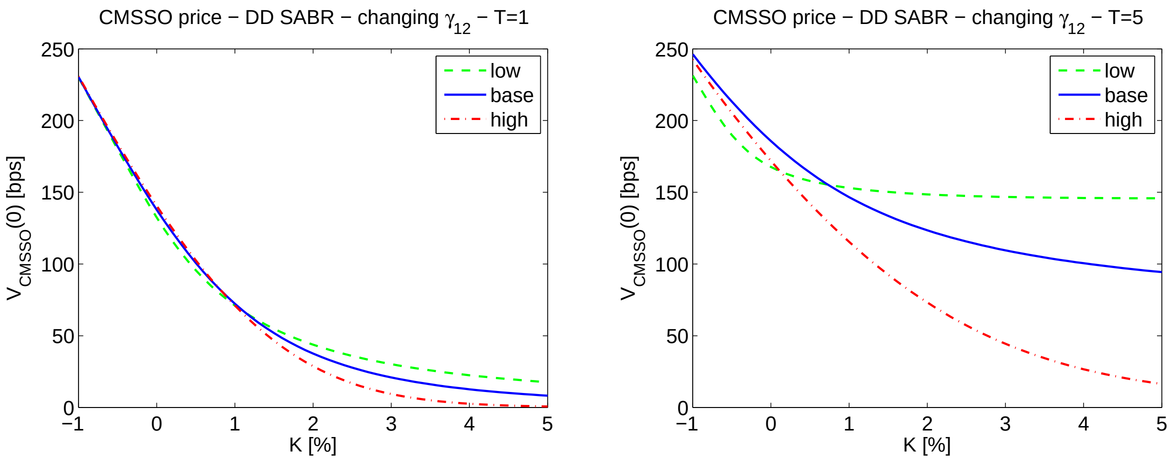


Figure 5.4: Effect of the cross-skew parameter γ_{12} in the DD SABR model. For γ_{12} three different values are chosen. Base: $\gamma_{12} = -0.3$, low: $\gamma_{12} = -0.8$ and high: $\gamma_{12} = 0.2$.

From Figure 5.4 we can see that already for short time to maturity there is a cross-skew effect. For longer time to maturity we see that the effect of the cross-skew parameter γ_{12} is significant, the prices differ much when shifting γ_{12}. In general, when γ_{12} increases the CMSSO prices decrease, and when γ_{12} decreases the CMSSO prices increase. For correlation parameter γ_{21} we follow the same approach as was done for parameter γ_{12}. The results for parameter γ_{21} are given in Figure 5.5. From Figure 5.5 we can see that shifting γ_{21} has the same effect

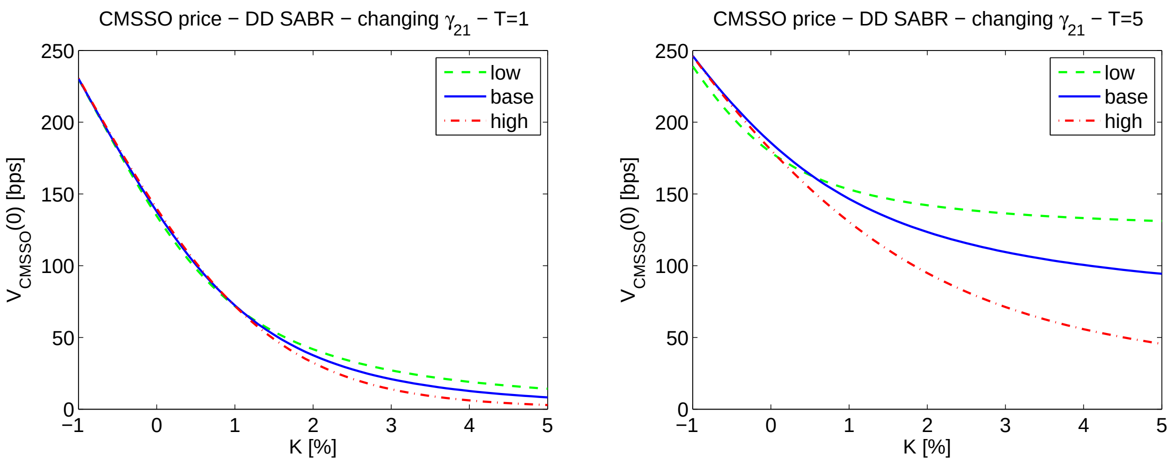


Figure 5.5: Effect of the cross-skew parameter γ_{21} in the DD SABR model. For γ_{21} three different values are chosen. Base: $\gamma_{21} = -0.3$, low: $\gamma_{21} = -0.8$ and high: $\gamma_{21} = 0.2$.

as shifting γ_{12}. The effect that parameter γ_{21} has on the CMSSO price seems weaker when compared with parameter γ_{12}.

Next, we look at the effect of the de-correlation parameter ξ. We again follow the same approach as was done for parameters γ_{12} and γ_{21}. The results when varying ξ are given in Figure 5.6. From Figure 5.6 it is clear that as the time to maturity increases the de-correlation

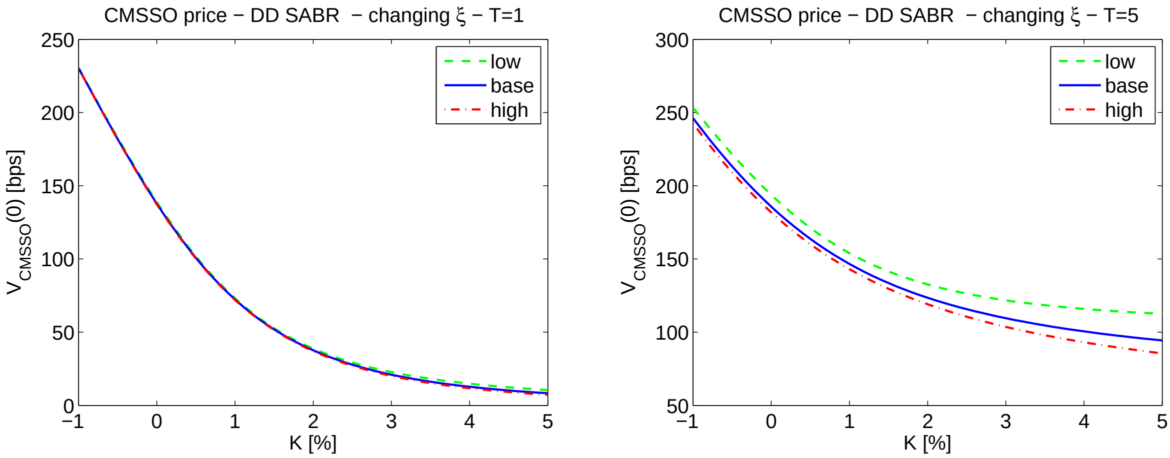


Figure 5.6: Effect of the de-correlation parameter ξ in the DD SABR model. For ξ three different values are chosen. Base: $\xi = 0.75$, low: $\xi = 0.25$ and high: $\xi = 0.99$.

effect becomes more pronounced. For a small time to maturity we hardly see any difference in the prices when shifting the parameter ξ. For longer time to maturity we see a difference. When ξ increases we see that the CMSSO prices decrease, on the other hand when ξ decreases the the CMSSO prices increase. It is also interesting to note that the effect of the de-correlation parameter seems smaller than the effect of the cross-skew parameters. Thus, there is indeed a cross-skew effect and a de-correlation effect.

5.6 Conclusions

In this chapter we showed a technique to obtain and calibrate a relatively simple stochastic volatility model which can be used for pricing CMS spread options, the DD SABR model.

First a modification of the multi-dimensional SABR model presented in [26] was obtained, the 2D SABR model, which can be used for the pricing of CMS spread options. In this model the CMS-adjusted forward rates are assumed to follow SABR dynamics. Most parameters of the 2D SABR model can be obtained by calibrating the SABR model to CMS caplet prices, where the CMS caplet prices are calculated using the replication method and a TSR model, as was described in Chapter 3. The remaining parameters, γ_{12}, γ_{21} and ξ, can be used to calibrate to additional market prices (if available).

Using the Markovian Projection method we obtained from the 2D SABR model the DD SABR model, which allows for more efficient simulation. Note that using Hagan's formula the prices of CMS spread options can be obtained even analytically.

The main advantage of the DD SABR model compared to the copula approach is that the full correlation structure is incorporated into the pricing.

Many numerical experiments were performed to study the performance of the DD SABR model. First European call spread options were priced, which showed that the DD SABR model is accurate up to 5 years to maturity. The results of the copula approach from Chapter 4 were compared to the results of the DD SABR model. From the results of the numerical experiments we can conclude that both the copula approach and the DD SABR model can be used to efficiently and accurately price CMS spread options. We have also seen that the cross-skew parameters and the de-correlation parameter have influence on the CMS spread option price.

We can conclude that generally the DD SABR model outperforms the copula approach.

Chapter 6

Conclusions

This chapter summarizes the main results and conclusions that we have obtained in this thesis.

In this thesis we considered models that can be used for efficient pricing of CMS and CMS spread derivatives. The first part of the research focuses on the pricing of CMS derivatives, while the second part focuses on the pricing of CMS spread derivatives. A CMS derivative is a financial instrument whose payoff is a function of a single swap rate. On the other hand a CMS spread derivative is a financial instrument whose payoff is a function of the spread between two swap rates of different maturity. We are specifically interested in pricing CMS and CMS spread options.

CMS options are widely used by insurance companies and pension funds in their Asset & Liability Management, because these institutions are very vulnerable to movements in the interest rates. CMS spread options are typically used by banks to hedge the CMS spread swaps that they have entered into with customers. The pricing of these products has to be accurate and efficient. The notional values for these kind of derivatives are usually quite large, so even small errors when pricing these products can lead to substantial losses. It is possible to use sophisticated models, for example the Libor market model, to price these products. A downside of these type of models is that they usually are not very efficient, the models have limited analytical tractability. Therefore, the aim of this thesis was to obtain accurate and efficient models that can be used to price CMS and CMS spread options.

For the pricing of CMS options we made use of the TSR approach. Using this approach TSR models can be obtained. We considered two types of TSR models that were described in the literature, the linear TSR model and the swap-yield TSR model. We also developed two new TSR models both based on interpolation, the linear interpolation TSR model and the log-linear interpolation TSR model. To study the performance of the respective TSR models market data from 2007 and 2013 was used. The results for both sets of market data were similar but we observed that the price differences for the year 2013 were larger than for the year 2007, which is probably due to the fact that the volatilities in 2013 were more extreme, indicating that nowadays correct valuation of CMS derivatives is of even more importance. We have also seen that depending on the chosen TSR model the computed price of the CMS option differs slightly. We also showed that there is a timing and a volatility effect. The further the start date is moved into the future, the bigger the differences are between the computed prices of the CMS derivative with the respective TSR models, indicating that there is a timing effect. The volatility effect follows from the fact that the higher the volatility is, the bigger the differences will be between the computed prices of the CMS derivative with the respective TSR models.

From the numerical experiments we have seen that all TSR models have their pros and cons. The swap-yield TSR model is most widely used in the financial industry. Its popularity stems from the fact that only a single assumption is necessary to derive the annuity mapping function. The assumption that is made, is that all underlying swap rates are approximated by a single swap rate. A downside of the swap-yield model is that it is not arbitrage-free. A rescaling has to be used to correctly calculate the price of the CMS option price, which doubles the computation time. The mean reversion linear TSR model is arbitrage-free by definition. Of the four considered TSR models the mean reversion linear TSR model is the only TSR model that incorporates a mean reversion parameter, making it the most flexible TSR model. However, calibrating this mean reversion parameter is not straight-forward and is an issue that should be further researched. The linear and log-linear interpolation TSR models on the other hand require no additional calibration. The linear interpolation TSR model is based on a linear interpolation of the zero-coupon bonds. Another advantage of this model compared to the swap-yield TSR model is that it is arbitrage-free by construction. The log-linear interpolation TSR model is based on a linear interpolation of the logarithm of zero-coupon bonds, which can be a better way to describe the future yield curve movement. For the log-linear interpolation TSR model the same rescaling as for the swap-yield model has to be used. Besides the necessary rescaling, also a numerical root finding algorithm has to be used to calculate the values for the annuity mapping function. Meaning that the log-linear TSR model has the highest computational cost of the considered TSR models, although the computational costs are still very low when compared to sophisticated models.

We would recommend the use of the log-linear interpolation TSR model to price CMS options, depending on the view of the movement of the yield curve. When it is important to reduce the calculation time, we recommend the use of the linear interpolation TSR model. If more flexibility needs to be added we would recommend the use of the mean reversion linear TSR model.

In the second part we made use of the copula approach to efficiently price CMS spread options. We determined the marginal distributions of the two swap rates under the forward measure, the joint distribution was obtained by linking the marginal distributions using a copula function. The Gaussian copula is the copula that we used for the pricing of CMS spread options. Using the copula approach and applying integration by parts we derived a one-dimensional pricing formula that can be used for the pricing of CMS spread options. A copula and a TSR model have to be selected in order to make use of this one-dimensional pricing formula. We also presented a Monte Carlo method which can be used to price CMS spread options, in case a Gaussian copula is assumed.

To study the performance of the copula approach we again made use of market data from 2007 and 2013. In general, the results for both sets of market data were similar. The differences in prices using the one-dimensional pricing formula and the MC method can be explained mostly by numerical issues, i.e. large number of MC paths, chosen values for the integration bounds.

All four TSR models were considered in the copula approach and generally we have seen that the computed CMS spread option price only differs slightly depending on the chosen TSR model. We also showed that there are timing, volatility and correlation effects. The further the start date is moved into the future, the bigger the differences will be between the computed prices of the CMS spread option with the respective TSR models. The volatility effect follows from the fact that the higher the volatility is, the bigger the differences will be between the computed prices of the CMS spread option with the respective TSR models. The correlation parameter has influence on the curvature. The higher the value of the correlation parameter is, the more convex the plot of the computed CMS spread option price becomes. It was shown that the behavior of the two new TSR models, the linear interpolation and log-linear interpolation TSR model, is highly satisfactory.

Finally, we considered a stochastic volatility model for the pricing of CMS spread options, the displaced diffusion SABR (DD SABR) model. A two-dimensional SABR (2D SABR) model was presented that can be used for the pricing of CMS spread options. However, the prices can only be calculated using a MC simulation. Using the Markovian projection method the DD SABR model was derived from the 2D SABR model. To use the 2D and DD SABR models for the pricing of CMS spread options, adjusted CMS-forward rates and the associated adjusted SABR parameters have to be used. The main advantage of the DD SABR model compared to the copula approach is that, unlike in the copula approach, now the full correlation structure is incorporated into the pricing. CMS-adjusted forward rates are defined such that each CMS-adjusted rate follows SABR dynamics. Therefore, most parameters of the DD SABR and 2D SABR models can be obtained by calibrating the SABR model to CMS caplets prices. The remaining parameters, γ_{12}, γ_{21} and ξ, can be used to calibrate to additional market prices (if available).

We can conclude that both the copula approach and the DD SABR model can be used to accurately and efficiently price CMS spread options. The DD SABR model generally outperforms the copula approach.

Chapter 7

Further Research

This chapter describes possible further research directions that could be taken for the efficient pricing of CMS and CMS spread derivatives.

We have seen that CMS derivatives can be priced efficiently using TSR models. Besides the TSR models we considered it might be interesting to look into alternative TSR models. We have seen that the mean reversion linear TSR model is the most flexible TSR model, due to the use of a mean reversion parameter. In this thesis we did not calibrate this parameter. Calibrating this mean reversion parameter is not straight-forward and deserves further research.

It would also be interesting to investigate if the other three TSR models could incorporate a mean reversion parameter. Perhaps alterations could be made to their respective annuity mapping functions. Although we feel that it would be difficult to make such alterations while still satisfying the no-arbitrage, consistency and reasonability conditions.

The two new TSR models we developed are based on interpolation, another possible research direction is to investigate other interpolation techniques which can lead to new TSR models.

For the pricing of CMS spread options we made use of the copula approach. The copula that we chose was the Gaussian copula. Obviously, research could be done into the application of other copulas in CMS spread option pricing. Making use of other copulas in the obtained one-dimensional pricing formula would perhaps lead to slightly more accurate CMS spread option prices. The downside is that these other types of copulas have more parameters which in turn need to be calibrated properly. Additionally, one of the main advantages of the copula approach is the transparency of the obtained model. The more complicated the considered copula is the less transparent the model becomes. In that case it would probably make more sense to use a stochastic volatility model.

We also priced CMS spread options using a stochastic volatility model, the DD SABR model. The DD SABR model was obtained by applying the Markovian projection method to the 2D SABR model. Most parameters of the 2D SABR model can be obtained by calibrating the SABR model to CMS caplet prices. The cross-skew parmaters (γ_{12}, γ_{21}) and de-correlation parameter (ξ) can only be calibrated if additional market prices are available. Therefore, it is worth to look into other possible calibration procedures for the two-dimensional and displaced diffusion SABR models.

When applying the Markovian projection method we are limited by the accuracy of the approximations of the conditional expectations. We followed the same approach as in [34], we made use of first-order Taylor expansions and Gaussian approximations to obtain approximations for the conditional expectations. However, quantifying the error of these approximations was not straight-forward and deserves further research. Additionally, research could be done on other approximation methods such that the accuracy of the approximated conditional expectations improves.

Finally, we would like to mention that the multi-dimensional SABR model and the DD SABR model are not only useful for the pricing of CMS spread options. Further research could be done in order to apply these models to the pricing of e.g. FX Asian options or equity basket options.

Bibliography

[1] A. Antonov and M. Arneguy. Analytical Formulas for Pricing CMS Products in the Libor Market Model with Stochastic Volatility. *Numerix Software Ltd., London*, 2009.

[2] A. Antonov, M. Konikov, and M. Spector. SABR Spreads its Wings. *Risk*, pages 58–63, August 2013.

[3] D. Belomestny, A. Kolodko, and J. Schoenmakers. Pricing CMS Spreads in a Libor Market Model. *International Journal of Theoretical and Applied Finance*, 13(1):45–62, February 2010.

[4] N.H. Bingham and R. Kiesel. Risk-Neutral Valuation - Pricing and Hedging of Financial Derivatives. *Springer Finance*, 1998.

[5] BIS. BIS Quarterly Review, June 2013.

[6] T. Björk. *Arbitrage Theory in Continuous Time.* Oxford University Press, 1998.

[7] F. Black. The Pricing of Commodity Contracts. *Journal of Financial Economics*, 3:167–179, March 1976.

[8] A. Brace, D. Gatarek, and M. Musiela. The Market Model of Interest Rate Dynamics. *Mathematical Finance*, 7(2):127–154, 1997.

[9] D. Brigo and F. Mercurio. *Interest Rate Models - Theory and Practice With Smile, Inflation and Credit.* Springer-Verlag Berline Heidelberg, 2007.

[10] P. Carr and D. Madan. Optimal Positioning in Derivative Securities. *Quantitative Finance*, 1:19–37, 2001.

[11] S. Cederval and V. Piterbarg. Full implications for CMS convexity. *Asia Risk*, 2011.

[12] B. Chen, C.W. Oosterlee, and H. van der Weide. A Low-Bias Simulation Scheme for the SABR Stochastic Volatility Model. *International Journal of Theoretical and Applied Finance*, 15(2), 2012.

[13] M.H.A. Davis and V. Mataix-Pastor. Negative Libor rates in the Swap Market Model. *Finance and Stochastics*, 11:181–193, 2007.

[14] B. Dupire. Pricing with a Smile. *Risk*, 7(1):18–20, 1994.

[15] V. Fleuriet and C. Lubochinsky. Interest Rate Risk Management by Life Insurance Companies and Pension Funds. *Banque de France, Financial Stability Review*, 6:95–111, 2005.

[16] H. Geman, N. El Karoui, and J. Rochet. Changes of Numéraire, Changes of Probability Measure and Option Pricing. *Journal of Applied Probability*, 32:443–458, 1995.

[17] P. Glasserman. *Monte Carlo Methods in Financial Engineering.* Springer, 2004.

[18] I. Gyöngy. Mimicking the One-Dimensional Marginal Distributions of Processes Having an Ito Differential. *Probability Theory and Related Fields*, 71:501–516, 1986.

[19] P.S. Hagan. Convexity Conundrums: Pricing CMS Swaps, Caps, and Floors. *Wilmott Magazine*, 62:38–44, 2003.

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

[21] P.S. Hagan, D. Kumar, A.S. Lesniewski, and D.E. Woodward. Arbitrage-Free SABR. *Wilmott Magazine*, pages 60–75, 2014.

[22] J.C. Hull. *Option Futures and other Derivatives.* Prentice Hall, seventh edition, 2006.

[23] P. Hunt, J. Kennedy, and A. Pelsser. *Financial Derivatives in Theory and Practice.* John Wiley & Sons, Chichester, 2000.

[24] H. Joe. *Multivariate Models and Dependence Concepts.* Chapman & Hall, London, 1997.

[25] M. Joshi and C. Yang. Fast Delta Computations in the Swap-Rate Market Model. *Journal of Economic Dynamics and Control*, 35(5):764–775, May 2011.

[26] J. Kienitz and D. Wetterau. *Financial Modelling Theory, Implementation and Practice with MATLAB Source.* John Wiley & Sons Ltd, 2012.

[27] R. Kiesel and M. Lutz. Efficient Pricing of Constant Maturity Swap Spread Options in a Stochastic Volatility LIBOR Market Model. *The Journal of Computational Finance*, 14(4):37–72, June 2011.

[28] P.E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations.* Springer Heidelberg, first edition, 1992.

[29] W. Margrabe. The Value of an Option to Exchange One Asset for Another. *Journal of Finance*, 33:177–186, 1978.

[30] A. McNeil, R. Frey, and P. Embrechts. *Quantitative Risk Management: Concepts, Techniques and Tools.* Princeton University Press, second edition, 2008.

[31] F. Mercurio and A. Pallavicini. Smiling at Convexity. *Risk*, page 64, August 2006.

[32] R.B. Nelsen. *An Introduction to Copulas.* Springer, New York, second edition, 2006.

[33] A. Pelsser. Mathematical Foundation of Convexity Correction. *Quantitative Finance*, pages 59–65, 2003.

[34] V. Piterbarg. Markovian Projection Method for Volatility Calibration. *Available at SSRN 906473*, 2006.

[35] V. Piterbarg and B. Andersen. *Interest Rate Modeling Volume I: Foundations and Vanilla Models.* Atlantic Financial Press, 2010.

[36] V. Piterbarg and B. Andersen. *Interest Rate Modeling Volume II: Term Structure Models*, chapter 10, pages 407–417. Atlantic Financial Press, 2010.

[37] V. Piterbarg and B. Andersen. *Interest Rate Modeling Volume III: Products and Risk Management.* Atlantic Financial Press, 2010.

[38] R. Rebonato, K. McKay, and R. White. *The SABR/LIBOR Market Model.* Wiley, first edition, 2009.

[39] A. Savine. A Theory of Volatility. *Recent Developments in Mathematical Finance*, pages 151–167, 2002.

[40] E. Schlögl. Arbitrage-Free Interpolation in Models of Market Observable Interest Rates. *Asia Risk*, 2002.

[41] B. Schweizer and A. Sklar. *Probabilistic Metric Spaces.* North-Holland/Elsevier, New York, 1983.

[42] S.E. Shreve. *Stochastic Calculusfor Finance II Continuous-Time Models.* Springer New York, 2004.

[43] A. Sklar. *Fonctions de Repartition à n Dimensions et leurs Marges.* Publications de l'Institut de Statistique de l'Université de Paris 8, 1959.

[44] J.M. Steel. *Stochastic Calculculus and Financial Applications.* Springer, New York, 2001.

[45] L. Tchuindjo. On Valuing Constant Maturity Swap Spread Derivatives. *Mathematical Finance*, 2:189–194, 2012.

[46] G. West. Calibration of the SABR Model in Illiquied Markets. *Applied Mathematical Finance*, 12(4):371–385, 2005.

Delft University of Technology
Faculty of Electrical Engineering, Mathematics and Computer Science
Delft Institute of Applied Mathematics

Appendix

The efficient pricing of CMS and CMS spread derivatives

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

for the degree

MASTER OF SCIENCE
in
APPLIED MATHEMATICS

by

Sebastiaan Borst
Delft, the Netherlands
September 2014

Appendix A

Proofs

In this appendix we list the proofs, which we have omitted in the main text.

A.1 Proof of Theorem 2.2.6

We first notice that for numéraires $M(T)$ and $N(T)$ we have the following two equalities:

$$\mathbb{E}^M\left[\left.\frac{X(T)}{M(T)}\right|\mathcal{F}_t\right] = \frac{X(t)}{M(t)}, \tag{A.1}$$

$$\mathbb{E}^N\left[\left.\frac{X(T)}{N(T)}\right|\mathcal{F}_t\right] = \frac{X(t)}{N(t)}. \tag{A.2}$$

From equation (2.6) it follows that

$$\frac{M(T)N(t)}{N(T)M(t)}\frac{d\mathbb{Q}^N}{d\mathbb{Q}^M} = 1.$$

This means that we can write,

$$\begin{aligned}
\mathbb{E}^M\left[\left.\frac{X(T)}{M(T)}\right|\mathcal{F}_t\right] &= \int_\Omega \frac{X(T)}{M(T)} d\mathbb{Q}^M \\
&= \int_\Omega \frac{X(T)}{M(T)}\left(\frac{M(T)N(t)}{N(T)M(t)}\frac{d\mathbb{Q}^N}{d\mathbb{Q}^M}\right) d\mathbb{Q}^M \\
&= \int_\Omega \frac{X(T)}{N(T)}\frac{N(t)}{M(t)} d\mathbb{Q}^N \\
&= \mathbb{E}^N\left[\left.\frac{X(T)}{N(T)}\frac{N(t)}{M(t)}\right|\mathcal{F}_t\right].
\end{aligned}$$

Equating the expectations above, we obtain the following expression for the Radon-Nikodym derivative:

$$\left.\frac{d\mathbb{Q}^N}{d\mathbb{Q}^M}\right|_{\mathcal{F}_t} = \frac{N(T)M(t)}{N(t)M(T)}. \tag{A.3}$$

For an arbitrary numéraire a probability measure, $\mathbb{Q}^N$, exists, equivalent to the initial measure, $\mathbb{Q}^0$, so that the price of an attainable claim, X, normalized by N, is a martingale under measure

$\mathbb{Q}^N$, and

$$\mathbb{E}^N\left[\left.\frac{X(T)}{N(T)}\right|\mathcal{F}_t\right]=\frac{\mathbb{E}^N\left[\left.\frac{X(T)}{N(T)}\frac{d\mathbb{Q}^N}{d\mathbb{Q}^M}\right|\mathcal{F}_t\right]}{\mathbb{E}^M\left[\left.\frac{d\mathbb{Q}^N}{d\mathbb{Q}^M}\right|\mathcal{F}_t\right]}$$
$$=\frac{\mathbb{E}^M\left[\left.\frac{X(T)}{N(T)}\frac{N(T)M(t)}{M(T)N(t)}\right|\mathcal{F}_t\right]}{\mathbb{E}^M\left[\left.\frac{N(T)M(t)}{M(T)N(t)}\right|\mathcal{F}_t\right]}.$$

We can rewrite this as follows:

$$\mathbb{E}^N\left[\left.\frac{X(T)}{N(T)}\right|\mathcal{F}_t\right]=\frac{\mathbb{E}^M\left[\left.\frac{X(T)}{M(T)}\frac{M(t)}{N(t)}\right|\mathcal{F}_t\right]}{\mathbb{E}^M\left[\left.\frac{N(T)}{M(T)}\frac{M(t)}{N(t)}\right|\mathcal{F}_t\right]}$$
$$=\frac{\mathbb{E}^M\left[\left.\frac{X(T)}{M(T)}\right|\mathcal{F}_t\right]}{\mathbb{E}^M\left[\left.\frac{N(T)}{M(T)}\right|\mathcal{F}_t\right]}.$$

Now, using equalities (A.1) and (A.2) we get

$$\frac{\mathbb{E}^M\left[\left.\frac{X(T)}{M(T)}\right|\mathcal{F}_t\right]}{\mathbb{E}^M\left[\left.\frac{N(T)}{M(T)}\right|\mathcal{F}_t\right]}=\frac{M(t)}{N(t)}\frac{X(t)}{M(t)} \tag{A.4}$$
$$=\frac{X(t)}{N(t)}. \tag{A.5}$$

Equation (A.4) follows from the assumption that a numéraire M exists and the price of any traded asset divided by its associated numéraire is a martingale.

A.2 Proof of Lemma 2.3.4

We can rewrite (2.12) as follows:

$$S_{n,m}(t)=\frac{P(t,T_n)-P(t,T_{n+m})}{A_{n,m}(t)}$$
$$=\frac{\sum_{i=n}^{n+m-1}P(t,T_i)-P(t,T_{i+1})}{A_{n,m}(t)}$$
$$=\frac{\sum_{i=n}^{n+m-1}\tau_iP(t,T_{i+1})\frac{1}{\tau_i}\frac{P(t,T_i)-P(t,T_{i+1})}{P(t,T_{i+1})}}{A_{n,m}(t)}$$
$$=\frac{\sum_{i=n}^{n+m-1}\tau_iP(t,T_{i+1})\frac{1}{\tau_i}\left(\frac{P(t,T_i)}{P(t,T_{i+1})}-1\right)}{A_{n,m}(t)}$$
$$=\frac{\sum_{i=n}^{n+m-1}\tau_iP(t,T_{i+1})\frac{1}{\tau(T_{i+1}-T_i)}\left(\frac{P(t,T_i)}{P(t,T_{i+1})}-1\right)}{A_{n,m}(t)} \tag{A.6}$$
$$=\frac{\sum_{i=n}^{n+m-1}\tau_iP(t,T_{i+1})L_i(t,T_i,T_{i+1})}{A_{n,m}(t)}$$
$$=\frac{\sum_{i=n}^{n+m-1}\tau_iP(t,T_{i+1})L_i(t)}{A_{n,m}(t)}, \tag{A.7}$$

where we made use of expression (2.9) in (A.6).

A.3 Proof of Lemma 2.5.1

We can rewrite (2.27) as follows:

$$\begin{aligned}
V_{\text{swap}}(t) &= B(t) \sum_{n=0}^{N-1} \tau_n \mathbb{E}^B \left[\frac{L_n(T_n) - K}{B(T_{n+1})} \middle| \mathcal{F}_t \right] && \text{(A.8)}\\
&= B(t) \sum_{n=0}^{N-1} \tau_n \mathbb{E}^B \left[\frac{L_n(T_n) - K}{B(T_n)/P(T_n, T_{n+1})} \middle| \mathcal{F}_t \right] && \text{(A.9)}\\
&= B(t) \sum_{n=0}^{N-1} \tau_n \mathbb{E}^B \left[\frac{\left(\frac{1}{\tau_n}\left(\frac{P(T_n,T_n)}{P(T_n,T_{n+1})} - 1\right) - K\right) P(T_n, T_{n+1})}{B(T_n)} \middle| \mathcal{F}_t \right] && \text{(A.10)}\\
&= B(t) \sum_{n=0}^{N-1} \mathbb{E}^B \left[\frac{P(T_n, T_n) - P(T_n, T_{n+1}) - \tau_n K P(T_n, T_{n+1})}{B(T_n)} \middle| \mathcal{F}_t \right] && \text{(A.11)}\\
&= B(t) \sum_{n=0}^{N-1} \frac{P(t, T_n) - P(t, T_{n+1}) - \tau_n K P(t, T_{n+1})}{B(t)} && \text{(A.12)}\\
&= \sum_{n=0}^{N-1} P(t, T_n) - P(t, T_{n+1}) - \tau_n K P(t, T_{n+1}) && \text{(A.13)}\\
&= \sum_{n=0}^{N-1} \tau_n P(t, T_{n+1}) \left(\frac{1}{\tau_n} \left(\frac{P(t, T_n) - P(t, T_{n+1})}{P(t, T_{n+1})} \right) - K \right) && \text{(A.14)}\\
&= \sum_{n=0}^{N-1} \tau_n P(t, T_{n+1})(L_n(t) - K) && \text{(A.15)}\\
&= \left(\sum_{n=0}^{N-1} \tau_n P(t, T_{n+1}) \right) \left(\frac{\sum_{n=0}^{N-1} \tau_n P(t, T_{n+1}) L_n(t)}{\sum_{n=0}^{N-1} \tau_n P(t, T_{n+1})} - K \right) && \text{(A.16)}\\
&= A(t)(S(t) - K), && \text{(A.17)}
\end{aligned}$$

where $A(t)$ is given by (2.11) and $S(t)$ is given by (2.13). The value of the receiver swap follows analogically,

$$V_{\text{swap-rec}}(t) = A(t)(K - S(t)). \tag{A.18}$$

We used the following to get from (A.8) to (A.17): expression (2.16), definition of $L_n(T_n)$, basic calculations, martingale property, basic calculations, definition of $L_n(t)$, basic calculations, definition of $A(t)$ and $S(t)$.

A.4 Proof of Lemma 2.5.2

From (2.39) we know that we can write

$$V_{\text{swaption-pay}}(t) = B(t) \mathbb{E}^B \left[\frac{1}{B(T_n)} A_{n,m}(T_n)(S_{n,m}(T_n) - K)^+ \middle| \mathcal{F}_t \right]. \tag{A.19}$$

Changing to the annuity measure by applying Theorem 2.2.6 we get

$$V_{\text{swaption-pay}}(t) = B(t)\mathbb{E}^A\left[\frac{1}{B(T_n)}A_{n,m}(T_n)(S_{n,m}(T_n)-K)^+\frac{B(T_n)A_{n,m}(t)}{B(t)A_{n,m}(T_n)}\middle|\mathcal{F}_t\right] \tag{A.20}$$

$$= \frac{B(t)}{B(t)}\mathbb{E}^A\left[A_{n,m}(t)(S_{n,m}(T_n)-K)^+\middle|\mathcal{F}_t\right] \tag{A.21}$$

$$= A_{n,m}(t)\mathbb{E}\left[(S_{n,m}(t)-K)^+\middle|\mathcal{F}_t\right]. \tag{A.22}$$

The value of the receiver swaption follows analogically,

$$V_{\text{swaption-rec}}(t) = A_{n,m}(t)\mathbb{E}\left[(K-S_{n,m}(t))^+\middle|\mathcal{F}_t\right]. \tag{A.23}$$

A.5 Proof of Lemma 3.3.1

We write,

$$\mathbb{E}^A\left[f(S(T_0))\middle|\mathcal{F}_0\right] = \int_a^b f(x)\psi^A(x)dx. \tag{A.24}$$

Let us start to analyze Case 3, $a \leq S(0) \leq b$, then we separate the integral and get

$$\mathbb{E}^A\left[f(S(T_0))\middle|\mathcal{F}_0\right] = \int_a^{S(0)} f(x)\psi^A(x)dx + \int_{S(0)}^b f(x)\psi^A(x)dx, \tag{A.25}$$

where $\psi^A(x)$ is given by

$$\psi^A(x) = \begin{cases} \frac{\partial^2 p(0,S(0),T_0,x)}{\partial x^2}, & \text{if } a \leq x < S_0, \\ \frac{\partial^2 c(0,S(0),T_0,x)}{\partial x^2}, & \text{if } S_0 \leq x \leq b. \end{cases} \tag{A.26}$$

Remember the formula for integration by parts, given by

$$\int udv = uv - \int vdu. \tag{A.27}$$

We are going to calculate the integral

$$\int_a^{S(0)} \underbrace{f(x)}_{u}\underbrace{\frac{\partial^2 p(0,S(0),T_0,x)}{\partial x^2}dx}_{dv}, \tag{A.28}$$

the other integral can be calculated analogically. Now let $u = f(x)$ and $dv = \frac{\partial^2 p(0,S(0),T_0,x)}{\partial x^2}dx$, so that

$$\begin{aligned} du &= f'(x)dx, \\ v &= \frac{\partial p(0,S(0),T_0,x)}{\partial x}. \end{aligned}$$

Therefore,

$$\int_a^{S(0)} f(x)\frac{\partial^2 p(0,S(0),T_0,x)}{\partial x^2}dx = f(x)\left.\frac{\partial p(0,S(0),T_0,x)}{\partial x}\right|_a^{S(0)} - \int_a^{S(0)} f'(x)\frac{\partial p(0,S(0),T_0,x)}{\partial x}dx. \tag{A.29}$$

We use integration by parts again to determine the integral

$$\int_a^{S(0)} \underbrace{f'(x)}_{u} \underbrace{\frac{\partial p(0,S(0),T_0,x)}{\partial x} dx}_{dv}.$$

Now let $u = f'(x)$ and $dv = \frac{\partial p(0,S(0),T_0,x)}{\partial x} dx$ so that

$$\begin{aligned} du &= f''(x)dx, \\ v &= p(0,S(0),T_0,x). \end{aligned}$$

Therefore,

$$\int_a^{S(0)} f'(x)\frac{\partial p(0,S(0),T_0,x)}{\partial x} dx = f'(x)p(0,S(0),T_0,x)\big|_a^{S(0)} - \int_a^{S(0)} f''(x)p(0,S(0),T_0,x)dx. \tag{A.30}$$

Substituting (A.30) in (A.29) we get

$$\begin{aligned} \int_a^{S(0)} f(x)\frac{\partial^2 p(0,S(0),T_0,x)}{\partial x^2} dx = f(x) \left.\frac{\partial p(0,S(0),T_0,x)}{\partial x}\right|_a^{S(0)} - f'(x)p(0,S(0),T_0,x)\big|_a^{S(0)} \\ + \int_a^{S(0)} f''(x)p(0,S(0),T_0,x)dx. \end{aligned} \tag{A.31}$$

Analogically for the other integral we get

$$\begin{aligned} \int_{S(0)}^{b} f(x)\frac{\partial^2 c(0,S(0),T_0,x)}{\partial x^2} dx = f(x) \left.\frac{\partial c(0,S(0),T_0,x)}{\partial x}\right|_{S(0)}^{b} - f'(x)c(0,S(0),T_0,x)\big|_{S(0)}^{b} \\ + \int_{S(0)}^{b} f''(x)c(0,S(0),T_0,x)dx. \end{aligned} \tag{A.32}$$

Substituting (A.31) and (A.32) in (A.25) we get the following expression

$$
\begin{aligned}
\mathbb{E}^A\left[f(S(T_0))|\,\mathcal{F}_0\right] = f(x)\left.\frac{\partial p(0,S(0),T_0,x)}{\partial x}\right|_a^{S(0)} &- f'(x)p(0,S(0),T_0,x)\big|_a^{S(0)} \\
&+ \int_a^{S(0)} f''(x)p(0,S(0),T_0,x)dx + f(x)\left.\frac{\partial c(0,S(0),T_0,x)}{\partial x}\right|_{S(0)}^b \\
&- f'(x)c(0,S(0),T_0,x)\big|_{S(0)}^b + \int_{S(0)}^b f''(x)c(0,S(0),T_0,x)dx \\
= f(S(0))\frac{\partial p(0,S(0),T_0,S(0))}{\partial x} &- f(a)\frac{\partial p(0,S(0),T_0,a)}{\partial x} - f'(S(0))p(0,S(0),T_0,S(0)) \\
&+ f'(a)p(0,S(0),T_0,a) + \int_a^{S(0)} f''(x)p(0,S(0),T_0,x)dx + f(b)\frac{\partial c(0,S(0),T_0,b)}{\partial x} \\
&- f(S(0))\frac{\partial c(0,S(0),T_0,S(0))}{\partial x} - f'(b)c(0,S(0),T_0,b) \\
&+ f'(S(0))c(0,S(0),T_0,S(0)) + \int_{S(0)}^b f''(x)c(0,S(0),T_0,x)dx \\
= f(S(0))\left[\frac{\partial p(0,S(0),T_0,S(0))}{\partial x} - \frac{\partial c(0,S(0),T_0,S(0))}{\partial x}\right] &- f(a)\frac{\partial p(0,S(0),T_0,a)}{\partial x} \\
&+ f(b)\frac{\partial c(0,S(0),T_0,b)}{\partial x} - f'(S(0))\Big[p(0,S(0),T_0,S(0)) - c(0,S(0),T_0,S(0))\Big] \\
&+ f'(a)p(0,S(0),T_0,a) - f'(b)c(0,S(0),T_0,b) + \int_a^{S(0)} f''(x)p(0,S(0),T_0,x)dx \\
&+ \int_{S(0)}^b f''(x)c(0,S(0),T_0,x)dx. \qquad \text{(A.33)}
\end{aligned}
$$

Using the following two identities

$$
p(0,S(0),T_0,S(0)) - c(0,S(0),T_0,S(0)) = 0, \qquad \text{(A.34)}
$$

$$
\frac{\partial p(0,S(0),T_0,S(0))}{\partial x} - \frac{\partial c(0,S(0),T_0,S(0))}{\partial x} = 1, \qquad \text{(A.35)}
$$

we can write (A.33) as follows

$$
\begin{aligned}
\mathbb{E}^A\left[f(S(T_0))|\,\mathcal{F}_0\right] = f(S(0)) &- f(a)\frac{\partial p(0,S(0),T_0,a)}{\partial x} + f(b)\frac{\partial c(0,S(0),T_0,b)}{\partial x} \\
&+ f'(a)p(0,S(0),T_0,a) - f'(b)c(0,S(0),T_0,b) \\
&+ \int_a^{S(0)} f''(x)p(0,S(0),T_0,x)dx + \int_{S(0)}^b f''(x)c(0,S(0),T_0,x)dx. \qquad \text{(A.36)}
\end{aligned}
$$

We still need to consider the other two cases.

For Case 1, $S(0) < a$, we have

$$\begin{aligned}
\mathbb{E}^A\left[f(S(T_0))\middle|\,\mathcal{F}_0\right] &= \int_{S(0)}^{b} f(x)\frac{\partial^2 c(0,S(0),T_0,x)}{\partial x^2}dx \\
&= \int_{a}^{b} f(x)\frac{\partial^2 c(0,S(0),T_0,x)}{\partial x^2}dx \\
&= f(x)\left.\frac{\partial c(0,S(0),T_0,x)}{\partial x}\right|_a^b - f'(x)c(0,S(0),T_0,x)\big|_a^b \\
&\quad + \int_a^b f''(x)c(0,S(0),T_0,x)dx \qquad \text{(A.37)} \\
&= f(b)\frac{\partial c(0,S(0),T_0,b)}{\partial x} - f(a)\frac{\partial c(0,S(0),T_0,a)}{\partial x} \\
&\quad - f'(b)c(0,S(0),T_0,b) + f'(a)c(0,S(0),T_0,a) \\
&\quad + \int_a^b f''(x)c(0,S(0),T_0,x)dx. \qquad \text{(A.38)}
\end{aligned}$$

For Case 2, $S(0) > b$, we have

$$\begin{aligned}
\mathbb{E}^A\left[f(S(T_0))\middle|\,\mathcal{F}_0\right] &= \int_{a}^{S(0)} f(x)\frac{\partial^2 p(0,S(0),T_0,x)}{\partial x^2}dx \\
&= \int_{a}^{b} f(x)\frac{\partial^2 p(0,S(0),T_0,x)}{\partial x^2}dx \\
&= f(x)\left.\frac{\partial p(0,S(0),T_0,x)}{\partial x}\right|_a^b - f'(x)p(0,S(0),T_0,x)\big|_a^b \\
&\quad + \int_a^b f''(x)p(0,S(0),T_0,x)dx \qquad \text{(A.39)} \\
&= f(b)\frac{\partial p(0,S(0),T_0,b)}{\partial x} - f(a)\frac{\partial p(0,S(0),T_0,a)}{\partial x} \\
&\quad - f'(b)p(0,S(0),T_0,b) + f'(a)p(0,S(0),T_0,a) \\
&\quad + \int_a^b f''(x)p(0,S(0),T_0,x)dx. \qquad \text{(A.40)}
\end{aligned}$$

A.6 Proof of Lemma 3.5.1

To obtain the linear TSR model the following relation is specified

$$\frac{\pi(s,M)}{\sum_{n=0}^{N-1}\tau_n\pi(s,T_{n+1})} = a(M)s + b(M), \quad M \geq T, \qquad \text{(A.41)}$$

where $a(\cdot)$ and $b(\cdot)$ are deterministic functions. In order for the model to satisfy the *no-arbitrage* condition we see after substituting (A.41) in (3.23) that the following relation must be satisfied

$$P(0,M) = A(0)\mathbb{E}^{\mathbb{A}}\left[a(M)S(T) + b(M)\right]. \qquad \text{(A.42)}$$

Using the fact that under the annuity measure $S(T)$ is a martingale we get the following condition on the free coefficient $b(\cdot)$,

$$b(M) = \frac{P(0,M)}{A(0)} - a(M)S(0). \qquad \text{(A.43)}$$

Now we move on to the *consistency* condition, to satisfy the *consistency* condition we must have

$$\begin{aligned} s &= \frac{1-\pi(s,T_N)}{\sum_{n=0}^{N-1}\tau_n\pi(s,T_{n+1})} \\ &= \frac{1}{\sum_{n=0}^{N-1}\tau_n\pi(s,T_{n+1})} - \frac{\pi(s,T_N)}{\sum_{n=0}^{N-1}\tau_n\pi(x,T_{n+1})} \\ &= \frac{\pi(s,T_0)}{\sum_{n=0}^{N-1}\tau_n\pi(s,T_{n+1})} - \frac{\pi(s,T_N)}{\sum_{n=0}^{N-1}\tau_n\pi(s,T_{n+1})} \\ &= (a(T_0)s+b(T_0)) - (a(T_N)s+b(T_N)). \end{aligned} \tag{A.44}$$

Rewriting expression (A.44) we get,

$$\begin{aligned} 0 &= (a(T_0)s+b(T_0)) - (a(T_N)s+b(T_N)) - s \\ &= (a(T_0)-a(T_N)-1)s + b(T_0) - b(T_N). \end{aligned}$$

The expression above must hold for all s (so also $s=0$) if follows that $b(T_0)=b(T_N)$. Now writing,

$$0 = (a(T_0)-a(T_N)-1)s + b(T_0) - b(T_0) = (a(T_0)-a(T_N)-1)s,$$

it follows that

$$a(T_0) = a(T_N)+1.$$

So we found the following two conditions

$$b(T_0) = b(T_N), \tag{A.45}$$

$$a(T_0) = 1 + a(T_N). \tag{A.46}$$

Also if (A.45) is satisfied then (A.46) is satisfied, this follows from (A.42). We have

$$\begin{aligned} b(T_0) &= b(T_N), \\ \frac{P(0,T_0)}{A(0)} - a(T_0)S(0) &= \frac{P(0,T_N)}{A(0)} - a(T_N)S(0). \end{aligned}$$

Rewriting gives,

$$P(0,T_0) - a(T_0)S(0)A(0) = P(0,T_N) - a(T_N)S(0)A(0).$$

So we have

$$\begin{aligned} a(T_0) &= \frac{P(0,T_0)-P(0,T_N)+a(T_N)S(0)A(0)}{S(0)A(0)} \\ &= \frac{P(0,T_0)-P(0,T_N)}{A(0)}\frac{1}{S(0)} + a(T_N) \\ &= \frac{S(0)}{S(0)} + a(T_N) \\ &= 1 + a(T_N). \end{aligned}$$

Proceeding in a similar fashion it is possible to show from (A.42) that if (A.46) is satisfied, then (A.45) is satisfied.

The specified relationship (A.41) imposes additional restrictions on $a(\cdot), b(\cdot)$. The following must now hold

$$\begin{aligned} 1 &= \sum_{n=0}^{N-1} \tau_n (a(T_{n+1})s + b(T_{n+1})) \\ &= \sum_{n=0}^{N-1} \tau_n a(T_{n+1})s + \sum_{n=0}^{N-1} \tau_n b(T_{n+1}). \end{aligned}$$

Since the equation above must again be valid for all s we must have

$$\sum_{n=0}^{N-1} \tau_n a(T_{n+1}) = 0$$
$$\sum_{n=0}^{N-1} \tau_n b(T_{n+1}) = 1.$$

A.7 Proof of Lemma 3.5.3

We want to connect the coefficients $a(\cdot)$ to mean reversion parameter $\varkappa$. Remember that

$$\frac{\pi(s, M)}{\sum_{n=0}^{N-1} \tau_n \pi(s, T_{n+1})} = a(M)s + b(M), \quad M \geq T, \tag{A.47}$$

and

$$P(T, M) = \pi(S(T), M), \qquad M \geq T. \tag{A.48}$$

Substituting (A.48) in (A.47) and differentiating we get

$$a(M) = \frac{\partial}{\partial S(T)} \frac{P(T, M)}{\sum_{n=0}^{N-1} \tau_n P(t, T_{n+1})}. \tag{A.49}$$

We can rewrite this in the context of a Gaussian one-factor model, as

$$a(M) = \frac{\partial}{\partial x} \frac{P(T, M, x)}{\sum_{n=0}^{N-1} \tau_n P(T, T_{n+1}, x)}\Bigg|_{S(T,x)=S(0)} \times \frac{1}{\frac{\partial}{\partial x} S(T, x)}\Bigg|_{S(T,x)=S(0)}, \tag{A.50}$$

where x is now the short rate state in the Gaussian model on which all discount bonds and swap rates depend. In [36] we have seen that the bond reconstitution formula is given by

$$P(t, T) = \frac{P(0, T)}{P(0, t)} \exp\left(-x(t)G(t, T) - \frac{1}{2} y(t) G^2(t, T)\right), \tag{A.51}$$

where

$$G(t, T) = \int_t^T e^{-\int_t^u \varkappa(s)ds} du.$$

We can easily see that we have

$$G(t, T) = \frac{1 - e^{-\varkappa(T-t)}}{\varkappa}, \tag{A.52}$$

since

$$\begin{aligned}
G(t,T) &= \int_t^T e^{-\int_t^u \varkappa(s)ds} du = \int_t^T e^{-\varkappa(u-t)} du = \left[-\frac{1}{\varkappa} e^{-\varkappa u + \varkappa t}\right]_t^T \\
&= \left(-\frac{1}{\varkappa} e^{-\varkappa T + \varkappa t}\right) - \left(-\frac{1}{\varkappa} e^{-\varkappa t + \varkappa t}\right) = -\frac{1}{\varkappa} e^{-\varkappa(T-t)} + \frac{1}{\varkappa} \\
&= \frac{1 - e^{-\varkappa(T-t)}}{\varkappa}.
\end{aligned}$$

Differentiating $P(t,T)$ we get

$$\begin{aligned}
\frac{\partial}{\partial x} P(t,T) &= \frac{\partial}{\partial x}\left(\frac{P(0,T)}{P(0,t)} \exp\left(-x(t)G(t,T) - \frac{1}{2} y(t) G^2(t,T)\right)\right) \\
&= -G(t,T) \frac{P(0,T)}{P(0,t)} \exp\left(-x(t)G(t,T) - \frac{1}{2} y(t) G^2(t,T)\right) \\
&= -G(t,T)P(t,T).
\end{aligned}$$

Now denote by $A(T,x)$ the annuity as the function of the short rate state x,

$$A(T,x) = \sum_{n=0}^{N-1} \tau_n P(T, T_{n+1}, x).$$

Differentiating we get

$$\begin{aligned}
\frac{\partial}{\partial x} A(T,x) &= \frac{\partial}{\partial x} \sum_{n=0}^{N-1} \tau_n P(T, T_{n+1}, x) = \sum_{n=0}^{N-1} \tau_n \frac{\partial}{\partial x} P(T, T_{n+1}, x) \\
&= -\sum_{n=0}^{N-1} G(T, T_{n+1}) \tau_n P(T, T_{n+1}, x).
\end{aligned}$$

Now we are ready to tackle expression (A.50). Firstly we have

$$\begin{aligned}
\frac{\partial}{\partial x} \frac{P(T,M,x)}{A(T,x)} &= \frac{A(T,x) \frac{\partial}{\partial x} P(T,M,x) - P(T,M,x) \frac{\partial}{\partial x} A(T,x)}{A^2(T,x)} \\
&= \frac{-A(T,x)G(T,M)P(T,M,x) + P(T,M,x) \sum_{n=0}^{N-1} G(T,T_{n+1}) \tau_n P(T,T_{n+1},x)}{A^2(T,x)} \\
&= -\frac{G(T,M)P(T,M,x)}{A(T,x)} + \frac{P(T,M,x) \sum_{n=0}^{N-1} G(T,T_{n+1}) \tau_n P(T,T_{n+1},x)}{A^2(T,x)} \\
&= -\frac{G(T,M)P(T,M,x)}{\sum_{n=0}^{N-1} \tau_n P(T,T_{n+1},x)} + \frac{P(T,M,x) \sum_{n=0}^{N-1} G(T,T_{n+1}) \tau_n P(T,T_{n+1},x)}{\sum_{n=0}^{N-1} \tau_n P(T,T_{n+1},x) \sum_{n=0}^{N-1} \tau_n P(T,T_{n+1},x)}.
\end{aligned}$$

Secondly,

$$\begin{aligned}
\frac{\partial S(T,x)}{\partial x} &= \frac{\partial}{\partial x}\left(\frac{P(T,T_0,x)-P(T,T_N,x)}{A(T,x)}\right)\\
&= \frac{A(T,x)\frac{\partial}{\partial x}(P(T,T_0,x)-P(T,T_N,x)) - (P(T,T_0,x)-P(T,T_N,x))\frac{\partial}{\partial x}A(T,x)}{A^2(T,x)}\\
&= \frac{A(T,x)\frac{\partial}{\partial x}(P(T,T_0,x)-P(T,T_N,x)) - A(T,x)S(T,x)\frac{\partial}{\partial x}A(T,x)}{A^2(T,x)}\\
&= \frac{\frac{\partial}{\partial x}P(T,T_0,x) - \frac{\partial}{\partial x}P(T,T_N,x) - S(T,x)\frac{\partial}{\partial x}A(T,x)}{A(T,x)}\\
&= \frac{G(T,T_N)P(T,T_N,x) - S(T,x)\frac{\partial}{\partial x}A(T,x)}{A(T,x)}\\
&= \frac{G(T,T_N)P(T,T_N,x) + S(T,x)\sum_{n=0}^{N-1}G(T,T_{n+1})\tau_n P(T,T_{n+1},x)}{\sum_{n=0}^{N-1}\tau_n P(T,T_{n+1},x)}.
\end{aligned}$$

Using for all $t \geq T$ the approximation

$$P(T,t,x)|_{S(T,x)=S(0)} \approx \frac{P(0,t)}{P(0,T)}, \tag{A.53}$$

we get

$$\begin{aligned}
\frac{\partial}{\partial x}\frac{P(T,M,x)}{A(T,x)}\bigg|_{S(T,x)=S(0)} &= -\frac{G(T,M)\frac{P(0,M)}{P(0,T)}}{\sum_{n=0}^{N-1}\tau_n\frac{P(0,T_{n+1})}{P(0,T)}} + \frac{\frac{P(0,M)}{P(0,T)}\sum_{n=0}^{N-1}G(T,T_{n+1})\tau_n\frac{P(0,T_{n+1})}{P(0,T)}}{\sum_{n=0}^{N-1}\tau_n\frac{P(0,T_{n+1})}{P(0,T)}\sum_{n=0}^{N-1}\tau_n\frac{P(0,T_{n+1})}{P(0,T)}}\\
&= -\frac{G(T,M)P(0,M)}{\sum_{n=0}^{N-1}\tau_n P(0,T_{n+1})} + \frac{P(0,M)\sum_{n=0}^{N-1}G(T,T_{n+1})\tau_n P(0,T_{n+1})}{\sum_{n=0}^{N-1}\tau_n P(0,T_{n+1})\sum_{n=0}^{N-1}\tau_n P(0,T_{n+1})},
\end{aligned} \tag{A.54}$$

and

$$\begin{aligned}
\frac{\partial}{\partial x}S(T,x)\bigg|_{S(T,x)=S(0)} &= \frac{G(T,T_N)\frac{P(0,T_N)}{P(0,T)} + S(0)\sum_{n=0}^{N-1}G(T,T_{n+1})\tau_n\frac{P(0,T_{n+1})}{P(0,T)}}{\sum_{n=0}^{N-1}\tau_n\frac{P(0,T_{n+1})}{P(0,T)}}\\
&= \frac{G(T,T_N)P(0,T_N) + S(0)\sum_{n=0}^{N-1}G(T,T_{n+1})\tau_n P(0,T_{n+1})}{\sum_{n=0}^{N-1}\tau_n P(0,T_{n+1})}.
\end{aligned}$$

So we have

$$\frac{1}{\frac{\partial}{\partial x}S(T,x)\big|_{S(T,x)=S(0)}} = \frac{\sum_{n=0}^{N-1}\tau_n P(0,T_{n+1})}{G(T,T_N)P(0,T_N) + S(0)\sum_{n=0}^{N-1}G(T,T_{n+1})\tau_n P(0,T_{n+1})}. \tag{A.55}$$

Substituting (A.54) and (A.55) in (A.50) we get

$$
\begin{aligned}
a(M) &= \frac{-G(T,M)P(0,M)}{G(T,T_N)P(0,T_N) + S(0)\sum_{n=0}^{N-1} G(T,T_{n+1})\tau_n P(0,T_{n+1})} \; + \\
&\frac{P(0,M)\sum_{n=0}^{N-1} G(T,T_{n+1})\tau_n P(0,T_{n+1})}{G(T,T_N)P(0,T_N)\sum_{n=0}^{N-1}\tau_n P(0,T_{n+1}) + S(0)\sum_{n=0}^{N-1} G(T,T_{n+1})\tau_n P(0,T_{n+1})\sum_{n=0}^{N-1}\tau_n P(0,T_{n+1})} \\
&= \frac{-G(T,M)P(0,M)\sum_{n=0}^{N-1}\tau_n P(0,T_{n+1})}{G(T,T_N)P(0,T_N)\sum_{n=0}^{N-1}\tau_n P(0,T_{n+1}) + S(0)\sum_{n=0}^{N-1} G(T,T_{n+1})\tau_n P(0,T_{n+1})\sum_{n=0}^{N-1}\tau_n P(0,T_{n+1})} \\
&+ \frac{P(0,M)\sum_{n=0}^{N-1} G(T,T_{n+1})\tau_n P(0,T_{n+1})}{G(T,T_N)P(0,T_N)\sum_{n=0}^{N-1}\tau_n P(0,T_{n+1}) + S(0)\sum_{n=0}^{N-1} G(T,T_{n+1})\tau_n P(0,T_{n+1})\sum_{n=0}^{N-1}\tau_n P(0,T_{n+1})} \\
&= \frac{P(0,M)\sum_{n=0}^{N-1} G(T,T_{n+1})\tau_n P(0,T_{n+1}) - P(0,M)G(T,M)\sum_{n=0}^{N-1}\tau_n P(0,T_{n+1})}{A(0)G(T,T_N)P(0,T_N) + A(0)S(0)\sum_{n=0}^{N-1} G(T,T_{n+1})\tau_n P(0,T_{n+1})} \\
&= \frac{P(0,M)\left(\sum_{n=0}^{N-1} G(T,T_{n+1})\tau_n P(0,T_{n+1}) - G(T,M)A(0)\right)}{A(0)G(T,T_N)P(0,T_N) + A(0)S(0)\sum_{n=0}^{N-1} G(T,T_{n+1})\tau_n P(0,T_{n+1})} \\
&= \frac{P(0,M)\left(\frac{\sum_{n=0}^{N-1} G(T,T_{n+1})\tau_n P(0,T_{n+1})}{A(0)} - G(T,M)\right)}{G(T,T_N)P(0,T_N) + S(0)\sum_{n=0}^{N-1} G(T,T_{n+1})\tau_n P(0,T_{n+1})}.
\end{aligned} \tag{A.56}
$$

We can rewrite (A.56) in the following form

$$
a(M) = \frac{P(0,M)(\gamma - G(T,M))}{P(0,T_N)G(T,T_N) + A(0)S(0)\gamma}, \quad \text{for all} \quad t \geq T,
$$

where

$$
\gamma = \frac{\sum_{n=0}^{N-1}\tau_n P(0,T_{n+1})G(T,T_{n+1})}{A(0)}.
$$

A.8 Proof of Lemma 3.5.4

We can rewrite (3.16) as follows:

$$
\begin{aligned}
V_{\text{gCMS}}(0) &= \frac{A(0)}{P(0,T_p)} \mathbb{E}^A \left[\alpha(S(T_n))S(T_n) \middle| \mathcal{F}_n \right] \\
&= \frac{A(0)}{P(0,T_p)} \mathbb{E}^A \left[(aS(T_n)+b)S(T_n) \middle| \mathcal{F}_n \right] \\
&= \frac{A(0)}{P(0,T_p)} \mathbb{E}^A \left[aS(T_n)^2 + bS(T_n) \middle| \mathcal{F}_n \right] \\
&= \frac{A(0)}{P(0,T_p)} \left(b\mathbb{E}^A \left[S(T_n) \middle| \mathcal{F}_n \right] + a\mathbb{E}^A \left[S(T_n)^2 \middle| \mathcal{F}_n \right] \right) \\
&= \frac{A(0)}{P(0,T_p)} \left(bS(0) + a\mathbb{E}^A \left[S(T_n)^2 \middle| \mathcal{F}_n \right] \right) \\
&= \frac{A(0)}{P(0,T_p)} \left(\left(\frac{P(0,T_p)}{A(0)} - aS(0) \right) S(0) + a\mathbb{E}^A \left[S(T_n)^2 \middle| \mathcal{F}_n \right] \right) \\
&= S(0) + \frac{A(0)}{P(0,T_p)} \left(a\mathbb{E}^A \left[S(T_n)^2 \middle| \mathcal{F}_n \right] - aS(0)^2 \right) \\
&= S(0) + \frac{A(0)}{P(0,T_p)} \left(a\mathbb{E}^A \left[S(T_n)^2 \middle| \mathcal{F}_n \right] - a \left(\mathbb{E}^A \left[S(T_n) \middle| \mathcal{F}_n \right] \right)^2 \right) \\
&= S(0) + \frac{A(0)}{P(0,T_p)} a\text{Var}^A \left(S(T_n) \right). \qquad \text{(A.57)}
\end{aligned}
$$

and

$$
\begin{aligned}
\Lambda_{\text{gCMS}}(0) &= \mathbb{E}^{T_p} \left[S(T_n) \middle| \mathcal{F}_n \right] - S(0) \\
&= V_{\text{gCMS}}(0) - S(0) \\
&= \frac{A(0)}{P(0,T_p)} a\text{Var}^A \left(S(T_n) \right). \qquad \text{(A.58)}
\end{aligned}
$$

We used the following to get (A.57): first two times basic calculations, linearity of expectation and taking out what is known, martingale property, no-arbitrage condition (3.29), basic calculations, martingale property, definition of the variance.

A.9 Proof of Lemma 3.5.6

We want to determine the first and second derivative of $\alpha(s)$. Using the quotient rule we obtain the following two expressions:

$$\frac{d\alpha}{ds} = \frac{z\frac{dy}{ds} - y\frac{dz}{ds}}{z^2}, \qquad \text{(A.59)}$$

and

$$\frac{d^2\alpha}{ds^2} = \frac{z\left(z\frac{d^2y}{ds^2} - y\frac{d^2z}{ds^2}\right) - 2\frac{dz}{ds}\left(z\frac{dy}{ds} - y\frac{dz}{ds}\right)}{z^3}. \qquad \text{(A.60)}$$

Expression (A.60) is valid since,

$$\begin{aligned}
\frac{d^2\alpha}{ds^2} &= \frac{z^2 \frac{d}{ds}\left(z\frac{dy}{ds} - y\frac{dz}{ds}\right) - \left(z\frac{dy}{ds} - y\frac{dz}{ds}\right)\frac{d}{ds}\left(z^2\right)}{\left(z^2\right)^2} \\
&= \frac{z^2\left[\frac{d}{ds}\left(z\frac{dy}{ds}\right) - \frac{d}{ds}\left(y\frac{dz}{ds}\right)\right] - \left(z\frac{dy}{ds} - y\frac{dz}{ds}\right)2z\frac{dz}{ds}}{z^4} \\
&= \frac{z\left[\frac{dz}{ds}\frac{dy}{ds} + z\frac{d^2y}{ds^2} - \left(\frac{dy}{ds}\frac{dz}{ds} + y\frac{d^2z}{ds^2}\right)\right] - 2\frac{dz}{ds}\left(z\frac{dy}{ds} - y\frac{dz}{ds}\right)}{z^3} \\
&= \frac{z\left(z\frac{d^2y}{ds^2} - y\frac{d^2z}{ds^2}\right) - 2\frac{dz}{ds}\left(z\frac{dy}{ds} - y\frac{dz}{ds}\right)}{z^3}.
\end{aligned}$$

The next step is to obtain expressions for $\frac{dy}{ds}$, $\frac{d^2y}{ds^2}$, $\frac{dz}{ds}$ and $\frac{d^2z}{ds^2}$. For $\frac{dy}{ds}$ and $\frac{d^2y}{ds^2}$ we have:

$$\begin{aligned}
\frac{dy}{ds} &= 1, \\
\frac{d^2y}{ds^2} &= 0.
\end{aligned}$$

To get an expression for $\frac{dz}{ds}$ we write:

$$z(s) = 1 - h(s), \tag{A.61}$$

where

$$h(s) = \prod_{i=0}^{N-1} \frac{1}{1+\tau_i s}. \tag{A.62}$$

So we have

$$\begin{aligned}
\frac{dz}{ds} &= \frac{d}{ds}(1 - h(s)) \\
&= -\frac{dh}{ds}.
\end{aligned} \tag{A.63}$$

Now taking the logarithm on both sides of equation (A.62), we can write:

$$\begin{aligned}
\log(h(s)) &= \log\left(\prod_{i=0}^{N-1}\frac{1}{1+\tau_i s}\right) \\
&= \sum_{i=0}^{N-1}\log\left(\frac{1}{1+\tau_i s}\right) \\
&= \sum_{i=0}^{N-1}\log\left((1+\tau_i s)^{-1}\right).
\end{aligned} \tag{A.64}$$

Differentiating equation (A.64) on both sides we get:

$$\begin{aligned}
\frac{1}{h(s)}\frac{dh}{ds} &= \sum_{i=0}^{N-1}\frac{1}{(1+\tau_i s)^{-1}}\cdot -(1+\tau_i s)^{-2}\cdot\tau_i \\
&= \sum_{i=0}^{N-1}(1+\tau_i s)\cdot -(1+\tau_i s)^{-2}\cdot\tau_i \\
&= \sum_{i=0}^{N-1}\frac{-\tau_i}{1+\tau_i s}.
\end{aligned}$$

So we have:

$$\begin{aligned}\frac{dh}{ds} &= h(s)\sum_{i=0}^{N-1}\frac{-\tau_i}{1+\tau_i s}\\ &= \prod_{i=0}^{N-1}\frac{1}{1+\tau_i s}\sum_{i=0}^{N-1}\frac{-\tau_i}{1+\tau_i s}.\end{aligned}$$

Thus,

$$\begin{aligned}\frac{dz}{ds} &= -\prod_{i=0}^{N-1}\frac{1}{1+\tau_i s}\sum_{i=0}^{N-1}\frac{-\tau_i}{1+\tau_i s}\\ &= \prod_{i=0}^{N-1}\frac{1}{1+\tau_i s}\sum_{i=0}^{N-1}\frac{\tau_i}{1+\tau_i s}. \qquad \text{(A.65)}\end{aligned}$$

For the second derivative of z we can write:

$$\begin{aligned}\frac{d^2z}{ds^2} &= \frac{d}{ds}\left(\prod_{i=0}^{N-1}\frac{1}{1+\tau_i s}\sum_{i=0}^{N-1}\frac{\tau_i}{1+\tau_i s}\right)\\ &= \frac{d}{ds}\left(h(s)\sum_{i=0}^{N-1}\frac{\tau_i}{1+\tau_i s}\right)\\ &= \frac{dh}{ds}\sum_{i=0}^{N-1}\frac{\tau_i}{1+\tau_i s} + h(s)\frac{d}{ds}\left(\sum_{i=0}^{N-1}\frac{\tau_i}{1+\tau_i s}\right)\\ &= \frac{dh}{ds}\sum_{i=0}^{N-1}\frac{\tau_i}{1+\tau_i s} + h(s)\frac{d}{ds}\left(\sum_{i=0}^{N-1}\tau_i(1+\tau_i s)^{-1}\right)\\ &= \frac{dh}{ds}\sum_{i=0}^{N-1}\frac{\tau_i}{1+\tau_i s} + h(s)\sum_{i=0}^{N-1}-\tau_i(1+\tau_i s)^{-2}\cdot\tau_i\\ &= \frac{dh}{ds}\sum_{i=0}^{N-1}\frac{\tau_i}{1+\tau_i s} + h(s)\sum_{i=0}^{N-1}-\left(\frac{\tau_i}{1+\tau_i s}\right)^2\\ &= \prod_{i=0}^{N-1}\frac{1}{1+\tau_i s}\sum_{i=0}^{N-1}\frac{-\tau_i}{1+\tau_i s}\sum_{i=0}^{N-1}\frac{\tau_i}{1+\tau_i s} + \prod_{i=0}^{N-1}\frac{1}{1+\tau_i s}\sum_{i=0}^{N-1}-\left(\frac{\tau_i}{1+\tau_i s}\right)^2\\ &= -\prod_{i=0}^{N-1}\frac{1}{1+\tau_i s}\left(\sum_{i=0}^{N-1}\frac{\tau_i}{1+\tau_i s}\right)^2 - \prod_{i=0}^{N-1}\frac{1}{1+\tau_i s}\sum_{i=0}^{N-1}\left(\frac{\tau_i}{1+\tau_i s}\right)^2\\ &= -\left(\prod_{i=0}^{N-1}\frac{1}{1+\tau_i s}\left(\sum_{i=0}^{N-1}\frac{-\tau_i}{1+\tau_i s}\right)^2 + \prod_{i=0}^{N-1}\frac{1}{1+\tau_i s}\sum_{i=0}^{N-1}\left(\frac{\tau_i}{1+\tau_i s}\right)^2\right).\end{aligned}$$

A.10 Proof of Lemma 3.5.9

We want to determine the first and second derivatives of $\alpha(s)$. For the first derivative of $\alpha(s)$ we get

$$\begin{aligned}
\frac{d}{ds}\alpha(s) &= \frac{d}{ds}\left(\frac{1}{\sum_{n=0}^{N-1}\tau_n z(s)^{\vartheta_{n+1}}}\right) \\
&= \frac{-1}{\left(\sum_{n=0}^{N-1}\tau_n z(s)^{\vartheta_{n+1}}\right)^2}\frac{d}{ds}\left(\sum_{n=0}^{N-1}\tau_n z(s)^{\vartheta_{n+1}}\right) \\
&= \frac{-\sum_{n=0}^{N-1}\frac{d}{ds}\left(\tau_n z(s)^{\vartheta_{n+1}}\right)}{\left(\sum_{n=0}^{N-1}\tau_n (z(s))^{\vartheta_{n+1}}\right)^2} \\
&= \frac{-\sum_{n=0}^{N-1}\tau_n\vartheta_{n+1}z(s)^{\vartheta_{n+1}-1}\frac{dz}{ds}}{\left(\sum_{n=0}^{N-1}\tau_n (z(s))^{\vartheta_{n+1}}\right)^2} \\
&= \frac{-\frac{dz}{ds}\sum_{n=0}^{N-1}\tau_n\vartheta_{n+1}z(s)^{\vartheta_{n+1}-1}}{\left(\sum_{n=0}^{N-1}\tau_n (z(s))^{\vartheta_{n+1}}\right)^2}.
\end{aligned} \tag{A.66}$$

Note that we need the first derivative of $z(s)$ with respect to s. Using expression (3.91) we can write:

$$z(s) = 1 - s\sum_{n=0}^{N-1}\tau_n z(s)^{\vartheta_{n+1}}. \tag{A.67}$$

Next we will make use of implicit differentiation and obtain:

$$\frac{dz}{ds} = -\sum_{n=0}^{N-1}\tau_n z(s)^{\vartheta_{n+1}} - s\frac{dz}{ds}\sum_{n=0}^{N-1}\tau_n\vartheta_{n+1}z(s)^{\vartheta_{n+1}-1}.$$

We can rewrite this as,

$$\frac{dz}{ds}\left(1 + s\sum_{n=0}^{N-1}\tau_n\vartheta_{n+1}z(s)^{\vartheta_{n+1}-1}\right) = -\sum_{n=0}^{N-1}\tau_n z(s)^{\vartheta_{n+1}}.$$

Thus we obtain,

$$\frac{dz}{ds} = \frac{-\sum_{n=0}^{N-1}\tau_n z(s)^{\vartheta_{n+1}}}{1 + s\sum_{n=0}^{N-1}\tau_n\vartheta_{n+1}z(s)^{\vartheta_{n+1}-1}}.$$

For notational convenience we define:

$$\alpha(s) = \frac{\Upsilon_1(s)}{\Upsilon_2(s)},$$

where

$$\begin{aligned}
\Upsilon_1(s) &= -\frac{dz}{ds}\sum_{n=0}^{N-1}\tau_n\vartheta_{n+1}z(s)^{\vartheta_{n+1}-1}, \\
\Upsilon_2(s) &= \left(\sum_{n=0}^{N-1}\tau_n (z(s))^{\vartheta_{n+1}}\right)^2.
\end{aligned}$$

For the second derivative of $\alpha(s)$ we then have,

$$\frac{d^2}{ds^2}\alpha(s) = \frac{\Upsilon_2(s)\frac{d\Upsilon_1}{ds} - \Upsilon_1(s)\frac{d\Upsilon_2}{ds}}{(\Upsilon_2(s))^2}. \tag{A.68}$$

For the first derivative of $\Upsilon_1(s)$ with respect to s we have,

$$\begin{aligned}
\frac{d\Upsilon_1}{ds} &= \frac{d}{ds}\left(-\frac{dz}{ds}\sum_{n=0}^{N-1}\tau_n\vartheta_{n+1}z(s)^{\vartheta_{n+1}-1}\right) \\
&= -\frac{d^2z}{ds^2}\sum_{n=0}^{N-1}\tau_n\vartheta_{n+1}z(s)^{\vartheta_{n+1}-1} - \frac{dz}{ds}\frac{d}{ds}\left(\sum_{n=0}^{N-1}\tau_n\vartheta_{n+1}z(s)^{\vartheta_{n+1}-1}\right) \\
&= -\frac{d^2z}{ds^2}\sum_{n=0}^{N-1}\tau_n\vartheta_{n+1}z(s)^{\vartheta_{n+1}-1} \\
&\quad - \frac{dz}{ds}\sum_{n=0}^{N-1}\frac{d}{ds}\left(\tau_n\vartheta_{n+1}z(s)^{\vartheta_{n+1}-1}\right) \\
&= -\frac{d^2z}{ds^2}\sum_{n=0}^{N-1}\tau_n\vartheta_{n+1}z(s)^{\vartheta_{n+1}-1} \\
&\quad - \frac{dz}{ds}\sum_{n=0}^{N-1}\tau_n\vartheta_{n+1}(\vartheta_{n+1}-1)z(s)^{\vartheta_{n+1}-2}\frac{dz}{ds} \\
&= -\frac{d^2z}{ds^2}\sum_{n=0}^{N-1}\tau_n\vartheta_{n+1}z(s)^{\vartheta_{n+1}-1} \\
&\quad - \left(\frac{dz}{ds}\right)^2\sum_{n=0}^{N-1}\tau_n\vartheta_{n+1}(\vartheta_{n+1}-1)z(s)^{\vartheta_{n+1}-2}.
\end{aligned} \tag{A.69}$$

Note that we need the second derivative of $z(s)$ with respect to s. We have,

$$\begin{aligned}
\frac{d^2z}{ds^2} &= \frac{\left(1+s\sum_{n=0}^{N-1}\tau_n\vartheta_{n+1}z(s)^{\vartheta_{n+1}-1}\right)\frac{d}{ds}\left(-\sum_{n=0}^{N-1}\tau_n z(s)^{\vartheta_{n+1}}\right)}{\left(1+s\sum_{n=0}^{N-1}\tau_n\vartheta_{n+1}z(s)^{\vartheta_{n+1}-1}\right)^2} \\
&- \frac{\left(-\sum_{n=0}^{N-1}\tau_n z(s)^{\vartheta_{n+1}}\right)\frac{d}{ds}\left(1+s\sum_{n=0}^{N-1}\tau_n\vartheta_{n+1}z(s)^{\vartheta_{n+1}-1}\right)}{\left(1+s\sum_{n=0}^{N-1}\tau_n\vartheta_{n+1}z(s)^{\vartheta_{n+1}-1}\right)^2} \\
&= \frac{-\frac{dz}{ds}\sum_{n=0}^{N-1}\tau_n\vartheta_{n+1}z(s)^{\vartheta_{n+1}-1}}{1+s\sum_{n=0}^{N-1}\tau_n\vartheta_{n+1}z(s)^{\vartheta_{n+1}-1}} \\
&+ \frac{\left(\sum_{n=0}^{N-1}\tau_n z(s)^{\vartheta_{n+1}}\right)\left(\sum_{n=0}^{N-1}\tau_n\vartheta_{n+1}z(s)^{\vartheta_{n+1}-1} + s\frac{d}{ds}\left(\sum_{n=0}^{N-1}\tau_n\vartheta_{n+1}z(s)^{\vartheta_{n+1}-1}\right)\right)}{\left(1+s\sum_{n=0}^{N-1}\tau_n\vartheta_{n+1}z(s)^{\vartheta_{n+1}-1}\right)^2} \\
&= \frac{-\frac{dz}{ds}\sum_{n=0}^{N-1}\tau_n\vartheta_{n+1}z(s)^{\vartheta_{n+1}-1}}{1+s\sum_{n=0}^{N-1}\tau_n\vartheta_{n+1}z(s)^{\vartheta_{n+1}-1}} \\
&\quad + \frac{\left(\sum_{n=0}^{N-1}\tau_n z(s)^{\vartheta_{n+1}}\right)\left(\sum_{n=0}^{N-1}\tau_n\vartheta_{n+1}z(s)^{\vartheta_{n+1}-1} + s\frac{dz}{ds}\sum_{n=0}^{N-1}\tau_n\vartheta_{n+1}(\vartheta_{n+1}-1)z(s)^{\vartheta_{n+1}-2}\right)}{\left(1+s\sum_{n=0}^{N-1}\tau_n\vartheta_{n+1}z(s)^{\vartheta_{n+1}-1}\right)^2}.
\end{aligned} \tag{A.70}$$

For the first derivative of $\Upsilon_2(s)$ with respect to s we have,

$$\begin{aligned}
\frac{d\Upsilon_2}{ds} &= \frac{d}{ds}\left(\left(\sum_{n=0}^{N-1}\tau_n(z(s))^{\vartheta_{n+1}}\right)^2\right) \\
&= 2\left(\sum_{n=0}^{N-1}\tau_n(z(s))^{\vartheta_{n+1}}\right)\frac{d}{ds}\left(\sum_{n=0}^{N-1}\tau_n(z(s))^{\vartheta_{n+1}}\right) \\
&= 2\left(\sum_{n=0}^{N-1}\tau_n(z(s))^{\vartheta_{n+1}}\right)\left(\frac{dz}{ds}\sum_{n=0}^{N-1}\tau_n\vartheta_{n+1}z(s)^{\vartheta_{n+1}-1}\right). \qquad \text{(A.71)}
\end{aligned}$$

A.11 Proof of Lemma 4.5.2

The proof we present here can be found in [37, pp. 737-738]. The value of the density $\psi^{T_p}(K)$ at point K is equal to the (undiscounted) value of the security with the delta-function payoff, $\delta(S(T)-K)$,

$$\psi^{T_p}(K) = \mathbb{E}^{T_p}\left[\delta(S(T)-K)\right]. \qquad \text{(A.72)}$$

By switching to the annuity measure, using the law of iterated conditional expectation, and the definition (3.27) of $\alpha(s)$, we obtain

$$\begin{aligned}
\psi^{T_p}(K) &= \frac{A(0)}{P(0,T_p)}\alpha(K)\mathbb{E}^A\left[\frac{P(T,T_p)}{A(T)}\delta(S(T)-K)\right] \\
&= \frac{A(0)}{P(0,T_p)}\alpha(K)\mathbb{E}^A\left[\alpha(S(T))\delta(S(T)-K)\right] \\
&= \frac{A(0)}{P(0,T_p)}\alpha(K)\mathbb{E}^A\left[\delta(S(T)-K)\right] \\
&= \frac{A(0)}{P(0,T_p)}\alpha(K)\psi^A(K).
\end{aligned}$$

Expression (4.31), follows trivially.

A.12 Proof of Lemma 4.4.4

The inverse of Ψ, denoted by Ψ^{-1}, exists since Ψ is continuous. Suppose that $X \sim H$, then

$$\begin{aligned}
\mathbb{P}(\Psi(X) \le u) &= \mathbb{P}(\Psi^{-1}(\Psi(X)) \le \Psi^{-1}(u)) \\
&= \mathbb{P}(X \le \Psi^{-1}(u)) \\
&= H(\Psi^{-1}(u)).
\end{aligned}$$

If $H = \Psi$, then

$$H(\Psi^{-1}(u)) = \Psi(\Psi^{-1}(u)) = u.$$

Now suppose that $\Psi(X) \sim U[0,1]$, then

$$\begin{aligned}
u &= \mathbb{P}(\Psi(X) \le u) \\
&= H(\Psi^{-1}(u)).
\end{aligned}$$

Hence, $H = \Psi$.

A.13 Proof of Theorem 4.4.5

A full proof of Sklar's Theorem can be found [41] and [32]. We are going to proof Sklar's Theorem for the two-dimensional case, where the marginal distribution functions are assumed to be continuous.

Let $\mathbf{X} = (X_1, X_2)'$ be a random vector with joint distribution function Ψ and continuous marginal distribution functions Ψ_1, Ψ_2. Since the marginal distribution functions are continuous, we have:

$$\mathbb{P}(X_1 \leq x_1, X_2 \leq x_2) = \mathbb{P}(\Psi_1(X_1) \leq \Psi_1(x_1), \Psi_2(X_2) \leq \Psi_2(x_2)). \tag{A.73}$$

From Lemma 4.4.4 we know that for $i = 1, 2$:

$$\Psi_i(X_i) \sim U[0, 1]. \tag{A.74}$$

The existence follows directly from the definition of the copula, Definition 4.4.3, since $(\Psi_1(X_1)\Psi_2(X_2))'$ has a copula C as its joint distribution function.

What is left to prove is the uniqueness. Let $x_i = \Psi_i^{-1}(u_i)$ for all $i = 1, 2$, then by continuity of the marginal distribution functions we have,

$$C(u_1, u_2) = \Psi(\Psi_1^{-1}(u_1), \Psi_2^{-1}(u_2)). \tag{A.75}$$

Since (A.75) is an explicit expression, C must be unique.

Conversely, let C be a copula and Ψ_1, Ψ_2 be continuous univariate distribution functions. Suppose the random vector $\mathbf{U} = (U_1, U_2)'$ has joint distribution function C. Let $\mathbf{X} = (\Psi_1^{-1}(U_1), \Psi_2^{-1}(U_2))'$ then by Lemma 4.4.4 we have,

$$\begin{aligned}\mathbb{P}(X_1 \leq x_1, X_2 \leq x_2) &= \mathbb{P}(\Psi_1^{-1}(U_1) \leq x_1, \Psi_2^{-1}(U_2) \leq x_2) \\ &= \mathbb{P}(U_1 \leq \Psi_1(x_1), U_2 \leq \Psi_2(x_2)) \\ &= C(\Psi_1(x_1), \Psi_2(x_2)).\end{aligned} \tag{A.76}$$

A.14 Proof of Lemma 4.4.6

The proof of the multi-dimensional case can be found in [37, pp. 771-772]. We are going to proof the two-dimensional case. From Sklar's Theorem, Theorem 4.4.5, we have

$$\Psi_C(x_1, x_2) = C(\Psi_1(x_1), \Psi_2(x_2)). \tag{A.77}$$

The PDF is the derivative of the CDF, so differentiating the copula function gives us

$$\begin{aligned}\psi_C(x_1, x_2) &= \frac{\partial^2}{\partial x_1 \partial x_2}\Psi_C(x_1, x_2) \\ &= \frac{\partial^2}{\partial x_1 \partial x_2}C(\Psi_1(x_1), \Psi_2(x_2)) \\ &= \frac{\partial^2 C(\Psi_1(x_1), \Psi_2(x_2))}{\partial \Psi_1(x_1)\partial \Psi_2(x_2)}\frac{\partial \Psi_1(x_1)}{\partial x_1}\frac{\partial \Psi_2(x_2)}{\partial x_2} \\ &= \frac{\partial^2 C(\Psi_1(x_1), \Psi_2(x_2))}{\partial \Psi_1(x_1)\partial \Psi_2(x_2)}\psi_1(x_1)\psi_2(x_2) \\ &= c(\Psi_1(x_1), \Psi_2(x_2))\psi_1(x_1)\psi_2(x_2),\end{aligned}$$

where

$$c(u_1, u_2) = \frac{\partial^2}{\partial u_1 \partial u_2}C(u_1, u_2).$$

A.15 Proof of Theorem 5.3.1

Piterbarg [34] mentions that the original proof in [18] is fairly involved. Instead he follows a different approach inspired by Savine (see [39]), which is a much more financially-motivated approach. This is also the approach we will follow.

Before we start with the proof we first have to address the problem of applying Itô's lemma to non-differentiable functions. Itô established a link between stochastic processes and differentiable solutions. For any stochastic process

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

and any twice-continuously differentiable function $f(t, X(t))$, we have

$$df(t, X(t)) = \frac{\partial}{\partial t} f(t, X(t))dt + \frac{\partial}{\partial X(t)} f(t, X(t))dX(t) + \frac{1}{2}\frac{\partial^2}{\partial X^2(t)} f(t, X(t))(dX(t))^2. \tag{A.78}$$

For call option prices we have the following expression

$$C(t, K) = \mathbb{E}[\max\{X(t) - K, 0\}]. \tag{A.79}$$

Now let $f(t, X(t)) = \max\{X(t) - K, 0\}$. Note, that in this case we cannot apply Itô's lemma. In [39] a new type of differentiation is described for these type of functions. The functions are differentiable in the sense of distributions. Using the results from [39] we obtain the following partial derivatives for $f(t, X(t)) = \max\{X(t) - K, 0\}$:

$$\frac{\partial}{\partial t} f(t, X(t)) = 0, \tag{A.80}$$

$$\frac{\partial}{\partial X} f(t, X(t)) = \mathbb{1}_{\{X(t)>K\}}, \tag{A.81}$$

$$\frac{\partial^2}{\partial X^2} f(t, X(t)) = \delta(X(t) - K). \tag{A.82}$$

We can apply the Itô-Tanaka formula to the function $f(t, X(t)) = \max\{X(t)-K, 0\}$ and obtain:

$$\max\{X(T) - K, 0\} = \max\{X(0) - K, 0\} + \int_0^T \mathbb{1}_{\{X(t)>K\}} dX(t) + \frac{1}{2}\int_0^t \delta(X(t) - K)(dX(t))^2. \tag{A.83}$$

Now, we are ready to start with the proof. In general we are mostly concerned about the volatility function, so we set $\mu(t) = 0$. The dynamic for the stochastic process $X(t)$ are now given by

$$dX(t) = \sigma(t)dW(t). \tag{A.84}$$

From Dupire's theorem, [14] we already now that in the local volatility model,

$$dY(t) = b(t, Y(t)dW(t)), \tag{A.85}$$

for a given call option $C(t, K)$ the local volatility $b(t, Y(t))$ can be expressed as:

$$b^2(t, K) = \frac{\frac{\partial}{\partial t} C(t, K)}{\frac{1}{2}\frac{\partial^2}{\partial K^2} C(t, K)}. \tag{A.86}$$

The dynamics of $f(t, X(t)) = \max\{X(t) - K, 0\}$ are derived from the Itô-Tanaka formula:

$$\begin{aligned} df(t, X(t)) &= \frac{\partial}{\partial t} f(t, X(t)) + \frac{\partial}{\partial X} f(t, X(t))dX(t) + \frac{1}{2}\frac{\partial^2}{\partial X^2} f(t, X(t))(dX(t))^2 \\ &= \mathbb{1}_{\{X(t)>K\}} dX(t) + \frac{1}{2}\delta(X(t) - K)(dX(t))^2 \\ &= \mathbb{1}_{\{X(t)>K\}} \sigma(t)dW(t) + \frac{1}{2}\delta(X(t) - K)\sigma^2(t)dt. \end{aligned} \tag{A.87}$$

On the other hand the expectation of the payoff is given by:

$$\begin{aligned}
\mathbb{E}[\max\{X(T)-K,0\}] &= \max\{X(0)-K,0\} + \mathbb{E}\left[\int_0^T \mathbb{1}_{\{X(t)>K\}} dX(t)\right] + \frac{1}{2}\mathbb{E}\left[\int_0^t \delta(X(t)-K)(dX(t))^2\right] \\
&= \max\{X(0)-K,0\} + \frac{1}{2}\int_0^t \mathbb{E}\left[\delta(X(t)-K)(dX(t))^2\right] \\
&= \max\{X(0)-K,0\} + \frac{1}{2}\int_0^t \mathbb{E}\left[\delta(X(t)-K)\sigma^2(t)\right] dt. \qquad \text{(A.88)}
\end{aligned}$$

We can rewrite $\mathbb{E}\left[\delta(X(t)-K)\sigma^2(t)\right]$ as follows:

$$\begin{aligned}
\mathbb{E}\left[\delta(X(t)-K)\sigma^2(t)\right] &= \mathbb{E}\left[\mathbb{E}\left[\delta(X(t)-K)\sigma^2(t)\middle|\, X(t)=K\right]\right] \\
&= \mathbb{E}\left[\delta(X(t)-K)\right]\mathbb{E}\left[\sigma^2(t)\middle|\, X(t)=K\right]. \qquad \text{(A.89)}
\end{aligned}$$

Making use of (A.81) and (A.82) we get:

$$\delta(X(t)-K) = \frac{\partial^2}{\partial K^2}\max\{X(t)-K,0\}. \qquad \text{(A.90)}$$

Substituting (A.90) in (A.89) we obtain:

$$\begin{aligned}
\mathbb{E}\left[\delta(X(t)-K)\sigma^2(t)\right] &= \mathbb{E}\left[\frac{\partial^2}{\partial K^2}\max\{X(t)-K,0\}\right]\mathbb{E}\left[\sigma^2(t)\middle|\, X(t)=K\right] \\
&= \frac{\partial^2}{\partial K^2}\mathbb{E}\left[\max\{X(t)-K,0\}\right]\mathbb{E}\left[\sigma^2(t)\middle|\, X(t)=K\right] \\
&= \frac{\partial^2}{\partial K^2}C(t,K)\mathbb{E}\left[\sigma^2(t)\middle|\, X(t)=K\right]. \qquad \text{(A.91)}
\end{aligned}$$

For the partial derivative of the call price with respect to the time of maturity we have:

$$\begin{aligned}
\frac{\partial}{\partial T}C(T,K) &= \frac{\partial}{\partial T}\left(\max\{X(0)-K,0\} + \frac{1}{2}\int_0^t \mathbb{E}\left[\delta(X(t)-K)\sigma^2(t)\right] dt\right) \\
&= \frac{1}{2}\mathbb{E}\left[\delta(X(t)-K)\sigma^2(t)\right] \\
&= \frac{1}{2}\frac{\partial^2}{\partial K^2}C(T,K)\mathbb{E}\left[\sigma^2(t)\middle|\, X(T)=K\right]. \qquad \text{(A.92)}
\end{aligned}$$

Therefore,

$$\mathbb{E}\left[\sigma^2(t)\middle|\, X(t)=K\right] = \frac{\frac{\partial}{\partial t}C(t,K)}{\frac{1}{2}\frac{\partial^2}{\partial K^2}C(t,K)}. \qquad \text{(A.93)}$$

Since matching European option prices for all strikes and maturities is equivalent to matching all one-dimensional distributions, [14], we have:

$$b^2(t,K) = \mathbb{E}\left[\sigma^2(t)\middle|\, X(t)=K\right]. \qquad \text{(A.94)}$$

A.16 Proof of Lemma 5.4.3

Using a first-order Taylor expansion for $f_{ij}(t) = f(S_i(t), S_j(t), u_i(t), u_j(t))$, we obtain:

$$\begin{aligned}
f_{ij}(t) &\approx f(S_i(0), S_j(0), u_i(0), u_j(0)) + (S_i(t) - S_i(0))\frac{\partial f}{\partial S_i}(S_i(0), S_j(0), u_i(0), u_j(0)) \\
&\quad + (S_j(t) - S_j(0))\frac{\partial f}{\partial S_j}(S_i(0), S_j(0), u_i(0), u_j(0)) \\
&\quad + (u_i(t) - u_i(0))\frac{\partial f}{\partial u_i}(S_i(0), S_j(0), u_i(0), u_j(0)) \\
&\quad + (u_j(t) - u_j(0))\frac{\partial g}{\partial u_j}(S_i(0), S_j(0), u_i(0), u_j(0)) \\
&= \varphi(S_i(0))\varphi(S_j(0))u_i(0)u_j(0) + (S_i(t) - S_i(0))\varphi'(S_i(0))\varphi(S_j(0))u_i(0)u_j(0) \\
&\quad + (S_j(t) - S_j(0))\varphi(S_i(0))\varphi'(S_j(0))u_i(0)u_j(0) \\
&\quad + (u_i(t) - u_i(0))\varphi(S_i(0))\varphi(S_j(0))u_i'(0)u_j(0) \\
&\quad + (u_j(t) - u_j(0))\varphi(S_i(0))\varphi(S_j(0))u_i(0)u_j'(0) \\
&= \varphi(S_i(0))\varphi(S_j(0)) + (S_i(t) - S_i(0))\varphi'(S_i(0))\varphi(S_j(0)) \\
&\quad + (S_j(t) - S_j(0))\varphi(S_i(0))\varphi'(S_j(0)) + (u_i(t) - 1)\varphi(S_i(0))\varphi(S_j(0)) \\
&\quad + (u_j(t) - 1)\varphi(S_i(0))\varphi(S_j(0)).
\end{aligned} \tag{A.95}$$

Now, using (5.38) and (5.39) we can write,

$$\begin{aligned}
f_{ij}(t) &\approx p_ip_j + (S_i(t) - S_i(0))q_ip_j + (S_j(t) - S_j(0))p_iq_j \\
&\quad + (u_i(t) - 1)p_ip_j + (u_j(t) - 1)p_ip_j \\
&= p_ip_j\left(1 + \frac{q_i}{p_i}(S_i(t) - S_i(0)) + \frac{q_j}{p_j}(S_j(t) - S_j(0)) + (u_i(t) - 1) + (u_j(t) - 1)\right).
\end{aligned} \tag{A.96}$$

Similarly, we obtain for $g_{ij}(t) = g(u_i(t), u_j(t))$:

$$\begin{aligned}
g_{ij}(t) &\approx g(u_i(0), u_j(0)) + (u_i(t) - u_i(0))\frac{\partial g}{\partial u_i}(u_i(0), u_j(0)) + (u_j(t) - u_j(0))\frac{\partial g}{\partial u_j}(u_i(0), u_j(0)) \\
&= \frac{1}{p^2}p_ip_ju_i(0)u_j(0) + (u_i(t) - u_i(0))\frac{1}{p^2}p_ip_ju_i'(0)u_j(0) + (u_j(t) - u_j(0))\frac{1}{p^2}p_ip_ju_i(0)u_j'(0) \\
&= \frac{1}{p^2}p_ip_j + (u_i(t) - 1)\frac{1}{p^2}p_ip_j + (u_j(t) - 1)\frac{1}{p^2}p_ip_j \\
&= \frac{p_ip_j}{p^2}\left(1 + (u_i(t) - 1) + (u_j(t) - 1)\right).
\end{aligned} \tag{A.97}$$

A.17 Proof of Lemma 5.4.4

Applying (5.25) to $\mathbb{E}[\bar{S}_i(t) - \bar{S}_i(0) \,|\, \bar{S}(t) = x]$, we obtain:

$$\mathbb{E}[\bar{S}_i(t) - \bar{S}_i(0) \,|\, \bar{S}(t) = x] = \mathbb{E}[\bar{S}_i(t) - \bar{S}_i(0)] + \frac{\mathrm{Cov}(\bar{S}_i(t) - \bar{S}_i(0), \bar{S}(t))}{\mathrm{Var}(\bar{S}(t))}(x - \mathbb{E}[\bar{S}(t)]). \tag{A.98}$$

To get an expression for $\mathbb{E}[\bar{S}_i(t) - \bar{S}_i(0) \,|\, \bar{S}(t) = x]$ we have to calculate the respective terms. First, we calculate $\mathbb{E}[\bar{S}_i(t) - \bar{S}_i(0)]$:

$$\mathbb{E}[\bar{S}_i(t) - \bar{S}_i(0)] = \mathbb{E}\left[\bar{S}_i(0) + \int_0^t p_i dW_i(s) - \bar{S}_i(0)\right] = \mathbb{E}\left[\int_0^t p_i dW_i(s)\right] = 0. \tag{A.99}$$

Secondly, we calculate $\mathbb{E}[\bar{S}(t)]$:

$$\mathbb{E}\left[\bar{S}(t)\right] = \mathbb{E}\left[\bar{S}(0) + \int_0^t pd\bar{W}(s) - \bar{S}_i(0)\right] = \bar{S}(0) + \mathbb{E}\left[\int_0^t pd\bar{W}(s) - \bar{S}_i(0)\right] = \bar{S}(0) = S(0). \tag{A.100}$$

Thirdly, we calculate $\mathrm{Var}(\bar{S}(t))$:

$$\begin{aligned}
\mathrm{Var}(\bar{S}(t)) &= \mathbb{E}[\bar{S}^2(t)] - \left(\mathbb{E}[\bar{S}(t)]\right)^2 = \mathbb{E}\left[\left(\bar{S}(0) + \int_0^t pd\bar{W}(s)\right)^2\right] - \bar{S}^2(0) \\
&= \mathbb{E}\left[\bar{S}^2(0) + 2\bar{S}(0)\int_0^t pd\bar{W}(s) + \left(\int_0^t pd\bar{W}(s)\right)^2\right] - \bar{S}^2(0) \\
&= \bar{S}^2(0) + 2\bar{S}(0)\mathbb{E}\left[\int_0^t pd\bar{W}(s)\right] + \mathbb{E}\left[\left(\int_0^t pd\bar{W}(s)\right)^2\right] - \bar{S}^2(0) \\
&= \mathbb{E}\left[\int_0^t p^2 ds\right] = \int_0^t \mathbb{E}\left[p^2\right] ds = \int_0^t p^2 ds \\
&= p^2 t.
\end{aligned} \tag{A.101}$$

Finally, we calculate $\mathrm{Cov}(\bar{S}_i(t) - \bar{S}_i(0), \bar{S}(t))$:

$$\begin{aligned}
\mathrm{Cov}(\bar{S}_i(t) - \bar{S}_i(0), \bar{S}(t)) &= \mathbb{E}[(\bar{S}_i(t) - \bar{S}_i(0))\bar{S}(t)] - \mathbb{E}[\bar{S}_i(t) - \bar{S}_i(0)]\mathbb{E}\left[\bar{S}(t)\right] = \mathbb{E}[(\bar{S}_i(t) - \bar{S}_i(0))\bar{S}(t)] \\
&= \mathbb{E}\left[\left(\bar{S}_i(0) + \int_0^t p_i dW_i(s) - \bar{S}_i(0)\right)\left(\bar{S}(0) + \int_0^t pd\bar{W}(s)\right)\right] \\
&= \mathbb{E}\left[\bar{S}(0)\int_0^t p_i dW_i(s) + \left(\int_0^t p_i dW_i(s)\right)\left(\int_0^t pd\bar{W}(s)\right)\right] \\
&= \bar{S}(0)\mathbb{E}\left[\int_0^t p_i dW_i(s)\right] + \mathbb{E}\left[\left(\int_0^t p_i dW_i(s)\right)\left(\int_0^t pd\bar{W}(s)\right)\right] \\
&= \mathbb{E}\left[\int_0^t p_i p \rho_i ds\right] = \int_0^t \mathbb{E}\left[p_i p \rho_i\right] ds = \int_0^t p_i p \rho_i ds \\
&= p_i p \rho_i t.
\end{aligned} \tag{A.102}$$

Note that we made use of Itô isometry and the fact that the expectation of an Itô integral is zero. Substituting (A.99), (A.100), (A.101) and (A.102) in (A.98) we obtain:

$$\mathbb{E}[\bar{S}_i(t) - \bar{S}_i(0) | \bar{S}(t) = x] = \frac{p_i \rho_i}{p}(x - S(0)). \tag{A.103}$$

Now, we apply (5.25) to $\mathbb{E}[\bar{u}_i(t) - 1 | \bar{S}(t) = x]$ we obtain:

$$\mathbb{E}[\bar{u}_i(t) - 1 | \bar{S}(t) = x] = \mathbb{E}[\bar{u}_i(t) - 1] + \frac{\mathrm{Cov}(\bar{u}_i(t) - 1, \bar{S}(t))}{\mathrm{Var}(\bar{S}(t))}(x - \mathbb{E}[\bar{S}(t)]). \tag{A.104}$$

To get an expression for $\mathbb{E}[\bar{u}_i(t) - 1 | \bar{S}(t) = x]$ we only have to calculate two terms, since the other terms are already calculated. First, we calculate $\mathbb{E}[\bar{u}_i(t) - 1]$:

$$\mathbb{E}[\bar{u}_i(t) - 1] = \mathbb{E}\left[\bar{u}_i(0) + \int_0^t \nu_i dZ_i(s) - 1\right] = \mathbb{E}\left[\int_0^t \nu_i dZ_i(s)\right] = 0. \tag{A.105}$$

Finally, we calculate $\mathrm{Cov}(\bar{u}_i(t)-1,\bar{S}(t))$:

$$\begin{aligned}\mathrm{Cov}(\bar{u}_i(t)-1,\bar{S}(t)) &= \mathbb{E}[(\bar{u}_i(t)-1)\bar{S}(t)] - \mathbb{E}[\bar{u}_i(t)-1]\mathbb{E}\left[\bar{S}(t)\right] = \mathbb{E}[(\bar{u}_i(t)-1)\bar{S}(t)] \\ &= \mathbb{E}\left[\left(\bar{u}_i(0)+\int_0^t \nu_i dZ_i(s) - 1\right)\left(\bar{S}(0)+\int_0^t p d\bar{W}(s)\right)\right] \\ &= \mathbb{E}\left[\bar{S}(0)\int_0^t \nu_i dZ_i(s) + \left(\int_0^t \nu_i dZ_i(s)\right)\left(\int_0^t p d\bar{W}(s)\right)\right] \\ &= \bar{S}(0)\mathbb{E}\left[\int_0^t \nu_i dZ_i(s)\right] + \mathbb{E}\left[\left(\int_0^t \nu_i dZ_i(s)\right)\left(\int_0^t p d\bar{W}(s)\right)\right] \\ &= \mathbb{E}\left[\int_0^t \nu_i p\rho_{i+2} ds\right] = \int_0^t \mathbb{E}\left[\nu_i p\rho_{i+2}\right] ds = \int_0^t \nu_i p\rho_{i+2} ds \\ &= \nu_i p\rho_{i+2} t. \end{aligned} \tag{A.106}$$

Substituting (A.105), (A.100), (A.101) and (A.106) in (A.104) we obtain:

$$\mathbb{E}[\bar{u}_i(t)-1|\,\bar{S}(t)=x] = \frac{\nu_i\rho_{i+2}}{p}(x-S(0)). \tag{A.107}$$

A.18 Proof of Lemma 5.4.5

For $\mathbb{E}\left[\sigma^2(t)\middle|\,S(t)=x\right]$ we have:

$$\begin{aligned}\mathbb{E}\left[\sigma^2(t)\middle|\,S(t)=x\right] &= \mathbb{E}\left[f_{11}(t)+f_{22}(t)-2\rho_{ij}f_{12}(t)|\,S(t)=x\right] \\ &= \mathbb{E}\left[f_{11}(t)|\,S(t)=x\right] + \mathbb{E}\left[f_{22}(t)|\,S(t)=x\right] - 2\rho_{12}\mathbb{E}\left[f_{12}(t)|\,S(t)=x\right]. \end{aligned} \tag{A.108}$$

Using (A.103) we can calculate the three terms in (A.108). For $\mathbb{E}\left[f_{11}(t)|\,S(t)=x\right]$ we have:

$$\begin{aligned}\mathbb{E}\left[f_{11}(t)|\,S(t)=x\right] &= \mathbb{E}\left[p_1^2\left(1+2\frac{q_1}{p_1}(S_1(t)-S_1(0))+2(u_1(t)-1)\right)\middle|\,S(t)=x\right] \\ &= p_1^2\left(1+2\frac{q_1}{p_1}\mathbb{E}\left[(S_1(t)-S_1(0))|\,S(t)=x\right]+2\mathbb{E}\left[(u_1(t)-1)|\,S(t)=x\right]\right) \\ &\approx p_1^2\left(1+2\frac{q_1}{p_1}\mathbb{E}\left[(\bar{S}_1(t)-\bar{S}_1(0))\middle|\,\bar{S}(t)=x\right]+2\mathbb{E}\left[(\bar{u}_1(t)-1)|\,\bar{S}(t)=x\right]\right) \\ &= p_1^2\left(1+2\frac{q_1}{p_1}\frac{p_1\rho_1}{p}(x-S(0))+2\frac{\nu_1\rho_3}{p}(x-S(0))\right) \\ &= p_1^2\left(1+2\frac{q_1\rho_1}{p}(x-S(0))+2\frac{\nu_1\rho_3}{p}(x-S(0))\right). \end{aligned} \tag{A.109}$$

For $\mathbb{E}\left[f_{22}(t)|\,S(t)=x\right]$ we have:

$$\begin{aligned}\mathbb{E}\left[f_{22}(t)|\,S(t)=x\right] &= \mathbb{E}\left[p_2^2\left(1+2\frac{q_2}{p_2}(S_2(t)-S_2(0))+2(u_2(t)-1)\right)\middle|\,S(t)=x\right] \\ &= p_2^2\left(1+2\frac{q_2}{p_2}\mathbb{E}\left[(S_2(t)-S_2(0))|\,S(t)=x\right]+2\mathbb{E}\left[(u_2(t)-1)|\,S(t)=x\right]\right) \\ &\approx p_2^2\left(1+2\frac{q_2}{p_2}\mathbb{E}\left[(\bar{S}_2(t)-\bar{S}_2(0))\middle|\,\bar{S}(t)=x\right]+2\mathbb{E}\left[(\bar{u}_2(t)-1)|\,\bar{S}(t)=x\right]\right) \\ &= p_2^2\left(1+2\frac{q_2}{p_2}\frac{p_2\rho_2}{p}(x-S(0))+2\frac{\nu_2\rho_4}{p}(x-S(0))\right) \\ &= p_2^2\left(1+2\frac{q_2\rho_2}{p}(x-S(0))+2\frac{\nu_2\rho_4}{p}(x-S(0))\right). \end{aligned} \tag{A.110}$$

For $\mathbb{E}\left[f_{12}(t)\middle|\, S(t)=x\right]$ we have:

$$\begin{aligned}
\mathbb{E}\left[f_{12}(t)\middle|\, S(t)=x\right] = \mathbb{E}\Bigg[p_1p_2\Bigg(&1+\frac{q_1}{p_1}(S_1(t)-S_1(0))+\frac{q_2}{p_2}(S_2(t)-S_2(0)) \\
&+\ (u_1(t)-1)+(u_2(t)-1)\Bigg)\Bigg|\, S(t)=x\Bigg] \\
= p_1p_2 &+ p_2q_1\mathbb{E}\left[S_1(t)-S_1(0)\middle|\, S(t)=x\right]+p_1q_2\mathbb{E}\left[S_2(t)-S_2(0)\middle|\, S(t)=x\right] \\
&+p_1p_2\mathbb{E}\left[u_1(t)-1\middle|\, S(t)=x\right]+p_1p_2\mathbb{E}\left[u_2(t)-1\middle|\, S(t)=x\right] \\
\approx p_1p_2 &+ p_2q_1\mathbb{E}\left[\bar{S}_1(t)-\bar{S}_1(0)\middle|\, \bar{S}(t)=x\right]+p_1q_2\mathbb{E}\left[\bar{S}_2(t)-\bar{S}_2(0)\middle|\, \bar{S}(t)=x\right] \\
&+p_1p_2\mathbb{E}\left[\bar{u}_1(t)-1\middle|\, \bar{S}(t)=x\right]+p_1p_2\mathbb{E}\left[\bar{u}_2(t)-1\middle|\, \bar{S}(t)=x\right] \\
= p_1p_2 &+ p_2q_1\frac{p_1\rho_1}{p}(x-S(0))+p_1q_2\frac{p_2\rho_2}{p}(x-S(0)) \\
&+p_1p_2\frac{\nu_1\rho_3}{p}(x-S(0))+p_1p_2\frac{\nu_2\rho_4}{p}(x-S(0)). \qquad \text{(A.111)}
\end{aligned}$$

Substituting (A.109), (A.110) and (A.111) in (A.108) we obtain:

$$\begin{aligned}
\mathbb{E}\left[\sigma^2(t)\middle|\, S(t)=x\right] \approx p_1^2 &+ p_2^2 - 2p_1p_2\rho_{12} \\
&+2\frac{p_1^2q_1\rho_1}{p}(x-S(0))+2\frac{p_1^2\nu_1\rho_3}{p}(x-S(0)) \\
&+2\frac{p_2^2q_2\rho_2}{p}(x-S(0))+2\frac{p_2^2\nu_2\rho_4}{p}(x-S(0)) \\
&-2\frac{p_2q_1p_1\rho_1\rho_{12}}{p}(x-S(0))-2\frac{p_1p_2q_2\rho_2\rho_{12}}{p}(x-S(0)) \\
&-2\frac{p_1p_2\nu_1\rho_3\rho_{12}}{p}(x-S(0))-2\frac{p_1p_2\nu_2\rho_4\rho_{12}}{p}(x-S(0)) \\
= p_1^2 &+ p_2^2 - 2p_1p_2\rho_{12} \\
&+\frac{2}{p}\left(p_1^2q_1\rho_1+p_1^2\nu_1\rho_3+p_2^2q_2\rho_2+p_2^2\nu_2\rho_4-p_2q_1p_1\rho_1\rho_{12}-p_1p_2q_2\rho_2\rho_{12}-\right. \\
&\left.p_1p_2\nu_1\rho_3\rho_{12}-p_1p_2\nu_2\rho_4\rho_{12}\right)(x-S(0)) \\
= p_1^2 &+ p_2^2 - 2p_1p_2\rho_{12} \\
&+\frac{2}{p}\left(p_1^2(q_1\rho_1+\nu_1\rho_3)+p_2^2(q_2\rho_2+\nu_2\rho_4)-p_1p_2\rho_{12}(q_1\rho_1+q_2\rho_2+\nu_1\rho_3\right. \\
&\left.+\nu_2\rho_4)\right)(x-S(0)) \\
= p_1^2 &+ p_2^2 - 2p_1p_2\rho \\
&+\frac{2}{p}\left(p_1^2(q_1\rho_1+\nu_1\rho_3)+p_2^2(q_2\rho_2+\nu_2\rho_4)-p_1p_2\rho(q_1\rho_1+q_2\rho_2+\nu_1\rho_3\right. \\
&\left.+\nu_2\rho_4)\right)(x-S(0)), \\
= p^2 &+ (x-S(0))\Theta_1, \qquad \text{(A.112)}
\end{aligned}$$

with

$$p=\sqrt{p_1^2+p_2^2-2p_1p_2\rho}, \qquad \text{(A.113)}$$

$$\Theta_1=\frac{2}{p}\left(p_1^2(q_1\rho_1+\nu_1\rho_3)+p_2^2(q_2\rho_2+\nu_2\rho_4)-p_1p_2\rho(q_1\rho_1+q_2\rho_2+\nu_1\rho_3+\nu_2\rho_4)\right). \qquad \text{(A.114)}$$

Next, we compute $\mathbb{E}\left[u^2(t)\middle|\, S(t)=x\right]$:

$$\begin{aligned}\mathbb{E}\left[u^2(t)\middle|\, S(t)=x\right] &= \mathbb{E}\left[g_{11}(t)+g_{22}(t)-2\rho_{ij}g_{12}(t)\middle|\, S(t)=x\right] \\ &= \mathbb{E}\left[g_{11}(t)\middle|\, S(t)=x\right]+\mathbb{E}\left[g_{22}(t)\middle|\, S(t)=x\right]-2\rho_{ij}\mathbb{E}\left[g_{12}(t)\middle|\, S(t)=x\right].\end{aligned} \tag{A.115}$$

Using (A.107) we can calculate the three terms in (A.115). For $\mathbb{E}\left[g_{11}(t)\middle|\, S(t)=x\right]$ we have:

$$\begin{aligned}\mathbb{E}\left[g_{11}(t)\middle|\, S(t)=x\right] &= \mathbb{E}\left[\frac{p_1^2}{p^2}\left(1+2(u_1(t)-1)\right)\middle|\, S(t)=x\right] = \frac{p_1^2}{p^2}\left(1+2\mathbb{E}\left[u_1(t)-1\middle|\, S(t)=x\right]\right) \\ &\approx \frac{p_1^2}{p^2}\left(1+2\mathbb{E}\left[\bar{u}_1(t)-1\middle|\, \bar{S}(t)=x\right]\right) \\ &= \frac{p_1^2}{p^2}\left(1+2\frac{\nu_1\rho_3}{p}(x-S(0))\right).\end{aligned} \tag{A.116}$$

For $\mathbb{E}\left[g_{22}(t)\middle|\, S(t)=x\right]$ we have:

$$\begin{aligned}\mathbb{E}\left[g_{22}(t)\middle|\, S(t)=x\right] &= \mathbb{E}\left[\frac{p_2^2}{p^2}\left(1+2(u_2(t)-1)\right)\middle|\, S(t)=x\right] = \frac{p_2^2}{p^2}\left(1+2\mathbb{E}\left[u_2(t)-1\middle|\, S(t)=x\right]\right) \\ &\approx \frac{p_2^2}{p^2}\left(1+2\mathbb{E}\left[\bar{u}_2(t)-1\middle|\, \bar{S}(t)=x\right]\right) \\ &= \frac{p_2^2}{p^2}\left(1+2\frac{\nu_2\rho_4}{p}(x-S(0))\right).\end{aligned} \tag{A.117}$$

For $\mathbb{E}\left[g_{12}(t)\middle|\, S(t)=x\right]$ we have:

$$\begin{aligned}\mathbb{E}\left[g_{12}(t)\middle|\, S(t)=x\right] &= \mathbb{E}\left[\frac{p_1p_2}{p^2}\left(1+(u_1(t)-1)+(u_2(t)-1)\right)\middle|\, S(t)=x\right] \\ &= \frac{p_1p_2}{p^2}\left(1+\mathbb{E}\left[u_1(t)-1\middle|\, S(t)=x\right]+\mathbb{E}\left[u_2(t)-1\middle|\, S(t)=x\right]\right) \qquad \text{(A.118)} \\ &\approx \frac{p_1p_2}{p^2}\left(1+\mathbb{E}\left[\bar{u}_1(t)-1\middle|\, \bar{S}(t)=x\right]+\mathbb{E}\left[\bar{u}_2(t)-1\middle|\, \bar{S}(t)=x\right]\right) \\ &= \frac{p_1p_2}{p^2}\left(1+\frac{\nu_1\rho_3}{p}(x-S(0))+\frac{\nu_2\rho_4}{p}(x-S(0))\right).\end{aligned}$$

Substituting (A.116), (A.117) and (A.119) in (A.115) we obtain:

$$\begin{aligned}\mathbb{E}\left[u^2(t)\middle|\, S(t)=x\right] &\approx \frac{p_1^2}{p^2}\left(1+2\frac{\nu_1\rho_3}{p}(x-S(0))\right)+\frac{p_2^2}{p^2}\left(1+2\frac{\nu_2\rho_4}{p}(x-S(0))\right) \\ &\quad -2\rho_{ij}\frac{p_1p_2}{p^2}\left(1+\frac{\nu_1\rho_3}{p}(x-S(0))+\frac{\nu_2\rho_4}{p}(x-S(0))\right) \\ &= \frac{1}{p^2}(p_1^2+p_2^2-2\rho_{12}p_1p_2) \\ &\quad +\frac{2}{p^3}\left(p_1^2\nu_1\rho_3+p_2^2\nu_2\rho_4-p_1p_2\nu_1\rho_3\rho_{12}-p_1p_2\nu_2\rho_4\rho_{12}\right)(x-S(0)) \\ &= \frac{p^2}{p^2}+\frac{2}{p^3}\left(\nu_1p_1(p_1-p_2\rho_{12})\rho_3+\nu_2p_2(p_2-p_1\rho_{12})\rho_4\right)(x-S(0)) \\ &= \frac{p^2}{p^2}+\frac{2}{p^3}\left(\nu_1p_1(p_1-p_2\rho)\rho_3+\nu_2p_2(p_2-p_1\rho)\rho_4\right)(x-S(0)) \\ &= 1+(x-S(0))\Theta_2,\end{aligned} \tag{A.119}$$

with

$$\Theta_2 = \frac{2}{p^3}\left(\nu_1p_1(p_1-p_2\rho)\rho_3+\nu_2p_2(p_2-p_1\rho)\rho_4\right). \tag{A.120}$$

A.19 Proof of Lemma 5.4.6

From Lemma 5.4.5 it follows that,

$$F^2(x) \approx \frac{p^2 + (x - S(0))\Theta_1}{1 + (x - S(0))\Theta_2}. \tag{A.121}$$

Therefore,

$$F(x) \approx \sqrt{\frac{p^2 + (x - S(0))\Theta_1}{1 + (x - S(0))\Theta_2}}. \tag{A.122}$$

We want to calculate $F(S(0))$ and $F'(S(0))$. For $F(S(0))$ we have

$$F(S(0)) = \sqrt{\frac{p^2 + (S(0) - S(0))\Theta_1}{1 + (S(0) - S(0))\Theta_2}} = p. \tag{A.123}$$

Next, we determine $F'(x)$:

$$\begin{aligned} F'(x) &= \frac{1}{2}\left(\frac{p^2 + (x - S(0))\Theta_1}{1 + (x - S(0))\Theta_2}\right)^{-1/2} \frac{d}{dx}\left(\frac{p^2 + (x - S(0))\Theta_1}{1 + (x - S(0))\Theta_2}\right) \\ &= \frac{1}{2\sqrt{\frac{p^2+(x-S(0))\Theta_1}{1+(x-S(0))\Theta_2}}} \frac{(1 + (x - S(0))\Theta_2)\Theta_1 - (p^2 + (x - S(0))\Theta_1)\Theta_2}{(1 + (x - S(0))\Theta_2)^2}. \end{aligned} \tag{A.124}$$

So for $F'(S(0))$ we obtain:

$$F'(S(0)) = \frac{1}{2p}\left(\Theta_1 - p^2\Theta_2\right). \tag{A.125}$$

Substituting (A.114) and (A.120) in (A.125) gives us

$$\begin{aligned} F'(S(0)) &= \frac{1}{p^2}\big(\, p_1^2 q_1\rho_1 + p_1\nu_1\rho_3 + p_2^2 q_2\rho_2 + p_2^2\nu_2\rho_4 - p_1p_2q_1\rho_1\rho - p_1p_2q_2\rho_2\rho - p_1p_2\nu_1\rho_3\rho \\ &\quad - p_1p_2\nu_2\rho_4\rho - p_1^2\nu_1\rho_3 + p_1p_2\nu_1\rho_3\rho - p_2^2\nu_2\rho_4 + p_1p_2\nu_2\rho_4\rho\big) \\ &= \frac{1}{p^2}\left(p_1^2q_1\rho_1 + p_2^2q_2\rho_2 - p_1p_2q_1\rho_1\rho - p_1p_2q_2\rho_2\rho\right) \\ &= \frac{1}{p^2}\left(p_1q_1\rho_1(p_1 - p_2\rho) + p_2q_2\rho_2(p_2 - p_1\rho)\right). \end{aligned}$$

Making use of

$$\rho_1 = \frac{p_1 - p_2\rho}{p}, \tag{A.126}$$

$$\rho_2 = \frac{p_1\rho - p_2}{\rho}, \tag{A.127}$$

we obtain:

$$F'(S(0)) = q, \tag{A.128}$$

where

$$q = \frac{p_1q_1\rho_1^2 - p_2q_2\rho_2^2}{p}. \tag{A.129}$$

Appendix B

Market Data

In this appendix we present parts of the market data that were used for the numerical experiments.

B.1 Market Data 2013

Start date	ZCB	Accrual
15-sep-14	0.995	0.997
14-sep-15	0.987	1.000
13-sep-16	0.972	1.000
13-sep-17	0.952	1.000
13-sep-18	0.929	1.000
13-sep-19	0.903	1.005
14-sep-20	0.875	0.997
13-sep-21	0.847	1.000
13-sep-22	0.818	1.000
13-sep-23	0.790	1.002
13-sep-24	0.761	1.005
15-sep-25	0.734	0.997
14-sep-26	0.707	0.997
13-sep-27	0.682	1.002
13-sep-28	0.658	1.000
13-sep-29	0.636	1.000
13-sep-30	0.615	1.005
15-sep-31	0.596	0.997
13-sep-32	0.578	1.000
13-sep-33	0.560	1.000
13-sep-34	0.544	-

Table B.1: Market data 2013 for 10Y and 2Y CMS rate with 12M frequency in 2013. Today's date is 11-sep-13.

Reset date	Pay date	α_0	β	ρ	ν
11-sep-14	14-sep-15	0.503	0.5	0.218	0.523
10-sep-15	13-sep-16	0.464	0.5	0.239	0.658
9-sep-16	13-sep-17	0.462	0.5	0.075	0.418
11-sep-17	13-sep-18	0.428	0.5	0.182	0.525
11-sep-18	13-sep-19	0.432	0.5	0.016	0.391
11-sep-19	14-sep-20	0.411	0.5	0.210	0.406
10-sep-20	13-sep-21	0.403	0.5	-0.010	0.365
9-sep-21	13-sep-22	0.389	0.5	0.106	0.375
9-sep-22	13-sep-23	0.387	0.5	-0.039	0.341
11-sep-23	13-sep-24	0.371	0.5	0.066	0.345

Table B.2: SABR parameters for the 10Y CMS rate in 2013. The SABR parameters were calibrated from given implied volatilities. Additionally, the reset dates and pay dates are reported.

Reset date	Pay date	α_0	β	ρ	ν
11-sep-14	14-sep-15	0.635	0.5	0.540	0.565
10-sep-15	13-sep-16	0.633	0.5	0.477	0.543
9-sep-16	13-sep-17	0.599	0.5	0.441	0.441
11-sep-17	13-sep-18	0.550	0.5	0.356	0.428
11-sep-18	13-sep-19	0.513	0.5	0.280	0.356
11-sep-19	14-sep-20	0.482	0.5	0.350	0.336
10-sep-20	13-sep-21	0.445	0.5	0.189	0.334
9-sep-21	13-sep-22	0.426	0.5	0.195	0.340
9-sep-22	13-sep-23	0.410	0.5	0.108	0.318
11-sep-23	13-sep-24	0.382	0.5	0.119	0.335

Table B.3: SABR parameters for the 2Y CMS rate in 2013. The SABR parameters were calibrated from given implied volatilities. Additionally, the reset dates and pay dates are reported.

period	$\tilde{S}_i(0)$	$\tilde{\alpha}_i^0$	$\tilde{\beta}_i$	$\tilde{\gamma}_{ii}$	$\tilde{\nu}_i$
10Y	0.026	0.510	0.5	0.253	0.531
2Y	0.011	0.641	0.5	0.548	0.571

Table B.4: Adjusted SABR parameters for the 10Y and 2Y CMS-adjusted rate in 2013. The start date is 1 year from today, $T_0 = 1$.

period	$\tilde{S}_i(0)$	$\tilde{\alpha}_i^0$	$\tilde{\beta}_i$	$\tilde{\gamma}_{ii}$	$\tilde{\nu}_i$
10Y	0.037	0.456	0.5	0.207	0.419
2Y	0.031	0.529	0.5	0.336	0.368

Table B.5: Adjusted SABR parameters for the 10Y and 2Y CMS-adjusted rate in 2013. The start date is 5 years from today, $T_0 = 5$.

Model	$\tilde{\rho}$	$\tilde{\gamma}_{12}$	$\tilde{\gamma}_{21}$	$\tilde{\xi}$
DD SABR	0.691	0.999	0.999	1.000
Copula	0.736	-	-	-

Table B.6: Calibrated correlation parameters DD SABR model and copula approach. The start date is 1 year from today, $T_0 = 1$.

Model	$\tilde{\rho}$	$\tilde{\gamma}_{12}$	$\tilde{\gamma}_{21}$	$\tilde{\xi}$
DD SABR	0.835	0.809	0.809	1.000
Copula	0.877	-	-	-

Table B.7: Calibrated correlation parameters DD SABR model and copula approach. The start date is 5 years from today, $T_0 = 5$.

B.2 Market Data 2007

Start date	ZCB	Accrual
15-sep-08	0.956	0.997
14-sep-09	0.918	0.997
13-sep-10	0.879	1.000
13-sep-11	0.842	1.002
13-sep-12	0.807	1.000
13-sep-13	0.772	1.005
15-sep-14	0.738	0.997
14-sep-15	0.705	1.000
13-sep-16	0.673	1.000
13-sep-17	0.642	1.000
13-sep-18	0.612	1.000
13-sep-19	0.583	1.005
14-sep-20	0.555	0.997
13-sep-21	0.528	1.000
13-sep-22	0.503	1.000
13-sep-23	0.479	1.002
13-sep-24	0.456	1.005
15-sep-25	0.434	0.997
14-sep-26	0.414	0.997
13-sep-27	0.395	1.002
13-sep-28	0.377	-

Table B.8: Market data 2013 for 10Y and 2Y CMS rate with 12M frequency in 2007. Today's date is 11-sep-07.

Reset date	Pay date	α_0	β	ρ	ν
11-sep-08	14-sep-09	0.275	0.5	-0.004	0.351
10-sep-09	13-sep-10	0.263	0.5	-0.116	0.358
9-sep-10	13-sep-11	0.259	0.5	-0.138	0.296
9-sep-11	13-sep-12	0.250	0.5	-0.116	0.349
11-sep-12	13-sep-13	0.248	0.5	-0.210	0.283
11-sep-13	15-sep-14	0.239	0.5	-0.131	0.326
11-sep-14	14-sep-15	0.240	0.5	-0.211	0.265
10-sep-15	13-sep-16	0.230	0.5	-0.137	0.304
9-sep-16	13-sep-17	0.231	0.5	-0.222	0.259
11-sep-17	13-sep-18	0.223	0.5	-0.132	0.278

Table B.9: SABR parameters for the 10Y CMS rate in 2007. The SABR parameters were calibrated from given implied volatilities. Additionally, the reset dates and pay dates are reported.

Reset date	Pay date	α_0	β	ρ	ν
11-sep-08	14-sep-09	0.335	0.5	-0.092	0.211
10-sep-09	13-sep-10	0.310	0.5	-0.106	0.248
9-sep-10	13-sep-11	0.294	0.5	-0.047	0.246
9-sep-11	13-sep-12	0.283	0.5	-0.010	0.268
11-sep-12	13-sep-13	0.275	0.5	-0.044	0.252
11-sep-13	15-sep-14	0.264	0.5	-0.028	0.273
11-sep-14	14-sep-15	0.258	0.5	-0.073	0.259
10-sep-15	13-sep-16	0.249	0.5	-0.055	0.276
9-sep-16	13-sep-17	0.244	0.5	-0.070	0.261
11-sep-17	13-sep-18	0.237	0.5	-0.054	0.265

Table B.10: SABR parameters for the 2Y CMS rate in 2007. The SABR parameters were calibrated from given implied volatilities. Additionally, the reset dates and pay dates are reported.

period	$\tilde{S}_i(0)$	$\tilde{\alpha}_i^0$	$\tilde{\beta}_i$	$\tilde{\gamma}_{ii}$	$\tilde{\nu}_i$
10Y	0.045	0.275	0.5	-0.050	0.366
2Y	0.043	0.333	0.5	-0.212	0.334

Table B.11: Adjusted SABR parameters for the 10Y and 2Y CMS-adjusted rate in 2007. The start date is 1 year from today, $T_0 = 1$.

period	$\tilde{S}_i(0)$	$\tilde{\alpha}_i^0$	$\tilde{\beta}_i$	$\tilde{\gamma}_{ii}$	$\tilde{\nu}_i$
10Y	0.048	0.246	0.5	-0.141	0.285
2Y	0.045	0.275	0.5	-0.087	0.256

Table B.12: Adjusted SABR parameters for the 10Y and 2Y CMS-adjusted rate in 2007. The start date is 5 years from today, $T_0 = 5$.