Calibration risk in pricing excess interest options

Author
R.Zeeman

Supervisors
Universiteit Utrecht
Prof. dr. ir. E.J. Balder
Ernst & Young Actuarissen
Ir. T.S. de Graaf

December 9, 2008

Preface

Proudly I present my thesis that finishes my Master's studies Mathematical Sciences at Universiteit Utrecht. The mathematical fields that have my attention and interest are optimization, numerical mathematics and financial and stochastic mathematics. I got the opportunity at Ernst & Young Actuarissen to combine these flavours of mathematics in an investigation on the calibration risk involved in pricing profit sharing contracts written by insurance companies.

Finishing this thesis could not have been possible without the advise and patience of professor Balder, my supervisor at Universiteit Utrecht and the person who got me enthusiastic for optimization problems.

Besides the support from the university, I appreciate the support from my supervisor Tony de Graaf of Ernst & Young Actuarissen. He spent loads of time to explain the financial industry that was totally new for me. Moreover, his insights and suggestions were very constructive. My colleague Wim Weijgertze friendly pushed me to finish this thesis before it would become a never-ending story. Moreover I want to mention the very pleasant atmosphere created by all people at Ernst & Young Actuarissen that motivated me during the writing of this thesis.

Finally, I would like to thank my family, friends and my fellow students for their interest, support and the great moments during my studies.

Executive Summary

Due to renewals in the financial reporting standards, financial instruments on the balance of insurance companies are reported at their market consistent, fair, values. Since there is not always an active market for these financial instruments, these fair values are often computed by particular models. One can imagine that the choice of the model will have impact on the fair value of the financial instrument. But when a model is chosen, it has to be adapted to the actual market situation in order to produce market consistent prices. This process of fitting is called *calibration*. Model calibration is not a straightforward process, several choices have to be made during the calibration process. We investigate the effect of these choices, in other words, we determine the *calibration risk*.

We look at the value of a particular embedded option in an insurance contract: an excess interest option. The value of this embedded option is calculated with two commonly used interest rate models: the two-factor Hull-White model and the Libor Market Model. We calibrate these two models to market values of interest rate swaptions and to market values of interest rate caps. The market values of these instruments can be either prices or implied volatilities. Cap data consist of market values for a whole range of caps or market values for only at-the-money caps. The choice between these collections of cap data is also investigated.

Actually, calibrating boils down to minimizing the 'difference' between model and market values over a set of model parameters. This difference can be measured in several ways. One can minimize absolute differences or one can minimize relative differences between market and model values.

In pricing the embedded option, two risks can be identified, see [DH07]. One is *model risk*. This is the impact of the model choice on the value of the option. We quantify model risk by the fraction between two model prices of the embedded option.
Calibration risk is the other risk which arises from the methods chosen in the calibration proces. Calibration risk is measured as the fraction between the option values within one model, but with different calibrations applied to that model.

The goal of this thesis is formulated as:

> Investigate the impact on pricing an excess interest option of different calibration methods by valuing the option with the two interest rate models, where each model is calibrated in different ways:
>
> a) with respect to absolute or relative differences of swaption prices or swaption implied volatilities;
>
> b) with respect to absolute or relative differences of all cap prices or only at-the-money cap prices.
>
> Combining these possibilities results in eight different ways to calibrate both interest rate models, which we will investigate.

Two important conclusions we draw from the investigation are:

- If one decides to calibrate to swaptions, then the impact of the model choice on the option price is larger than the impact of the calibration choice, i.e. model risk is larger than calibration risk.
- Calibration to caps must be performed with respect to all caps by minimizing absolute price differences, in order to obtain proper model parameters.

Contents

Chapter 1

Introduction

Insurance companies are obliged to report and value their assets and liabilities market consistently. This way of reporting requires the use of interest rate models to value embedded options in profit sharing contracts as there (still) exists no active market for these embedded options.

Ernst & Young Actuarissen wants to have a thorough knowledge of interest rate models and the techniques to value embedded options with interest rate models. The valuation of embedded options can be performed by analytic formulas, but is more often done by stochastic scenarios. Ernst & Young Actuarissen has a license for the Economic Scenario Generator (ESG), developed by Barrie and Hibbert[1], to broaden the knowledge on interest rate models and to improve valuation skills.

The ESG is a widely used tool that models capital structures by Monte-Carlo simulations of economic stochastic models. Among these models there are three interest rate models: the two-factor Vasicek model, the two-factor Black-Karasinski and a two-factor Libor Market Model. In [Kes07], Kessels investigated these interest rate models and determined the value of the liabilities for profit sharing contracts with these three interest rate models.

Interest rate models are to be configured to the market before they can be deployed for valuation purposes. This process is called *market consistent calibration*. There are numerous ways to perform calibrations of interest rate models. In this thesis, we focus on two interest rate models: the Libor Market Model as it is used by Barrie and Hibbert and the two-factor Hull-White model, a commonly used extension of the two-factor Vasicek model. With these interest models, we value an embedded option in a profit sharing contract: an excess interest option. The objective of this thesis is to determine the calibration risk involved in pricing this excess interest option. As in [DH07], we define calibration risk as the fraction

[1] see www.barrhibb.com

between model prices for the derivative to value, where the model is calibrated in two different ways.

The investigation of calibration risk is thus performed by valuing the option by the two interest models where both models are calibrated in several ways to the market.

In chapter two, financial definitions and products are explained. Several types of interest rates are presented. The two interest rate models are introduced and the way to value interest rate derivatives is investigated.

The calibration routine is the subject of chapter three. In an intermezzo we start with the calibration of a simple interest rate model. The points that are taken into consideration for a market consistent calibration arise from this intermezzo and are treated in the remainder of chapter three.

The excess interest option that is valued with the interest rate models is introduced in chapter four. The calibration results are presented and the option values that result from the calibrations are discussed. The calibration risk is investigated and finally compared with model risk: the effect of the model choice on the value of the excess interest option.

The conclusions that we draw from the results in chapter four are stated in chapter five. Points for further investigation are suggested after the conclusions.

In the appendix, some formulas and derivations are given for the Two-Additive-Factor Gaussian model are given. Also some background information on (Monte Carlo) simulation can be found there.

Chapter 2

Financial setting

We can mainly distinguish between three types of interest models. The first approach is to take the short-rate as the basis for modeling the term-structure of the interest rates. The short-rate, also called instantaneous spot rate, is the interest earned over a infinitesimal time-interval dt. This type of modeling is convenient from a mathematical point of view and leads to tractable models. However, one cannot trade in short rates and hence valuation of financial instruments becomes a complicated job, just like calibration of the model.
The second approach avoids these problems by basing the model on observable market rates, like Libor or swap rates. Therefore, these models are called *market models*. These models have a more complex setup, but can be made to fit the market prices perfectly. Furthermore, some options are priced by these models with the standard market model formulas. This is convenient and does not require numerical procedures.
The third way is modeling instantaneous forward rates in stead of short rates or observable market rates. We do not focus on this class of interest rate models since these models are less used.

2.1 Financial instruments

To develop the interest models, we first need to define some rates and specify financial instruments.

2.1.1 Bonds

A T-maturity **zero-coupon bond**, also called *pure discount bond*, is a contract that guarantees its holder the payment of one unit of currency at time T. The contract value at time $t < T$ is denoted by $P(t, T)$. Consequently, $P(T, T) = 1$ for all T.
Next to zero-coupon bonds there are **coupon-bearing bonds**. These are agree-

ments with interest payments over a principal at a number of dates until maturity. A stylized example: a two-year bond with principal of €100 and semi-annual coupon payments at 5% provides €5 at three dates, in six months, one year and eighteen months, and €105 in two years.

The **continuously-compounded spot interest rate** $R(t,T)$ at time t for maturity T is the constant rate at which an investment of value $P(t,T)$ at time t accrues continuously to yield 1 at maturity T. Therefore, it is also called the *yield* over time interval $[t,T]$. Thus,

$$R(t,T) = -\frac{\log P(t,T)}{\tau(t,T)}, \tag{2.1}$$

where $\tau(t,T)$ is the daycount fraction between t and T and equals approximately $T-t$.
For a fixed t, the graph of the function $T \mapsto R(t,T)$ is called the *yield curve*.

Next to continuous compounding, there is a concept called *simple compounding*, which is applied when an investment grows proportionally to the time of the investment. Hence, we define the **simply-compounded spot interest rate** $L(t,T)$ at time t for the maturity T as the constant rate at which an investment has to be made to produce an amount of 1 unit of currency at maturity, starting with $P(t,T)$ at time t, when accruing occurs proportionally to the investment time. The formula for this is

$$L(t,T) = \frac{1-P(t,T)}{\tau(t,T)P(t,T)}. \tag{2.2}$$

The **short rate** is now defined as

$$r(t) = \lim_{T\downarrow t} R(t,T) = \lim_{T\downarrow t} L(t,T). \tag{2.3}$$

Hence, the short rate can be interpreted as the interest earned over an infinitesimal interval dt.

The **simply-compounded forward interest rate** at time t for maturity T and expiry U is that value of the fixed rate r at which one unit of currency at time T accrues to $(1+\tau(T,U)r)$, determined at time t. The forward rate can be expressed as

$$F(t;T,S) = \frac{1}{\tau(T,U)}\left(\frac{P(t,T)}{P(t,U)} - 1\right). \tag{2.4}$$

Rates of this kind are also called *forward Libor rates* or just *forward rates*. Libor means “London Interbank Offer Rate”, the rate at which banks are willing to lend money to other banks in the London money market. Time T is known as the *maturity* of the forward rate and $(U-T)$ is called the *tenor*.

The **instantaneous forward interest rate**, short *forward rate*, $f(t,T)$ at T contracted at time t , is defined as the limit of the simply-compounded forward interest rate as maturity S collapses to expiry T. So

$$\begin{aligned} f(t,T) &= \lim_{U\downarrow T} F(t;T,U) \\ &= -\frac{1}{P(t,T)}\frac{\partial P(t,T)}{\partial T} \\ &= -\frac{\partial \log P(t,T)}{\partial T}. \end{aligned} \tag{2.5}$$

2.1.2 Bank account

We define $B(t)$ to be the value of a bank account at time $t \geq 0$. We assume that $B(0) = 1$ and that the account evolves according the differential equation

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

where $r(t)$ is the short rate. As a consequence,

$$B(t) = \exp\left(\int_0^t r(s)ds\right). \tag{2.6}$$

The **stochastic discount factor** $D(t,T)$ is defined for two time instants t and T as the amount of currency at t that equals one unit of currency at time T. It is given by

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

If the rates $r(t)$ are deterministic, then D is deterministic and necessarily $D(t,T) = P(t,T)$. But if the rates are stochastic, $D(t,T)$ is a random quantity, depending on future behavior of the rates $r(t)$. On the other hand, $P(t,T)$ has to be known at time t and hence is deterministic.

2.1.3 Interest Rate Swaps

An interest rate swap (IRS) is an agreement between two parties to exchange cash flows in the future; e.g. for a number of years, a cash flow equal to the interest on a notional amount at a fixed interest rate is swapped with a floating rate on the same principal. Common choices for the floating rate are the Libor rate or the Euro Interbank Offered Rate (Euribor) which is the equivalent for the Libor on the Euro money market. We will use the Libor rate (2.4).

The fixed rate S in the contract is called the *par swap rate* and is determined such that the swap is a fair contract, that means, the value of the swap at the start is zero.
The two parts of the contract are named legs, such one can distinguish a fixed leg, paying the fixed interest rate payments and a floating leg, paying the Libor rate payments.

When the fixed leg is paid and the floating leg is received, the swap is called a *payer swap*. In the other case we call it a *receiver swap*.

So the value V_{swap} at time t of a receiver swap swap can be expressed as

$$V_{\text{swap}}(t) = B_{\text{fix}}(t) - B_{\text{float}}(t). \tag{2.7}$$

For convenience, assume that the payments of both legs occur on the same dates $T_1, \dots, T_n$ and the Libor rate is set on the dates $T_0, T_1, \dots, T_{n-1}$. Denote the daycount fraction between T_{i-1} and T_i by δ_i. The set of dates $\{T_i\}$ is denoted as the *tenor structure*.

$$\begin{aligned} V_{\text{swap}}(t) &= \sum_{i=1}^{n} \delta_i (S - F(t; T_{i-1}, T_i)) P(t, T_i) \\ &= \sum_{i=1}^{n} \delta_i S P(t, T_i) - P(t, T_0) + P(t, T_n), \qquad t \leq T_0 \end{aligned} \tag{2.8}$$

which matches (2.7), since

$$\begin{aligned} B_{\text{fix}}(t) &= \sum_{i=1}^{n} \delta_i S P(t, T_i) + P(t, T_n), \\ B_{\text{float}}(t) &= P(t, T_0). \end{aligned}$$

The (par) swap rate S is determined at $t = T_0 = 0$ by the equation $V_{\text{swap}}(0) = 0$. We obtain

$$S = \frac{P(0, T_0) - P(0, T_n)}{\sum_{i=1}^{n} \delta_i P(0, T_i)}. \tag{2.9}$$

The *forward swap rate* $S_{m,n}(t)$ is the rate over the period $[T_m, T_n]$ that satisfies at $t < T_m$:

$$S_{m,n}(t) = \frac{P(t, T_m) - P(t, T_n)}{\sum_{i=m+1}^{n} \delta_i P(t, T_i)}. \tag{2.10}$$

2.1.4 Swaptions

A *swaption* is an agreement that gives the owner the right to enter into a swap. The specifications of a swaption are

- the principal amount L of the underlying swap;
- the term structure $\{T_0, \ldots, T_n\}$ of the swap;
- the expiry date of the option, we assume it to be T_0;
- the strike rate, i.e. the swap rate at which one can enter.

Because the floating leg in the swap underlying the swaption is worth par, the swaption can be regarded as a European option on the fixed leg with par strike.

2.1.5 Caps and Floors

A *cap* can be seen as a payer swap with payoff only if it has a positive value. The discounted payoff is therefore

$$\sum_{i=1}^{n} D(t, T_i) N \delta_i [F(t; T_{i-1}, T_i) - K]^+. \tag{2.11}$$

Here we use the notation $[\cdot]^+ := \max(\cdot, 0)$.
A cap protects its holder against high Libor rates. Suppose that he is obliged to pay the Libor rate on the notional amount N, resetting at dates $\{T_i\}$. If one expects that the Libor rate will increase, he can enter into a cap such that at times T_i he pays L and receives $(L - K)^+$. High Libor rates are thus capped at rate K.
Each term of the sum (2.11) defines a contract named *caplet*. A caplet is thus one option on one forward rate.

Where caps protect against high interest rates, *floors* are used against low interest rates. A floor is like a receiver swap, but exchange only for positive values. Analogously, a floorlet is the counterpart of a caplet.

In [Hul00] more background and practical information about these interest rate derivatives is given.

2.2 Interest models

Interest models can be split up into several classes. We focus on two classes of interest rate models: short-rate models and market models. We calculate the calibration risk of a profit-sharing option with the use of a short-rate model and a market model. These two kind of interest models are explained in this section.

2.2.1 Short-rate models

Short rate models assume that under some specified measure, e.g. the real world measure $\mathbb{P}$ or the risk-neutral measure $\mathbb{Q}$, the short rate satisfies the stochastic differential equation

$$dr(t) = \mu(t, r(t))dt + \sigma(t, r(t))dW(t), \tag{2.12}$$

where W is a Brownian motion under the specified measure. μ, the *drift*, is a real-valued function of t and r and σ, the *volatility*, is a positive real-valued function of t and r.
We define all short rate dynamics under the risk-neutral measure $\mathbb{Q}$. For a description of the term measure and the terms σ-algebra and (natural) filtration below, please refer to [Wil05].

The Vasicek model

The Vasicek model assumes time reversion of the interest rate. This model was introduced in 1977 by Vasicek in [Vas77]. The short rate follows the following dynamics:

$$dr(t) = a(\theta - r(t))dt + \sigma dW(t), \tag{2.13}$$

where $a > 0$ controls the speed of reversion to the constant mean $\theta > 0$. $\sigma > 0$ is the volatility parameter and $r(0) = r_0$, with $r_0 > 0$ a constant.

Integrating (2.13), we obtain, for each $s \leq t$,

$$r(t) = r(s)e^{-a(t-s)} + \theta(1 - e^{-a(t-s)}) + \sigma \int_s^t e^{-a(t-u)} dW(u). \tag{2.14}$$

Define $\mathcal{F}_s$ as the natural filtration of process r. Then we see that $r(t)$, conditional on the σ-algebra $\mathcal{F}_s$, is normally distributed with mean and variance given by

$$\mathbb{E}[r(t)|\mathcal{F}_s] = r(s)e^{-a(t-s)} + \theta(1 - e^{-a(t-s)}), \tag{2.15}$$

$$\mathrm{Var}[r(t)|\mathcal{F}_s] = \frac{\sigma^2}{2a}(1 - e^{-2a(t-s)}). \tag{2.16}$$

A major disadvantage of this model is that the short rate can attain negative values.

Nevertheless, the model is attractive since there are analytical formulas for bond prices and prices of European style options on bonds. That saves quite a lot computation time, because without these formulas prices must be retrieved by numerical procedures which need to run a lot of times to obtain plausible outcomes.

The price at time t of a pure discount bond that matures at time $T > t$ can be derived from the general pricing formula for a claim at time t with payoff H_T at time T:

$$H_t = \mathbb{E}\left[e^{-\int_t^T r(s)ds} H_T | \mathcal{F}_t\right], \tag{2.17}$$

where $\mathbb{E}_t$ denotes the t-conditional expectation under the risk-neutral measure. Thus

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

Vasicek shows in [Vas77] that from (2.18) one derives

$$P(t,T) = A(t,T)e^{-B(t,T)r(t)}, \tag{2.19}$$

where

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

and

$$\log A(t,T) = \frac{(B(t,T) - T + t)(a^2 b - \sigma^2/2)}{a^2} - \frac{\sigma^2 B(t,T)^2}{4a}.$$

Equation (2.19) now results in the following expression for the continuously compounded interest rate:

$$R(t,T) = -\frac{1}{\tau(t,T)} \log A(t,T) + \frac{1}{\tau(t,T)} B(t,T) r(t). \tag{2.20}$$

This shows us that the entire term structure can be determined as a function of $r(t)$. Short rate models for which $R(t,T)$ is an affine function of $r(t)$, like (2.20), are called *affine models*.

The Vasicek model is a one-factor model. There is one stochastic factor that drives all prices. So all yields to maturity are perfectly correlated, which might not be in line with the reality where yields to maturity could be highly correlated but are likely to be not perfectly correlated.

A way to prevent this perfect correlation is introducing another stochastic factor into the model, which results in the so-called two-factor Vasicek model.
Furthermore, an additional factor can explain more variations, like inverted, twisted or humped versions of the yield curve.

Two-factor Vasicek

Barrie and Hibbert describe (see e.g. [BH00]) the short rate dynamics under the risk neutral measure as follows:

$$\begin{aligned} dr(t) &= a_1(m(t) - r(t))dt + \sigma_1 dW_1(t), \\ dm(t) &= a_2(\mu - m(t)))dt + \sigma_2 dW_2(t). \end{aligned} \tag{2.21}$$

In this model $r(t)$ does not have a constant reversion level, but a stochastic time-reversing level which is driven by another Brownian motion. The two Brownian motions W_1 and W_2 are assumed to be independent.
This two-factor model can produce a larger set of yield curves than a single factor model.

Two-factor Vasicek still can value bonds analytically. The price of a discount bond at time t under two-factor Vasicek is given by

$$P(t,T) = \exp\left[A(T-t) - B_1(T-t)r(t) - B_2(T-t)m(t)\right], \tag{2.22}$$

where

$$\begin{aligned} B_1(s) &= \frac{1-e^{-a_1 s}}{a_1}, \\ B_2(s) &= \frac{a_1}{a_1 - a_2}\left(\frac{1-e^{a_1 s}}{a_2} - \frac{1-e^{a_1 s}}{a_1}\right), \\ A(s) &= (B_1(s) - s)(\mu - \frac{\sigma_1^2}{2a_1^2}) + B_2(s)\mu - \frac{\sigma_1^2 B_1(s)^2}{4a_1} \\ &\quad + \frac{\sigma_2^2}{2}\left[\frac{s}{a_2^2} - 2\frac{(B_2(s) + B_1(s))}{a_2^2} + \frac{1}{(a_1 - a_2)^2}\right. \\ &\quad - \frac{2a_1}{a_2(a_1 - a_2)^2}\frac{1-e^{-(a_1+a_2)s}}{a_1 + a2} \\ &\quad \left. + \frac{a_1^2}{a_2^2(a_1 - a_2)^2}\frac{1-e^{-2a_2 s}}{2a_2}\right]. \end{aligned}$$

For a derivation of this expression for $P(t,T)$, see Chapter 4 of [BM01].

Hull-White model

Besides the drawback of just one stochastic factor in the single-factor Vasicek model, there is another shortcoming of the model. The three parameters of the model are not enough to reproduce the initial term structure satisfactorily, so one cannot expect the model to reproduce other yield curves better.

A possibility is then to make the parameters time dependent, as was introduced by Hull and White in [HW90].

One version of the so-called *extended Vasicek model* that Hull and White consider is

$$dr(t) = (\theta(t) - ar(t))dt + \sigma dW(t) \tag{2.23}$$

where a and σ are constants and the drift or reversion-level θ is time dependent. This model is called the *Hull-White model*.

The reversion-level is chosen in such a way that the model perfectly fits the current term structure of interest rates. One obtains

$$\theta(t) = \frac{\partial f(0,t)}{\partial T} + af(0,t) + \frac{\sigma^2}{2a}\left(1 - e^{-2at}\right). \tag{2.24}$$

Integrating (2.23) yields

$$\begin{aligned} r(t) &= r(s)e^{-a(t-s)} + \int_s^t e^{-a(t-s)}\theta(u)du + \sigma \int_s^t e^{-a(t-u)}dW(u) \\ &= r(s)e^{-a(t-s)} + \alpha(t) - \alpha(s)e^{-a(t-s)} + \sigma \int_s^t e^{-a(t-u)}dW(u), \end{aligned} \tag{2.25}$$

where

$$\alpha(t) = f(0,t) + \frac{\sigma^2}{2a^2}(1 - e^{-at})^2. \tag{2.26}$$

The forward rates f are observed in the market.

Notice that (2.25) implies that $r(t)$ conditional on $\mathcal{F}_s$ is normal distributed with mean and variance given by

$$\mathbb{E}[r(t)|\mathcal{F}_s] = r(s)e^{-a(t-s)} + \alpha(t) - \alpha(s)e^{-a(t-s)} \tag{2.27}$$

$$\mathbb{E}[r(t)|\mathcal{F}_s] = \frac{\sigma^2}{2a^2}\left[1 - e^{-2a(t-s)}\right]. \tag{2.28}$$

This model possesses the same analytical tractability as Vasicek's. Bond prices are given by

$$P(t,T) = A(t,T)e^{-B(t,T)r(t)}, \tag{2.29}$$

where

$$B(t,T) = \frac{1 - e^{-a(T-t)}}{a} \tag{2.30}$$

and

$$\log A(t,T) = \log \frac{P^M(0,T)}{P^M(0,t)} - B(t,T)\frac{\partial P(0,t)}{\partial t} - \frac{1}{4a^3}\sigma^2(e^{-aT} - e^{-at})^2(e^{2at} - 1), \tag{2.31}$$

where $P^M(0,t)$ is the given market price at time t for a T-bond.
Consider a European call option with expiry T and strike K on a zero-coupon bond that matures at time S. As shown by Hull and White in [HW94], the Hull-White price of this product is

$$V_{\text{ZBC}}^{\text{HW}}(t,T,S,X) = P(t,S)\Phi(h) - XP(t,T)\Phi(h-\sigma_p), \tag{2.32}$$

where

$$\begin{aligned} \sigma_p &= \sigma\sqrt{\frac{1-e^{-2a(T-t)}}{2a}}B(T,S), \\ h &= \sigma_p = \frac{1}{\sigma_p}\log\frac{P(t,S)}{P(t,T)X} + \frac{\sigma_p}{2}. \end{aligned}$$

The price of the corresponding put option is thus

$$V_{\text{ZBP}}^{\text{HW}}(t,T,S,X) = XP(t,T)\Phi(-h+\sigma_p) - P(t,S)\Phi(-h).$$

Two-factor Hull White

Also the Hull-White model can be extended with an additional stochastic factor. The model keeps the property that there are analytical formulas for bond prices and prices of interest rate derivatives, including options on bonds, swaptions and caps. Hull and White set up their two factor model as follows:

$$dr(t) = [\theta(t) + u(t) - a_1 r(t)]dt + \sigma_1 dW_1(t), \tag{2.33}$$

with u the stochastic mean-reversion level that satisfies

$$du(t) = -a_2 dt + \sigma_2 dW_2(t), \tag{2.34}$$

where (W_1, W_2) is a two dimensional Brownian motion with instantaneous correlation $-1 \leq \rho \leq 1$ such that $dW_1(t)dW_2(t) = \rho dt$.
The additional factor makes it harder to derive formulas for interest rate options.

Things get easier when regarding another two-factor short rate model which can be rewritten into two-factor Hull-White. This specific model is the so-called *Two-Additive-Factor Gaussian model.*

Two-Additive-Factor Gaussian

Under the Two-Additive-Factor Gaussian model the short rare is assumed to satisfy

$$r(t) = x(t) + y(t) + \phi(t), \quad r(0) = r_0, \tag{2.35}$$

in the risk-neutral world. The processes $x(t)$ and $y(t)$ follow the dynamics

$$\begin{aligned} dx(t) &= -ax(t)dt + \sigma dW_1(t), \quad x(0) = 0, \\ dy(t) &= -by(t)dt + \eta dW_2(t), \quad y(0) = 0. \end{aligned} \tag{2.36}$$

where again W is a two-dimensional Brownian motion with instantaneous correlation ρ for which $-1 \leq \rho \leq 1$ and r_0, a, b, σ, η are positive constants.
The function ϕ is deterministic and is chosen as in two-factor Hull-White: to fit the initial term structure. Pricing interest rate derivatives is easier with this model because of its symmetry and the fact that x and y are not nested like u and r for two-factor Hull-White.
This model was first published by Brigo and Mercurio in 2001 in [BM01].

Connection between two-factor Hull-White and Two-Additive-Factor Gaussian

Remember that for two-factor Hull-White we had

$$dr(t) = [\theta(t) + u(t) - \bar{a}_1 r(t)]dt + \sigma_1 dW_1(t),$$

with u the stochastic mean-reversion level that satisfied

$$du(t) = -\bar{a}_2 dt + \sigma_2 dW_2(t), \tag{2.37}$$

and (W_1, W_2) a two dimensional Brownian motion with instantaneous correlation $\bar{\rho}$.

Assume that $\bar{a} \neq \bar{b}$. The analogy between two-factor Hull-White and Two-Additive-Factor Gaussian can now be shown by considering a new stochastic process

$$\chi(t) = r(t) + \delta u(t),$$

where $\delta = 1/(\bar{b} - \bar{a})$ with the assumption $\bar{a} > \bar{b}$ (the other case is analogous). The new stochastic process satisfies

$$\begin{aligned} d\chi(t) &= [\theta(t) + u(t) - \bar{a}r(t)]dt + \sigma_1 dW_1(t) - \delta\bar{b}u(t)dt + \delta\sigma_2 dW_2(t) \\ &= [\theta(t) + u(t) - \bar{a}\chi(t) + \bar{a}\delta u(t) - \bar{b}\delta u(t)]dt + \sigma_1 dW_i(t) + \delta\sigma_2 dW_2(t) \\ &= [\theta(t) - \bar{a}\chi(t)]dt + \sigma_3 dW_3(t), \end{aligned}$$

where

$$\sigma_3 = \sqrt{\sigma_1^2 + \frac{\sigma_2^2}{(\bar{a} - \bar{b})^2} + 2\bar{\rho}\frac{\sigma_1\sigma_2}{\bar{b} - \bar{a}}}$$

and

$$dW_3(t) = \frac{\sigma_1 dW_1(t) - \frac{\sigma_2}{(\bar{a}-\bar{b})^2} dW_2(t)}{\sigma_3}.$$

Furthermore, define

$$\psi(t) = \frac{u(t)}{\bar{a} - \bar{b}} = -\delta u(t),$$

then

$$\begin{aligned} d\psi(t) &= -\frac{\bar{b}}{\bar{a} - \bar{b}} u(t)dt + \frac{\sigma_2}{\bar{a} - \bar{b}} dW_2(t) \\ &= -\bar{b}\psi(t)dt + \sigma_4 dW_2(t), \end{aligned}$$

with

$$\sigma_4 = \frac{\sigma_2}{\bar{a} - \bar{b}}.$$

Finally we obtain that

$$r(t) = \tilde{\chi}(t) + \psi(t) + \varphi(t),$$

where

$$\begin{aligned} d\tilde{\chi}(t) &= -\bar{a}\tilde{\chi}(t)dt + \sigma_3 dW_3(t), \\ d\psi(t) &= -\bar{b}\psi(t)dt + \sigma_4 dW_2(t), \end{aligned}$$

and

$$\varphi(t) = r_0 e^{-\bar{a}t} + \int_0^t \theta(v) e^{-\bar{a}(t-v)} dv.$$

We thus see that we can rewrite the parameters of Two-Additive-Factor Gaussian into the ones of two-factor Hull-White by the relationship:

$$\begin{aligned} \bar{a} &= a, \\ \bar{b} &= b, \\ \sigma_1 &= \sqrt{\sigma^2 + \eta^2 + 2\rho\sigma\eta}, \\ \sigma_2 &= \eta(a - b), \\ \bar{\rho} &= \frac{\sigma\rho + \eta}{\sqrt{\sigma^2 + \eta^2 + 2\rho\sigma\eta}}, \\ \theta(t) &= \frac{d\varphi(t)}{dt} + a\varphi(t). \end{aligned}$$

This holds if $\bar{a} > \bar{b}$ which is equivalent to $a > b$. The Two-Factor-Additive Gaussian model is symmetric, so if $a < b$ change the roles of x and y to obtain the following parameters of two-factor Hull-White:

$$\begin{aligned} \bar{a} &= b, \\ \bar{b} &= a, \\ \sigma_1 &= \sqrt{\sigma^2 + \eta^2 + 2\rho\sigma\eta}, \end{aligned}$$

$$\sigma_2 = \eta(b-a),$$
$$\bar{\rho} = \frac{\eta\rho + \sigma}{\sqrt{\sigma^2 + \eta^2 + 2\rho\sigma\eta}},$$
$$\theta(t) = \frac{d\varphi(t)}{dt} + b\varphi(t).$$

Because Two-Additive-Factor Gaussian is a symmetric and tractable version of the two-factor Hull-White model, in the sequel we apply it for valuing the excess interest option.

2.2.2 Market Models

From a mathematical point of view, the short rate models are practical, as far as one can describe the whole behavior of the value of money. But from an economical point of view, short rate models are not practical. Short rates are not directly observable in the market. Valuation formulas for instruments are complicated for short rate models. Hence calibration of short rate models requires complicated numerical methods, as becomes clear in the sequel.
These drawbacks inspired some people not to model the short rate, but instead rates that are observable in the market. This resulted in so-called *market models.*
The first class of market models are Libor Market Models, introduced in 1997 by Miltersen, Sandmann and Sondermann and Brace, Gatarek and Musiela who described the forward Libor rates as lognormal processes. These models lead to Black's pricing formula for caps and floors.
Jamshidian developed also in 1997 a similar model for swap rates. This model leads to Black's formula for swaption prices.
Both models can be made to fit market prices perfectly. Only volatility parameters have to be estimated.
The focus in the sequel is on the Libor Market Model, which is popular with market practitioners and is implemented in the introduction mentioned Barrie and Hibbert's ESG.

Libor Market Model

The setup of the Libor Market Model is started by selecting a finite set of dates, the *tenor structure*

$$0 = T_0 < T_1 < T_2 < \ldots < T_N. \tag{2.38}$$

We denote the daycount fraction between two successive dates by $\delta_i = \tau(T_i, T_{i+1})$. In the sequel, we assume that $\delta_i = \delta$: all tenors have the same length.

With each tenor date T_n we can associate a bond that matures at that date, which price at time t is $P(t, T_n)$. Jamshidian defined the Libor Market Model by describing the bond price dynamics under the real world measure $\mathbb{P}$:

$$\frac{dP(t,T_n)}{P(t,T_n)} = \mu_n^P dt + \sigma_n^P dW^{\mathbb{P}}(t), \quad n = 1,\dots,N, \tag{2.39}$$

where $W^{\mathbb{P}}$ is a standard $\mathbb{P}$-Brownian motion and we assume it to be one or two dimensional. In the latter case, σ_n is a two-dimensional vector. The drift term μ^P and the volatility σ^P are time dependent and can depend on the bond price $P(t,T_n)$ itself.

We write $F_n(t)$ as the forward Libor rate for $[T_n, T_{n+1}]$ at time t. So

$$F_n(t) = \frac{1}{\delta}\left(\frac{P(t,T_n) - P(t,T_{n+1})}{P(t,T_{n+1})}\right), \quad n = 1,\dots,N-1. \tag{2.40}$$

Applying Ito's rule to (2.40), it follows that we have the following dynamics for the Libor rates under the real world measure:

$$\begin{aligned}
\frac{dF_n(t)}{F_n(t)} &= \mu_n(t)dt + \gamma_n(t)dW(t), \quad n = 1,\dots,N-1 \\
\gamma_n(t) &= \frac{P(t,T_n)}{P(t,T_n) - P(t,T_{n+1})}(\sigma_n^P(t) - \sigma_{n+1}^P(t)) \\
\mu_n(t) &= \frac{P(t,T_n)}{P(t,T_n) - P(t,T_{n+1})}(\mu_n^P(t) - \mu_{n+1}^P(t)) - \gamma_n(t)\sigma_{n+1}^P(t).
\end{aligned} \tag{2.41}$$

The *forward measure* $\mathbb{Q}^n$ is defined to be the equivalent martingale measure where the numeraire is the bond that matures at T_n. For a clear description of equivalent martingale measures and numeraires, see [Bjö98]. It is obvious that under this measure the forward Libor rate F_n is a martingale, since it can be seen as the relative value of a portfolio of zero coupon bonds

$$\frac{1}{\delta}(P(t,T_n) - P(t,T_{n+1})).$$

The measure $\mathbb{Q}^N$ is called the *terminal measure*. The Libor rate F_{N-1} is a martingale under this measure and has thus zero drift.

Jamshidian in [Jam97] shows that under $\mathbb{Q}^N$ we have the following dynamics for rate F_n:

$$dF_n(t) = F_n(t)\left(-\sum_{i=n+1}^{N-1}\frac{\delta F_i(t)\gamma_i(t)\gamma_n(t)}{1+\delta F_i(t)}dt + \gamma_n(t)dW^*(t)\right), \quad n = 1,\dots,N-1, \tag{2.42}$$

where W^* is a standard Wiener process under the terminal measure.

For practical reasons, we define an other measure: the *rolling forward risk neutral world*. In [BM01] this measure is called the Spot Libor Measure. It means that for this measure we can discount from time T_{n+1} to time T_n with the zero rate observed at time T_n for maturity T_{n+1}. This is in fact the measure that Barrie and Hibbert use for simulations in the ESG and to which they refer as *risk neutral world*. The numeraire is the *cash rollup*: the amount of cash at time t if one starts at $t = 0$ with €1 and invests in a bond maturing at T_1. Then at time T_1 one invests the proceeds in a bond maturing at T_2, etc. In this way,

$$\mathrm{CR}(t) = \prod_{j=0}^{m(t)-1} (1 + \delta F_j(T_j))P(t, T_{m(t)}). \tag{2.43}$$

As one easily sees, the cash rollup is the discrete version of the bank account introduced in section 2.1.2.

The dynamics for the rolling forward risk neutral world by Girsanov's theorem are as follows:

$$\frac{dF_n(t)}{F_n(t)} = \sum_{j=m(t)}^{n} \frac{\delta F_j(t)\gamma_j(t)\gamma_n(t)}{1 + \delta F_j(t)} dt + \gamma_n(t)dW(t), \tag{2.44}$$

where $m(t)$ is the index of the next reset date at time t: $m(t) = \inf\{m : t \leq T_m\}$. To make the model manageable, Barrie and Hibbert make the following assumptions about the forward rate volatilities $\gamma_n^q(t)$.

time-homogeneous: The volatilities of the forward depend only on the time to expiry of the forward rate, i.e. $\gamma_n^q(t)$ is of the form $\gamma^q(T_n - t)$.

piece-wise constant: The volatilities are constant over each interval $(T_i, T_{i+1}]$, i.e. $\gamma^q(T_n - t) = \Lambda^q_{n-m(t)+1}$ where Λ_n^q are step functions for $q = 1, 2$.

	Time				
Forward Rate	$t \in (0, T_1]$	$t \in (T_1, T_2]$	$t \in (T_2, T_3]$	...	$t \in (T_{N-1}, T_N]$
$F_1(t)$	Λ_1^q	expired	expired	...	expired
$F_2(t)$	Λ_2^q	Λ_1^q	expired	...	expired
$F_3(t)$	Λ_3^q	Λ_2^q	Λ_1^q	...	expired
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$F_N(t)$	Λ_N^q	Λ_{N-1}^q	Λ_{N-2}^q	...	Λ_1^q

mean reversion: Finally, with this step functions there are still many parameter choices available. So it is assumed that the volatility structure has a functional form, namely

$$\Lambda_n^1 = \sqrt{1-\rho^2}\sigma_1 e^{-\alpha_1 n\delta}, \tag{2.45}$$
$$\Lambda_n^2 = \sigma_2 e^{-\alpha_2 n\delta} + \rho\sigma_1 e^{-\alpha_1 n\delta}, \tag{2.46}$$

where $\alpha_1, \alpha_2, \sigma_1$ and σ_2 are positive constants and $-1 \leq \rho \leq 1$.

With these assumptions, we just need to specify five parameter values and put in N initial forward rates to simulate other forward rates.

We use these parametrization of the volatilities also in the Libor Market Model for pricing the embedded option in the interest contract. With this configuration of the model, calibration boils down to finding the 'appropriate' values for $\alpha_1, \alpha_2, \sigma_1, \sigma_2$ and ρ.

2.3 Pricing swaptions and caps

2.3.1 Black's model

European options are in the market usually priced with Black's model. The price at $t = 0$ of a European call option on a variable with value V and forward price at time t F_t of a contract with maturity T, with strike K and maturity T is given by

$$P(0,T)\left(F_0\Phi(d_1) - K\Phi(d_2)\right) \tag{2.47}$$

where

$$\begin{aligned} d_1 &= \frac{\log(F_0/K) + \sigma^2 T/2}{\sigma\sqrt{T}}, \\ d_2 &= \frac{\log(F_0/K) - \sigma^2 T/2}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T}. \end{aligned}$$

The σ is the volatility of the forward price F.

This model captures the assumptions that

- the value of the asset at time t, V_t, has a log-normal distribution with standard deviation of $\log V_t$ equal to $\sigma\sqrt{t}$;
- the expected value of V_t is F_0.

Black's model is a variant of the Black-Scholes pricing formula but with forward prices in stead of spot prices.

At $t = 0$ a payer swaption is a call option on a payer swap. In the market these are also priced under Black's model as

$$V_{\text{PS}}(0, N, K) = N(S\Phi(d_1) - K\Phi(d_2)), \tag{2.48}$$

where

$$\begin{aligned} d_1 &= \frac{\log(S/K) + \sigma^2 T_0/2}{\sigma\sqrt{T_0}}, \\ d_2 &= d_1 - \sigma\sqrt{T_0}. \end{aligned}$$

In this case, σ is a volatility parameter which is usually quoted in the market, rather than the price itself. There is a one-to-one relation between the implied volatility and the price, hence swaption prices are retrieved from the volatility data.

Analogously, the price of a receiver swaption is

$$V_{\mathrm{RS}}(0, N, K) = N(-S\Phi(-d_1) + K\Phi(-d_2)).$$

Swaptions for which the maturity and tenor are multiples of one year are denoted as $x \times y$-swaptions where x is the time to maturity and y equals the length of the underlying swap. Under the former notation, $x = \tau(0, T_0) \approx T_0$ and $y = \tau(T_0, T_n) \approx T_n - T_0$.

A cap is under Black's formula priced as

$$V_{\mathrm{cap}}(0, N, K) = N \sum_{i=1}^{n} P(0, T_i)\tau_i \left(F_i\Phi(d_{1i}) - K\Phi(d_{2i})\right), \tag{2.49}$$

where

$$\begin{aligned} d_{1i} &= \frac{\log(F_i/K) + \sigma^2 T_{i-1}}{\sigma\sqrt{T_{i-1}}} \\ d_{2i} &= \frac{\log(F_i/K) - \sigma^2 T_{i-1}}{\sigma\sqrt{T_{i-1}}} = d_{1i} - \sigma\sqrt{T_{i-1}} \end{aligned}$$

and F_i is the forward rate at time $t = 0$, which is omitted for clarity, for the period $[T_i, T_{i+1}]$.

Market values of caps are commonly quoted by the implied volatilities for caps. That is the σ in above expressions. We have the following Black pricing formula for floors:

$$V_{\mathrm{floor}}(0, N, K) = N \sum_{i=1}^{n} P(0, T_i)\tau_i \left(-F_i\Phi(d_{1i}) + K\Phi(d_{2i})\right), \tag{2.50}$$

where d_{1i} and d_{2i} are same as for the cap formula.

2.3.2 Two-Additive-Factor Gaussian

The price at time t of a zero-coupon bond maturing at T is given by

$$P(t,T) = \frac{P^M(0,T)}{P^M(0,t)} \exp\{\mathcal{A}(t,T)\} \tag{2.51}$$

$$\mathcal{A}(t,T) := \frac{1}{2}(V(t,T) - V(0,T) + V(0,t)) - \frac{1-e^{-a(T-t)}}{a}x(t) - \frac{1-e^{-b(T-t)}}{b}y(t). \tag{2.52}$$

In this formula, $V(t,T)$ is the variance of the random variable $\int_t^T x(u) + y(u)du$ conditional on $\mathcal{F}_t$. It has a closed form, dependent on the parameters of Two-Additive-Factor Gaussian, t and T. The formula of $V(t,T)$ can be found in the appendix.
A derivation of this expression is provided by Brigo and Mercurio in [BM01].

A cap can be viewed as a portfolio of several European put options on the Libor rate. So first consider the pricing of a European option on a zero-coupon bond. The value of a European call option with maturity T and strike K on an S-bond is given by

$$\begin{aligned} V_{\text{ZBO}}^{\text{G2}} &= \omega N P(t,S)\Phi\left(\omega\frac{\log\frac{NP(t,S)}{KP(t,T)}}{\Sigma(t,T,S)} + \frac{\omega}{2}\Sigma(t,T,S)\right) \\ &\quad -\omega P(t,T)K\Phi\left(\omega\frac{\log\frac{NP(t,S)}{KP(t,T)}}{\Sigma(t,T,S)} - \frac{\omega}{2}\Sigma(t,T,S)\right), \end{aligned} \tag{2.53}$$

where $\Sigma(t,T,S)$ is the standard deviation of $\log P(T,S)$ conditional on $\mathcal{F}_t$ under the forward measure $\mathbb{Q}^T$. For the expression of $\Sigma(t,T,S)$ refer to the appendix. For a call option, $\omega = 1$ and a put option, $\omega = -1$.

Consider a cap with reset dates $\{T_0, \ldots, T_{n-1}\}$ and payment dates $\{T_1, \ldots, T_n\}$, strike rate X and nominal amount N. Denote by τ_i the year fraction between T_{i-1} and T_i. Then,

$$\begin{aligned} V_{\text{cap}}^{\text{G2}} &= \sum_{i=1}^{n}\left[-N(1+X\tau_i)P(t,T_i)\Phi\left(\frac{\log\frac{P(t,T_{i-1})}{(1+X\tau_i)P(t,T_i)}}{\Sigma(t,T_{i-1},T_i)} - \frac{1}{2}\Sigma(t,T_{i-1},T_i)\right)\right. \\ &\quad \left. +P(t,T_{i-1})N\Phi\left(\frac{\log\frac{P(t,T_{i-1})}{(1+X\tau_i)P(t,T_i)}}{\Sigma(t,T_{i-1},T_i)} + \frac{1}{2}\Sigma(t,T_{i-1},T_i)\right)\right]. \end{aligned} \tag{2.54}$$

Pricing European swaptions is much more complicated. There is an analytic

formula for the price of a European swaption. Evaluating that formula however requires the computation of a one-dimensional integral which only can be evaluated numerically. But that is still faster than valuation by building a tree or performing simulations.
The formula for the price of a payer swaption is as follows:

$$V_{PS}^{G2} = P(0,T)\int_{-\infty}^{\infty} \frac{e^{-\frac{1}{2}\left(\frac{x-\mu_x}{\sigma_x}\right)^2}}{\sigma_x\sqrt{2\pi}} \left[\Phi(-h_1(x)) - \sum_{i=1}^{n} \lambda_i(x) e^{\kappa_i(x)} \Phi(-h_2(x))\right] dx. \tag{2.55}$$

The definitions of the functions $h_1(x), h_2(x), \lambda_i(x), \kappa_i(x)$ and the parameters $\mu_x, \mu_y, \sigma_x, \sigma_y, \rho_{xy}$ and $\bar{y}$ are in the appendix. A complete proof for this expression of the value of a payer swaption is recorded in the appendix of [BM01].

2.3.3 Libor Market Model

Market values of caps are quoted in implied volatilities. These volatilities are input for Black's pricing formula. But actually, each caplet has its own implied volatility. The pricing formula for a cap then becomes

$$V_{\text{cap}}(0, N, K) = N \sum_{i=1}^{n} P(0, T_i)\delta\left(F_i(0)\Phi(d_{1i}) - K\Phi(d_{2i})\right) \tag{2.56}$$

with

$$\begin{aligned} d_{1i} &= \frac{\log(F_i(0)/K) + \sigma_i^2 T_{i-1}}{\sigma_i\sqrt{T_{i-1}}}, \\ d_{2i} &= \frac{\log(F_i(0)/K) - \sigma_i^2 T_{i-1}}{\sigma_i\sqrt{T_{i-1}}}. \end{aligned} \tag{2.57}$$

In Black's model, σ_i represents the volatility of the forward rate $F_i(t)$. In the Libor Market Model, forward rates are also assumed to be log-normally distributed. The volatility of forward rate $F_i(t)$ equals

$$\sqrt{\frac{1}{i}\sum_{j=1}^{i}\left\{{\Lambda_j^1}^2 + {\Lambda_j^2}^2\right\}}.$$

Plugging this expression in Black's formula for caplets, i.e. one term of (2.49), as the implied volatility results in the Libor Market Model's caplet price. Summing this over all the caplets that form a cap yields the Libor Market Model's cap price.

For swaptions an analogy does not exist between the Libor Market Model and Black's model. Black's model assumes log-normal swap rates. It can be shown that under Libor Market Model swap rates are not log-normal. There is even not an analytic closed-from formula for the Libor Market Model price of a European swaption, for details see [dJDP01]. Under the Libor Market Model, swaptions can be valued by using approximation formulas and by numerical algorithms like binomial or trinomial trees or simulation. Numerical methods for pricing swaptions are time consuming. For calibrating the Libor Market Model to swaptions, we will need to evaluate the Libor Market Model's price of a swaption numerous times. So a quick way of pricing swaptions is desirable.

The approximation formulas are closed-form and easy to evaluate. The approximation should however be accurate for calibration purposes. Hull and White introduced in [HW99] the following approximation.

This relationship holds between bond prices and forward rates:

$$\frac{P(t,T_n)}{P(t,T_k)} = \prod_{j=k}^{n-1} \frac{1}{1+\delta F_j(t)}, \tag{2.58}$$

for $n \geq k+1$. The swap rate of (2.10), where $T_m = 0$, can be written as

$$S_{0,n}(0) = \frac{1 - \prod_{j=0}^{n-1} \frac{1}{1+\delta F_j(0)}}{\sum_{i=0}^{n-1} \delta \prod_{j=0}^{i} \frac{1}{1+\delta F_j(0)}},$$

or, equivalently as

$$S_{0,n}(0) = \frac{\prod_{j=0}^{n-1} (1+\delta F_j(0)) - 1}{\sum_{i=0}^{n-1} \delta \prod_{j=i+1}^{n-1} (1+\delta F_j(0))}.$$

Taking the logarithm of both sides results in

$$\log S_{0,n}(0) = \log \left\{ \prod_{j=0}^{n-1} (1+\delta F_j(0)) - 1 \right\} - \log \left\{ \sum_{i=0}^{n-1} \delta \prod_{j=i+1}^{n-1} (1+\delta F_j(0)) \right\}.$$

From this follows

$$\frac{1}{S_{0,n}(0)} \frac{\partial S(0)_{0,n}}{\partial F_k(0)} = \frac{\delta \xi_k(0)}{1+\delta F_k(0)},$$

where

$$\xi_k(t) = \frac{\prod_{j=0}^{n-1}(1+\delta F_j(t))}{\prod_{j=0}^{n-1}(1+\delta F_j(t)) - 1} - \frac{\sum_{i=0}^{k-1} \delta \prod_{j=i+1}^{n-1}(1+\delta F_j(t))}{\sum_{i=0}^{n-1} \delta \prod_{j=i+1}^{n-1}(1+\delta F_j(t))}.$$

From Ito's lemma we obtain that the qth component (where $q = 1, 2$) of the volatility of $S(t)_{m,n}$ equals

$$\sum_{k=m}^{n-1} \frac{1}{S_{m,n}(t)} \frac{\partial S_{m,n}(t)}{\partial F_k(t)} \gamma_k^q(t) F_k(t),$$

where $\gamma_k^q(t)$ is the qth component of the volatility of $F_k(t)$.
Above expression can be rewritten into

$$\sum_{k=m}^{n-1} \frac{\delta \gamma_k^q(t) F_k(t) \xi_k(t)}{1 + \delta F_k(t)}.$$

We have assumed in section 2.2.2 that $\gamma_k^q(t) = \Lambda_{k-m(t)+1}^q$. So the variance rate of $S_{m,n}(t)$ is

$$\left(\sum_{k=m}^{n-1} \frac{\delta \Lambda_{k-m(t)+1}^1 F_k(t) \xi_k(t)}{1 + \delta F_k(t)} \right)^2 + \left(\sum_{k=m}^{n-1} \frac{\delta \Lambda_{k-m(t)+1}^2 F_k(t) \xi_k(t)}{1 + \delta F_k(t)} \right)^2. \tag{2.59}$$

This expression for the variance rate of the swap rate is stochastic. The forward rates $F_k(t)$ are log-normally distributed under the Libor Market Model, this formula shows us that swap rates are not log-normally distributed.
To obtain an approximation, assume that $F_k(t) = F_k(0)$, i.e. the forward rates are constant. This assumption implies that the volatility of $S_{m,n}(t)$ is constant within each period between any T_j and T_{j+1} for $j \leq m-1$. These intervals are so-called accrual periods.

The average variance rate of $S_{m,n}(t)$ between time 0 and time T_m is

$$\frac{1}{T_m} \sum_{j=0}^{m-1} \left[\delta \left(\sum_{k=m}^{n-1} \frac{\delta \Lambda_{k-j}^1 F_k(t) \xi_k(t)}{1 + \delta F_k(t)} \right)^2 + \delta \left(\sum_{k=m}^{n-1} \frac{\delta \Lambda_{k-j}^2 F_k(t) \xi_k(t)}{1 + \delta F_k(t)} \right)^2 \right].$$

The volatility that is to be used as input for Black's model to price the swaption is thus

$$\sqrt{\frac{\delta}{T_m} \sum_{j=0}^{m-1} \left[\left(\sum_{k=m}^{n-1} \frac{\delta \Lambda_{k-j}^1 F_k(t) \xi_k(t)}{1 + \delta F_k(t)} \right)^2 + \left(\sum_{k=m}^{n-1} \frac{\delta \Lambda_{k-j}^2 F_k(t) \xi_k(t)}{1 + \delta F_k(t)} \right)^2 \right]}. \tag{2.60}$$

The accuracy of this approximation is investigated by Hull and White. They calibrated the Libor Market Model, with the same configuration for the forward rate volatilities but with three factors, to a zero curve with an average of 5% and to swaption implied volatilities of about 20% and the mismatch between the approximation and the value obtained by Monte Carlo simulation amounted to be less than 0.1% for a 5×5-swaption.

We have performed the same check; for a 5×5-swaption the mismatch between the swaption price based on simulation and based on the approximation was 0.3% in price.

There are other approximation formulas available. Brigo and Mercurio mention these in [BM01]. Upon their tests they conclude that the differences between the approximations is “practically negligible in most situations”.

Chapter 3

Calibration

The interest rate models treated in chapter two cannot be used without first determining the values of the parameters in the model. For example, the function $\theta(t)$ and the volatility parameters a and σ in the single-factor Hull-White model need to be chosen before the model is ready for use. The art of finding the parameter values is called *calibrating*.

Interest rate models are used in several areas. Two main utilities can be distinguished, first analyzing future behavior of interest rates and second valuing interest rate derivatives.

In the first case, one is interested in the dynamics of interest rates that drive risks involving investments as for Asset Liability Management. For this objective, the *real world dynamics* are significant: observable movements of interest rates.

The second field in which interest rate models are used is valuing interest rate derivatives. For example, one wants to compute a fair price of some swaption for which the price is not available. In this situation, the *risk-neutral dynamics* of the interest rates are studied.

A model can be very suitable for a certain purpose, but without a good calibration the model will return poor results which might not be realistic. So good and efficient calibration is as important as the choice of the model.

Both applications of interest rate models require different calibration techniques. For the first area of applications the calibration consists of adapting the interest rate model to the current yield curve or outside insights from experts. The volatility parameters are based on historical data, for some applications even data that are decades old. From these data, experts retrieve the volatilities with mathematical methods and then analyze the obtained values. Outcomes that do not seem reasonable are adjusted to match the opinion of the experts. This way of calibrating is called *best-estimate*. Calibrations of this type are usually not frequently executed, for example quarterly. The interest rate curve is not that volatile that more calibrations are needed and the volatility parameters would not change a lot, since a large buffer of historical data is part of the calibration input.

The second type of applications, pricing financial derivatives, requires a different approach. In principle, only current market data are needed. The goal is valuing a derivative that is consistent with other quotes in the market, hence historical data are not needed. The parameters that define the drift of the process, as $\theta(t)$ for Hull-White and the initial forward rates for the Libor Market Model, are computed from given interest rates and zero-coupon bond prices. This part is often not called calibration, it is usually referred to as fitting the initial term structure or, for the Libor Market Model, fitting the initial forward rates.

After this, the volatility parameters of the model are obtained by fitting the model prices of a set of financial instruments to the market values of those *calibration instruments*. With the so obtained parameters, we expect the model to accurately price derivatives that are similar to the market values of the calibration instruments.
For this calibration technique in principle no input of experts is used. But if the resulting parameter values are not sensible, one might wonder whether the market data is internally consistent or whether the model is appropriate for valuing the calibration instruments.

This latter approach of calibration used for pricing is the one we concentrate on. It is applied for pricing the excess interest option.

To get insight in the calibration process, we take a look at an example of such a calibration for the relatively simple one-factor Hull-White model.

Intermezzo: calibration of 1FHW to swaptions

Consider the one-factor Hull-White model (2.23) with constant mean reversion and volatility parameters. The model price of a Bermudan option on a swap depends on the calibration technique that is used, since different calibration methods could lead to different parameter values.
The Bermudan swaption we investigate is an option on the Euro market on a receiver swap with yearly exercise dates from 9 years to 28 years. The underlying swap matures in 29 years. The value date is 29 December 2006.

To obtain sensible parameter values for pricing the Bermudan swaption, the model is calibrated to market values of $9 \times 20, 10 \times 19, \ldots, 28 \times 1$ swaptions. Each option maturity on the swaps corresponds to the possible exercise dates of the Bermudan swaption and the underlying swaps have that length for which they end in 29 years.

The input for the calibration process are zero-coupon bond prices, also called discount factors, and the swaption prices for the mentioned swaptions, both in the Euro market.

Euro market data 29-12-2006

Swaptions	Black vols	Price
9×20	10.90 %	5.42
10×19	10.70 %	5.15
11×18	10.61 %	4.91
12×17	10.52 %	4.64
13×16	10.41 %	4.36
14×15	10.30 %	4.05
15×14	10.28 %	3.78
16×13	10.35 %	3.54
17×12	10.42 %	3.28
18×11	10.47 %	3.00
19×10	10.52 %	2.63
20×9	10.50 %	2.43
21×8	10.42 %	2.12
22×7	10.34 %	1.83
23×6	10.26 %	1.54
24×5	10.20 %	1.27
25×4	10.20 %	1.00
26×3	10.50 %	0.76
27×2	10.70 %	0.51
28×1	10.90 %	0.26

Time	Discount factor
1	0.948955838
2	0.902678516
3	0.859952637
4	0.818480359
5	0.776782639
6	0.737893683
7	0.700899051
8	0.666548685
9	0.632215109
10	0.599160456
11	0.567042652
12	0.536633647
13	0.508229798
14	0.4813382
15	0.455868378
20	0.347129297
25	0.266248855
30	0.20551631

To measure the quality of the estimated parameter values we define a *measure of fit*. The parameter values found depend on the choice of this measure of fit. For this example we take as measure of fit for each instrument the difference between the quoted market price and the model price and finally sum up the squares of the differences.

A European receiver swaption is actually a call option on a coupon-bearing bond. The Hull-White model has analytic formulas for options on zero-coupon bonds. There is a way to obtain formulas for options on coupon-bearing bonds. This idea has been published by Jamshidian in [Jam89]. Write $P_{\mathrm{HW}}(t, T, r(t))$ for the analytical bond price under the Hull-White model at time t for short rate value at time t, $r(t)$ and maturity T.

Consider a coupon-bearing bond paying $c_1, \ldots, c_n$ at times $T_1, \ldots, T_n$. Let T be some time $T \leq T_1$. The price of the coupon-bearing bond is obviously given by

$$\sum_{i=1}^{n} c_i P(T, T_i) = \sum_{i=1}^{n} c_i P_{\text{HW}}(T, T_i, r(T)).$$

The payoff of a European put option on this coupon-bearing bond with strike price K is

$$\left[K - \sum_{i=1}^{n} c_i P_{\text{HW}}(T, T_i, r(T)) \right]^+ . \tag{3.1}$$

Jamshidian came up with a trick to write this positive part of a sum as a sum of positive parts, which then enables us to apply Hull-White's formula for options on zero-coupon bonds. Find the value r^* for which holds that

$$\sum_{i=1}^{n} c_i P_{\text{HW}}(T, T_i, r^*) = K, \tag{3.2}$$

and rewrite the payoff (3.1) as

$$\left[\sum_{i=1}^{n} c_i (P_{\text{HW}}(T, T_i, r^*) - P_{\text{HW}}(T, T_i, r(T))) \right]^+ .$$

Under the sufficient condition

$$\frac{\partial P_{\text{HW}}(t, s, r)}{\partial r} < 0 \qquad \text{for all } 0 < t < s, \tag{3.3}$$

we can rewrite the payoff as

$$\sum_{i=1}^{n} c_i [P_{\text{HW}}(T, T_i, r^*) - P_{\text{HW}}(T, T_i, r(T))]^+ . \tag{3.4}$$

If P_{HW} satisfies (3.3), we have for all $i \in \{1, \ldots, n\}$ that

$$r^* < r(T) \Leftrightarrow P_{\text{HW}}(T, T_i, r^*) > P_{\text{HW}}(T, T_i, r(T)).$$

A payoff (3.1) equal to zero implies that there exists a j for which $P_{\text{HW}}(T, T_j, r^*) - P_{\text{HW}}(T, T_j, r(T)) \leq 0$. From this fact and by the condition, we conclude that $r^* \leq r(T)$. Again by the condition, we now have for all i that $P_{\text{HW}}(T, T_i, r^*) - P_{\text{HW}}(T, T_i, r(T)) \leq 0$.
The converse holds as well, if the payoff is positive we have for all i the inequality $P_{\text{HW}}(T, T_i, r^*) - P_{\text{HW}}(T, T_i, r(T)) > 0$.

From these observations it is shown that under the condition (3.3) we can write the payoff as (3.4).

That the sufficient conditions hold for the Hull-White model is intuitively clear. For fixed parameters in the Hull-White model, the Hull-White price for a zero-coupon bond only depends on the initial forward rate. The higher that initial rate is, the faster the process of discounting goes and thus the lower the bond prices are.
From a mathematical point of view it is clear as well. Recall

$$P_{\text{HW}}(t,s,r) = A(t,s)e^{-B(t,s)r}. \tag{3.5}$$

Then

$$\frac{\partial P_{\text{HW}}(t,s,r)}{\partial r} = -B(t,s)A(t,s)e^{-B(t,s)r} < 0, \tag{3.6}$$

since $A(t,s) > 0$ and $B(t,s) > 0$ for all $0 \leq t < s$.

Concluding, we can price a swaption under the Hull-White model since pricing of an option on a coupon-bearing bond can be accomplished by pricing a portfolio of options on zero-coupon bonds with strike $X_i = P_{\text{HW}}(T, T_i, r^*)$.
The Hull-White price of a receiver swaption is thus

$$V_{\text{RS}}^{\text{HW}} = N \sum_{i=1}^{n} c_i V_{\text{ZBC}}^{\text{HW}}(0, T, T_i, X_i), \tag{3.7}$$

where $c_i = X\tau_i$ for $1 \leq i < n$, $c_n = 1 + X\tau_n$ and τ_i is the year fraction between T_{i-1} and T_i.

Define $t = 0$ for 29 December 2006. The underlying swaps have annual payments, so approximately $\tau_i = 1$ for all i.

The quality of the parameters is measured by summing up the squares of the differences between market prices and model prices. A good fit corresponds to a low measure of fit. So the following expression has to be minimized over the parameter values:

$$\chi^2(\mathbf{a}) = \sum_{j=1}^{20} (p_j - V_{\text{RS}}^{\text{HW}}(\mathbf{a}, x_j))^2, \tag{3.8}$$

where $\mathbf{a}$ is a vector consisting of the two volatility parameter values a and σ that determine the model behavior. We say that $\mathbf{a}$ is an element of the *parameter space* $[0,\infty) \times [0,\infty)$. The variable x_j represents the input of the function for the specific instrument j: the expiry of the option, tenor of the underlying swap.

The next step is finding the minimum of χ^2. This is done by a minimization algorithm. For this case we choose the Levenberg-Marquardt method. This optimization method searches for a local minimum. The method is explained in section 3.3.1.

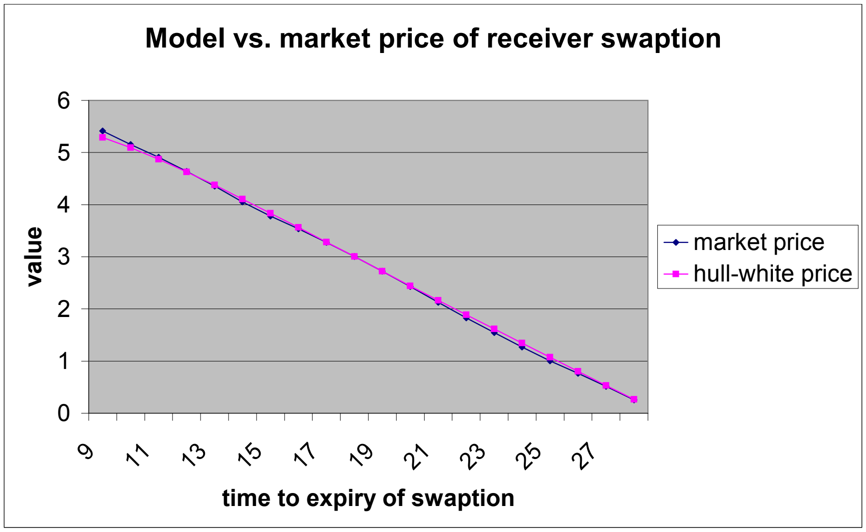
Levenberg-Marquardt gives the following values

$$\begin{aligned} a &= 0.044528426, \\ \sigma &= 0.009387939. \end{aligned} \tag{3.9}$$

The model prices and pricing errors with these parameters are:

Swaptions	**price**	**HW price**	**difference**	**squared diff.**
9X20	5.415133716	5.2891	0.126034	0.015885
10X19	5.151568901	5.09226	0.059309	0.003518
11X18	4.907258036	4.869061	0.038197	0.001459
12X17	4.640211622	4.629915	0.010296	0.000106
13X16	4.356012626	4.375548	-0.01953	0.000382
14X15	4.053499457	4.107082	-0.05358	0.002871
15X14	3.78389946	3.836941	-0.05304	0.002813
16X13	3.536505377	3.561834	-0.02533	0.000642
17X12	3.276482193	3.284038	-0.00756	5.71E-05
18X11	3.00409832	3.002993	0.001105	1.22E-06
19X10	2.725324089	2.722439	0.002885	8.32E-06
20X9	2.428697545	2.443549	-0.01485	0.000221
21X8	2.126463084	2.16704	-0.04058	0.001647
22X7	1.828705692	1.889563	-0.06086	0.003704
23X6	1.544210041	1.618253	-0.07404	0.005482
24X5	1.267072391	1.34538	-0.07831	0.006132
25X4	1.003355917	1.073848	-0.07049	0.004969
26X3	0.76479559	0.802564	-0.03777	0.001426
27X2	0.513580523	0.533997	-0.02042	0.000417
28X1	0.257647104	0.265979	-0.00833	6.94E-05
			$\chi^2 =$	0.051808

In a chart:



Valuing the Bermudan swaption can now be done by simulating the Hull-White model by the Monte Carlo techniques with the parameter values in (3.9).
The function θ from (2.23) does not need to be found explicitly since at each iteration time step t_i of one Monte Carlo simulation run, just the curve $T \mapsto P(t,T)$ is important and in the expression for $P(t,T)$ no $\theta(t)$ is included.

The actual valuing of a Bermudan swaption using these calibration parameters and one-factor Hull-White is outside the scope of this thesis and requires additional theorems and derivations.

As we see from the intermezzo, several considerations are made for the calibration process. Given the interest rate model and the derivative to price, we have to consider several possible choices. Namely,

1. the selection of calibration instruments;
2. the function to minimize, the measure of fit;
3. the minimization algorithm.

One can imagine that each of these choices would have impact on the parameter outcomes of the calibration and thus on the price of the derivative that is obtained with the calibrated model.

In the following sections, we investigate these choices. In chapter four, the impact of the choices on the price of an embedded option value in an insurance contract is investigated.

3.1 Selecting the calibration instruments

We deploy two interest rate models for valuing a profit sharing contract. The volatility parameters and the correlation parameter are to be estimated for both models. Calibrating should be an efficient process, there should be balance between the quality and computation time. If that is the case, calibration of the model can be performed frequently, which is preferable in the case of financial reporting. Usually, internal financial reporting happens on a quarterly, or even on a monthly basis in some insurance companies.
One reason is to give the management insight in the value of the liabilities of the company. Another motivation is the possibility to obtain information about the sensitivity of the liabilities to market factors, like the interest rate and volatility of interest, and control the net position of the liabilities by hedging them with a proper asset strategy.
So we need calibration instruments that give us reliable and sufficient information, instruments that have a recent market price and the model should be able to value the calibration instruments quickly.

If the model is used for pricing interest rate derivatives, then the model must be calibrated to interest rate derivatives that have a price that is quoted in the market. Such derivatives could be interest rate futures, interest rate forwards, interest rate swaps, interest rate swaptions, interest caps and floors and more exotic interest rate derivatives.
The value of interest rate forwards, futures and swaps depend on the interest yield curve on valuation moment. The price gives no information about volatility of interest rates. For example, in the price formula for swaps (2.8) only zero-coupon bond prices are input and no further input is required. Market expectations about future interest volatility should be retrieved to obtain a market consistent price for the derivative that is priced with the calibrated model.

Swaptions, caps and floors clearly contain volatility information. Swaptions are quoted with their Black implied volatilities, that can be interpreted as market expectations of the forward rate underlying the swaption. Usually, data providers daily publish an implied volatility matrix with swaption implied volatilities. In the rows the option maturities are given (in years) and in the columns, the underlying swap term is shown. See chapter four for the swaption matrix used in the calibrations of the interest rate models.

Caps are also often quoted with their Black implied volatilities, as we saw in (2.49). This volatility is not directly linked to a particular interest rate. But clearly, underlying interest rate volatilities play a large role in the price of a cap.

Caps are quoted in one volatility, but the underlying caplets have their own volatilities as well. These caplet Black implied volatilities are seldom quoted in the market. They can be obtained by bootstrapping from several cap quotes.
A caplet for the period $[T_i, T_{i+1}]$ has as Black's pricing formula:

$$V_{\text{caplet}}(0, N, K) = NP(0, T_i)\tau_i \left(F_i\Phi(d_{1i}) - K\Phi(d_{2i})\right),$$

where,

$$\begin{aligned} d_{1i} &= \frac{\log(F_i/K) + \sigma_i^2 T_{i-1}}{\sigma_i\sqrt{T_{i-1}}} \\ d_{2i} &= \frac{\log(F_i/K) - \sigma_i^2 T_{i-1}}{\sigma_i\sqrt{T_{i-1}}} = d_{1i} - \sigma\sqrt{T_{i-1}}. \end{aligned}$$

In this case, σ_i is in fact Black's volatility of the forward rate F_i. Compare this with formula (2.49). Since we have no caplet implied volatilities at our disposal, we should use bootstrapping. However, that would require a lot of time, analyses and could even add more subjectivity to the market values. Moreover, caplet volatilities correspond with interest rates for short term, usually one year, whereas profit sharing contracts cover much longer periods.
Hence, we do only use cap prices and not cap implied volatilities for the calibration.

Cap quotes are also given in a matrix. On the vertical axis, the rows, the terms of the cap are set out. In the horizontal direction, the cap rates (strike rates) are set out.

In chapter four, a cap matrix is given. This cap matrix does supply quotes for in-the-money caps as well for out-of-the-money caps. The swaption matrix only has at-the-money quotes. This extra dimension in the sense of additional information can be taken into account in the calibration process or can be omitted. This is also investigated in this thesis: the effect of including in- and out-of-the-money caps as calibration instruments.

The quotes of interest rate calibration instruments at valuation date should be actual. Market data providers generate the market prices based on quotes of parties in the market and the prices at which the products were traded. If a certain derivative is not traded frequently, its quote distributed by data providers might be stale. For example, a 40×50-swaption is certainly not a commonly traded product and its quote also cannot be based on several prices of recently traded products. To overcome this, quotes for these products can be neglected, or the measure of fit gives less significance to the quotes that might be stale.

Finally, the calibration instruments should not be too hard to value by the models. If the model price is only obtained by time-consuming numerical procedures, the minimization algorithm that searches the optimal parameters becomes too slow. In a minimization algorithm, the function that gives the model price of a calibration instrument, is evaluated numerous times.

In our analysis on the calibration risk, we calibrate the two interest models to swaptions and to caps. These interest rate derivatives fulfill all of the above preferences and there are techniques that make it relatively easy to price these instruments with the Two-Additive-Factor Gaussian model as well as the Libor Market Model.

3.2 Choice of measure of fit

Finding the parameters of an interest rate model based on prices of financial instruments quoted in the market is done by minimizing a function χ^2. This function gives a measure of the error between observed market values y_i of n financial instruments and the values of the instruments given by the interest rate model, $y(x_i; \mathbf{y})$, where x_i are specifications of the i-th instrument to be put in the pricing formula of the model and $\mathbf{a}$ is a vector of parameters.

These functions can be of several forms. One could minimize absolute differences between market values and model values. But also minimizing the squares between market values and model values is possible besides other expressions of market values and model values. Commonly used is the least squares method. Then the function is of the form

$$\chi^2(\mathbf{a}) = \sum_{i=1}^{n} w_i \left(y_i - y(x_i; \mathbf{a})\right)^2, \tag{3.10}$$

where $w_i > 0$ are weights. These weights make it possible to get a better fit at some maturities of the instrument. The higher the weight w_i, the better the fit for instrument i will be.

If you use the model to price interest rate derivatives, it is sensible to put more weight on those instruments with the best matching maturity and expiry. In that case, we can expect the model to get accurate prices for the exotic to price. In the intermezzo, we selected several European swaptions with execution and payment dates equal to the possible execution and payment dates of the Bermudan swaption.

This squared error measure is derived from statistics. In statistics, the least squares method is used to find the maximum likelihood estimator for specific problems and their samples. It is also connected to the chi-square test, where for n measurements g_i of model predicted values y_i are assumed to be normally distributed around g_i with variances σ_i^2. Then

$$\sum_{i=1}^{n} \frac{(y_i - g_i)^2}{\sigma_i^2}, \tag{3.11}$$

follows the χ^2-distribution and testing the null-hypothesis is performed by comparing the sum of the squared differences with the $1 - \alpha$-quantile of the χ^2 distribution.

FINCAD XL, a widely used software program that contains "financial functions and prebuilt workbooks used for valuing and measuring the risk of financial securities and derivatives", uses the notion *data uncertainty* for the weights[1]. Data uncertainty δy_i is the estimated uncertainty in the quoted price of the i-th calibrating instrument. The choice of FINCAD is then $w_i = 1/\delta y_i$.
Suppose, for example, you have M market quotes for the i-th instrument at your disposal. The data uncertainty δy_i can be taken as the standard deviation of this sample of M prices. Compare this method with (3.11). The lower the uncertainty, the better the fit at the corresponding instrument will be, which makes sense.

Another possible choice for the weights w_i is the reciprocal of the quoted price: $1/y_i$. This gives more weight to lower priced swaptions or caps. In this thesis, we focus on this choice as this approach stresses the relativity of the error and is commonly applied. Hence we investigate calibration results by minimizing relative and absolute differences between market values and model values.

Assume we have r parameters to choose in the model, all being time independent. Hence $\mathbf{a}$ can be seen as a point in $\mathbb{R}^r$. Our goal is now to minimize the function $\chi^2 : \mathbb{R}^r \to \mathbb{R}$. It is possible to put constraints on the parameters. In that case, we restrict the function χ^2 to a set $C \subset \mathbb{R}^r$.
Usual constraints are to require that the value of a parameter lies at one side of

[1]see: www.fincad.com/support/developerFunc/mathref/Calibration.htm

a certain boundary.
For the Two-Additive-Factor Gaussian model, these boundaries are $a > 0$, $b > 0$, $\sigma > 0$, $\eta > 0$ and $-1 \leq \rho \leq 1$.
For the Libor Market Model the restrictions for the parameters are $\alpha_1 > 0$, $\alpha_2 > 0$, $\sigma_1 > 0$, $\sigma_2 > 0$ and $-1 \leq \rho \leq 1$.

The market values we mentioned so far, can be either the market prices of the swaptions and caps or the market implied volatilities of these interest rate derivatives. Because until now, there are no common ideas about best practice with this, we compared both methods.

3.3 Minimization algorithms

In this section, we take a look at three commonly known and widely used minimization algorithms. The first two algorithms are local minimization algorithms: the Levenberg-Marquardt algorithm and the Downhill Simplex method. Finally, we discuss a global minimization method, simulated annealing.

3.3.1 Levenberg-Marquardt

Levenberg-Marquardt is an iterative method developed by Marquardt in 1963 and based on an idea of Levenberg. The model is designed for least squares problems and is the standard of non-linear least-squares routines, see [PFTV92].

Suppose a (local) minimum of χ^2 is attained in $\mathbf{a}_{min}$. Then the Taylor expansion of χ^2 around the current point in the iteration, $\mathbf{a}_i$, is

$$\chi^2(\mathbf{a}) = \chi^2(\mathbf{a}_i) + \nabla\chi^2(\mathbf{a}_i)(\mathbf{a} - \mathbf{a}_i) + \frac{1}{2}(\mathbf{a} - \mathbf{a}_i)^T H(\mathbf{a} - \mathbf{a}_i) + \ldots, \quad (3.12)$$

where $\nabla\chi^2$ is the gradient of χ^2 and $H = \nabla^2\chi^2(\mathbf{a}_i)$ is the Hessian at $\mathbf{a}_i$. Hence we assume that $y(x; \cdot)$ in (3.10) is differentiable.

With formula (3.12), we can approximate $\chi^2(\mathbf{a}_{min})$. This results in

$$\chi^2(\mathbf{a}_{min}) \approx \chi^2(\mathbf{a}_i) + \nabla\chi^2(\mathbf{a}_i)(\mathbf{a}_{min} - \mathbf{a}_i) + \frac{1}{2}(\mathbf{a}_{min} - \mathbf{a}_i)^T H(\mathbf{a}_{min} - \mathbf{a}_i) \quad (3.13)$$

In a local minimum, the gradient of χ^2 is equal to zero. So we have that for $\mathbf{a}_{min}$ should hold that $\nabla\chi^2(\mathbf{a}_{min}) = 0$. Using the approximation (3.13) yields

$$\nabla\chi^2(\mathbf{a}_{min}) \approx \nabla\chi^2(\mathbf{a}_i) + H(\mathbf{a}_{min} - \mathbf{a}_i). \quad (3.14)$$

Now setting $\nabla\chi^2(\mathbf{a}_{min}) = 0$, gives the following approximation for $\mathbf{a}_{min}$:

$$\mathbf{a}_{min} \approx \mathbf{a}_i + H^{-1}\left(-\nabla\chi^2(\mathbf{a}_i)\right). \quad (3.15)$$

We write $-\nabla\chi^2(\mathbf{a}_i)$, that is the *steepest descent direction*. An iteration based on (3.15) is called *Newton's method*. The steps are as follows:

$$\mathbf{a}_{i+1} = \mathbf{a}_i + H^{-1}(-\nabla\chi^2(\mathbf{a}_i)). \tag{3.16}$$

This reasoning is based on the approximation (3.13). But maybe it is a poor estimation. Taking too large steps in the steepest descent direction causes a slow convergence, or even not a convergence at all. In that case, one could also take a small step in the steepest descent direction:

$$\mathbf{a}_{i+1} = \mathbf{a}_i - \mu\nabla\chi^2(\mathbf{a}_i), \quad 0 < \mu < 1 \tag{3.17}$$

This method is called the *steepest descent method*.
The gradient of χ^2 with respect to the parameters $\mathbf{a}$ has elements

$$\frac{\partial\chi^2}{\partial a_k} = 2\sum_{i=1}^{n} \frac{\partial y(x_i;\mathbf{a})}{\partial a_k}\frac{y_i - y(x_i;\mathbf{a})}{w_i^2} \quad k = 1, 2, \ldots, r \tag{3.18}$$

The Hessian H has components

$$\frac{\partial^2\chi^2}{\partial a_k \partial a_l} = 2\sum_{i=1}^{n} \frac{1}{w_i^2}\left(\frac{\partial y(x_i;\mathbf{a})}{\partial a_k}\frac{\partial y(x_i;\mathbf{a})}{\partial a_l} - (y_i - y(x_i;\mathbf{a}))\frac{\partial^2 y(x_i;\mathbf{a})}{\partial a_k \partial a_l}\right). \tag{3.19}$$

To ease notation, we define

$$\beta_k := -\frac{1}{2}\frac{\partial\chi^2}{\partial a_k} \qquad \alpha_{kl} := \frac{1}{2}\frac{\partial^2\chi^2}{\partial a_k \partial a_l}, \tag{3.20}$$

such that β is half times the steepest descent direction and $[\alpha] = \frac{1}{2}H$.

Let's write $\Delta\mathbf{a}_i = \mathbf{a}_{i+1} - \mathbf{a}_i$. With the new notation, while omitting the subscript i for $[\alpha]$ and β to prevent any confusion, we have

$$\begin{array}{lr} \text{Newton method:} & [\alpha]\Delta\mathbf{a} = \beta \\ \text{Steepest descent method:} & \Delta\mathbf{a} = \mu\beta \end{array}$$

The Levenberg-Marquardt method combines these two methods. Far from the minimum, the approximation (3.13) is rather poor and the possibility that the steepest descent method steps over the minimum is nihil. So, let the iteration step particularly be determined by the steepest descent method. Close to the minimum, the second approximation gets better and the steepest descent method could step too far. So stress the inverse Hessian matrix more.
If there are exact analytical formulas for the partial second derivatives of $y(x;\mathbf{a})$ to $\mathbf{a}$, the inverse Hessian method is still based on approximation (3.13). If those formulas do not exist, one should numerically compute the Hessian. The numerical

estimate will not be accurate if χ^2 is highly non-linear. Therefore, Levenberg-Marquardt changes the definition for $[\alpha]$ into

$$\alpha_{kl} = \sum_{i=1}^{n} \frac{1}{w_i^2} \left(\frac{\partial y(x_i;\mathbf{a})}{\partial a_k} \frac{\partial y(x_i;\mathbf{a})}{\partial a_l} \right). \tag{3.21}$$

Since Newton is used close to the minimum, the difference $y_i - y(x_i;\mathbf{a})$ is very small. The new $[\alpha]$ will not deviate much from the old one.
The advantage is that no second derivatives have to be computed. The change does not lead to other minima. Only the route of the iteration is changed slightly.

The Levenberg-Marquardt algorithm switches between the two algorithms by adding a multiple of the identity matrix to $[\alpha]$ getting $[\alpha'] = [\alpha] + \lambda I$ and in step i solving

$$[\alpha']\Delta\mathbf{a} = \beta. \tag{3.22}$$

For large $\lambda \gg 1$, equation (3.22) transforms into the steepest descent with constant $1/\lambda$ and for λ approaching zero, equation (3.22) goes over to Newton's method.
This is the principle Levenberg-Marquardt uses. The scalar λ is adjusted in each iteration step. If at an iteration step the χ^2 value decreases, so we get closer to a minimum, decrease λ by a predefined substantial factor (e.g. 10). If, on the other hand, the χ^2 value increase, increase λ by the predefined factor, stressing the steepest descent direction.

The algorithm, in pseudo code, of Levenberg-Marquardt is:

1. Compute $\chi^2(\mathbf{a})$ with $\mathbf{a} = \mathbf{a}_{initial}$;
2. Choose an initial value for λ, e.g. $\lambda = 0.001$;
3. Solve (3.22) for $\Delta\mathbf{a}$;
4. If $\chi^2(\mathbf{a} + \Delta\mathbf{a}) \geq \chi^2(\mathbf{a})$, increase λ and go to 3;
   If $\chi^2(\mathbf{a} + \Delta a) < \chi^2(\mathbf{a})$, decrease λ and update the solution $\mathbf{a} = \mathbf{a} + \Delta\mathbf{a}$ and go to 3;
5. Stopping criterium

The iteration stops if the χ^2 value did not decrease more than a value *tol*. But if in the last step χ^2 **increased**, the method should continue.
The method has good convergence properties, see for details [Mar63] and [Haü83].

3.3.2 Downhill Simplex

The Levenberg-Marquardt algorithm is specially adapted to minimize functions of the form (3.10). The Downhill Simplex method, also known as Nelder-Mead algorithm, is designed to solve unconstrained minimization problems of the form

$$\min_{x \in \mathbb{R}^r} f(x), \tag{3.23}$$

where f is continuous. In the case of a measure of fit, we have

$$\min_{\mathbf{a} \in \mathbb{R}^r} \chi^2(\mathbf{a}). \tag{3.24}$$

The downhill simplex method is also iterative and starts with $r+1$ initial points. The method does not use evaluations of derivatives. Only function values are needed.
The $r+1$ points in $\mathbb{R}^r$ form a simplex. In each iteration step, the simplex is moved or shrinked. It moves into a direction opposite to the point in which χ^2 has the largest value or shrinks into the direction of the point with lowest χ^2 value.

We enumerate the $r+1$ initial points $\mathbf{a}_0, \mathbf{a}_1, \ldots, \mathbf{a}_r$, in increasing order of χ^2 value. So $\chi^2(\mathbf{a}_0) \leq \chi^2(\mathbf{a}_1) \leq \ldots \leq \chi^2(\mathbf{a}_r)$. At each iteration, four possible points are considered to replace $\mathbf{a}_r$. Those possible points are obtained by mirroring $\mathbf{a}_r$ in the point $\bar{\mathbf{a}}$, the middle of the other points:

$$\bar{\mathbf{a}} = \frac{\mathbf{a}_0 + \ldots \mathbf{a}_r}{r}.$$

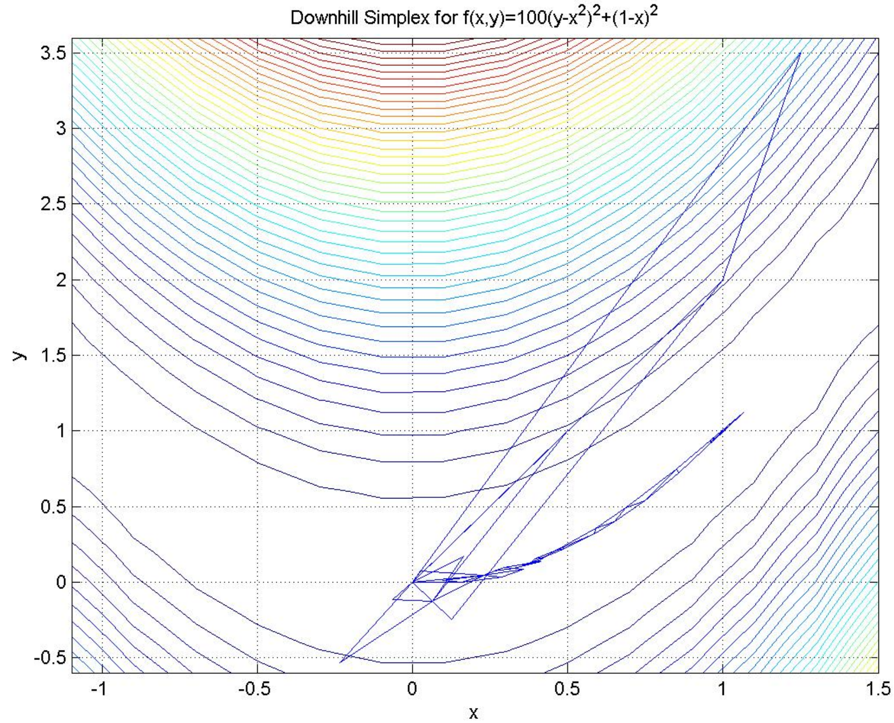
The four candidates are

$\mathbf{a}_m = \bar{\mathbf{a}} + (\bar{\mathbf{a}} - \mathbf{a}_r)$	reflection of $\mathbf{a}_r$ in the midpoint $\bar{\mathbf{a}}$;
$\mathbf{a}_e = \bar{\mathbf{a}} + 2(\bar{\mathbf{a}} - \mathbf{a}_r)$	expansion of the reflected point $\mathbf{a}_m$;
$\mathbf{a}_{rc} = \bar{\mathbf{a}} + 0.5(\bar{\mathbf{a}} - \mathbf{a}_r)$	contraction of the simplex, after reflection;
$\mathbf{a}_c = \bar{\mathbf{a}} - 0.5(\bar{\mathbf{a}} - \mathbf{a}_r)$	contraction, $\mathbf{a}_r$ is replaced by the midpoint of the segment $(\mathbf{a}_r, \bar{\mathbf{a}})$.

The algorithm is:

1. Sort the initial points such that the χ^2 value are in ascending order;
2. Calculate $\mathbf{a}_m$ and $\chi^2(\mathbf{a}_m)$;
3. Depending on $\chi^2(\mathbf{a}_m)$, choose one of the following actions:
   a) If $\chi^2(\mathbf{a}_m) < \chi^2(\mathbf{a}_0)$, first calculate $\mathbf{a}_e$ and $\chi^2(\mathbf{a}_e)$, second, if $\chi^2(\mathbf{a}_e) < \chi^2(\mathbf{a}_0)$ replace $\mathbf{a}_r$ by $\mathbf{a}_e$, otherwise, replace $\mathbf{a}_r$ by $\mathbf{a}_m$.
   b) If $\chi^2(\mathbf{a}_0) \leq \chi^2(\mathbf{a}_m) \leq \chi^2(a_{r-1})$, replace $\mathbf{a}_r$ by $\mathbf{a}_m$.
   c) If $\chi^2(r-1) < \chi^2(\mathbf{a}_m) \leq \chi^2(\mathbf{a}_r)$, first calculate $\mathbf{a}_{mc}$ and$\chi^2(\mathbf{a}_{mc})$, if then $\chi^2(\mathbf{a}_{mc}) \leq \chi^2(\mathbf{a}_r)$, replace $\mathbf{a}_r$ by $\mathbf{a}_{mc}$.
   d) If $\chi^2(\mathbf{a}_r) < \chi^2(\mathbf{a}_m)$, first calculate $\mathbf{a}_c$ and $\chi^2(\mathbf{a}_c)$, if $\chi^2(\mathbf{a}_c) \leq \chi^2(\mathbf{a}_r)$, then replace $\mathbf{a}_r$ by $\mathbf{a}_c$.
4. If after step 4 no new vertex is defined, shrink the simplex. Replace every point, except $\mathbf{a}_0$, by its midpoint of the segment between itself and $\mathbf{a}_0$.
5. If the stopping criterium is satisfied, stop, else go to 1.

In the graph below, a Downhill Simplex path is shown for the function $f(x, y) = 100(y - x^2)^2 + (1 - x)^2$.



The usual stopping criterium looks like: $|f(x_{i+1})-f(x_i)| < tol$, where tol is some small positive value, e.g. 0.001. So the iteration stops if the function value does not improve much more. But in the case of Downhill Simplex, there are $r+1$ function values in stead of just one. Hence we take as criterium

$$\chi^2(\mathbf{a}_r) - \chi^2(\mathbf{a}_0) < tol \tag{3.25}$$

But with this criterium, there is a possibility that the r points are scattered and not concentrated around a minimum. So a second criterium is

$$\|\mathbf{a}_i - \mathbf{a}_0\| < tol'. \tag{3.26}$$

The method does not possess very good convergence properties, see [PCB02]. In some cases, the method gets stuck while it did not find a local minimum. Therefore, when the iteration ends and comes up with a point $\mathbf{a}_{min}$, it is advised in [PFTV92] to run the method an other time but with one initial point, namely $\mathbf{a}_r$, replaced by $\mathbf{a}_{min}$. If the simplex shrinks again towards $\mathbf{a}_{min}$, we can assume that $\mathbf{a}_{min}$ is indeed a local minimum. But if the algorithm shows up with an other end point $\mathbf{a}'_{min}$, just start the program again with the initial point $\mathbf{a}_r$ replaced by the smallest of $\mathbf{a}_{min}$ and $\mathbf{a}'_{min}$.

As the algorithm shows, the method does not evaluate derivatives. To estimate derivatives numerically, especially in many dimensions, is a rather tremendous job. A disadvantage is however that Downhill Simplex needs many iteration steps for convergence, since it does not use any information of the function χ^2. But for very wild functions, it is preferable not to evaluate the derivatives and make use of a 'robust' method like Downhill Simplex.

Furthermore, as for Levenberg-Marquardt, the minimum found by Downhill Simplex is only a local minimum. It is easy to see why. The method just looks for a downhill direction and then rolls down the hill. In this way, only local minima are found.

3.3.3 Simulated annealing

The Levenberg-Marquardt algorithm and Downhill Simplex method search for local minima. But the function to minimize, the measure of fit, could be non-convex such that a local minimum does not have to be automatically a global minimum.

One of the best known algorithms for global minimization is *simulated annealing*. The method has originally been designed for combinatorial global optimization and published in 1983 by Kirkpatrick, Gelatt and Vecchi, see [KGV83]. The method is derived from thermodynamics, in special the physical process that is called 'annealing' by which molten metals cool and crystallize. This process is performed by lowering the temperature slowly and gradually in stages, in order to attain a state in which the metal has a rigid structure. This state is a natural state in which the metal has minimum energy. By cooling the metal slowly, this state is almost certainly reached. But if the temperature is decreased too fast, the metal ends up in some amorphous state with higher energy.

The link between physical annealing and minimization of some function by simulated annealing is that the function should be regarded the energy and the parameter space as the possible states of the metal.

To draw conclusions about the calibration risk, we want the measure of fit to be in a global minimum. Hence, we use simulated annealing to find the parameters that fit the market best. What is meant with 'fit the market best', is defined by the measure of fit and which calibration instruments are selected.

There are more global optimization algorithms, like Classical multi-start eamped Levenberg-Marquardt, Classical Tunneling and Exponential Tunneling and Multi-start selective search. In [VPTZ01], Veláquez et al. investigated these methods and their results show that for a set of nonlinear least-square problems simulated annealing is the fasted if the number of variables is not more then ten. Simulated annealing also converged to the global minimum. Based on this article and the general application of simulated annealing, this method is used for the calibrations of the interest rate models.

Consider the following optimization problem:

$$\min_{\mathbf{a}\in C} \chi^2(\mathbf{a}),$$

where $C \subset \mathbb{R}^n$ and χ^2 a continuous function.

The general simulated annealing algorithm is the following.

1. Let $\mathbf{a}_0 \in C$ be a certain starting point. Let $y_0 = \chi^2(\mathbf{a}_0)$. The initial temperature is T_0 and $k = 0$.
2. Choose randomly another point $\mathbf{b}_{k+1}$ in the parameter space according to some next candidate distribution $D(\cdot, \mathbf{a}_k, y_k)$.
3. Calculate $y_{k+1} = \chi^2(\mathbf{b}_{k+1})$. Draw a uniform $[0,1]$ distributed variate U. Set
$$\mathbf{a}_{k+1} = \begin{cases} \mathbf{b}_{k+1} & \text{if } U < e^{-\frac{y_{k+1}-y_k}{T_k}} \\ \mathbf{a}_k & \text{otherwise} \end{cases}$$
4. Set $T_{k+1} = C(T_k)$. This is cooling, lowering the temperature. Cooling can be dependent on the number of successes, that is, the number of times a next candidate was accepted.
5. If the stopping criterium is satisfied, stop. Else, set $k = k + 1$ and go to 2.

In simulated annealing, one moves to a new state if the function value, energy, is lower. One might move to a point with higher energy, dependent on an exponential random variable with mean T_k. The higher the temperature, the more likely a state with higher energy is chosen as next iteration step.
The acceptance criterium in step 3 is called the *Metropolis criterium*. It is a widely used criterium and one of the first developed criteria.

A practical overview of the application of simulated annealing is found in [BM95]. Brooks and Morgan advise in this article to run steps 2 and 3 until the system is in some equilibrium and after that, the temperature should be lowered. To recognize an equilibrium, one runs after s successes another N steps. These s and N should be chosen in advance by checking the best values.
The cooling schedule Brooks and Morgan mention is multiplying T_k by ρ where

$0 \leq \rho \leq 1$, if an equilibrium state is reached.
The next candidate schedule we used for the calibration is adding a random number to a random coordinate of $\mathbf{a}_k$, taking into account that $\mathbf{b}_k \in C$. This is idea also derived from [BM95].

If the annealing schedule is sufficiently slow, then the system will end up in a state of minimum energy, the global minimum of the measure of fit χ^2. This state Brooks and Morgan and is set out by Ingber in [Ing93] using Markov chains. Locatelli gives in [Loc00] conditions for convergence of simulated annealing. There are some conditions on the function that is to be minimized, for which it is not easy to show that they hold for the measure of fits we defined in section 3.2. Moreover, the measure of fit is in some cases obtained by numerical procedures which might affect any smoothness of those functions. We tested the convergence quality of the simulated annealing procedures by applying them to simple minimization problems and to the calibration problems. The conclusion of these investigations is that with a cooling schedule of $C(t) = 0.8t$, $N = 300$ and $s = 20$, initial temperature $T_0 = 1$ and stopping criterium "$T_k < 10^{-8}$" the simulated annealing procedure is not fast but converges.

Chapter 4

Profit sharing contracts

Insurers provide ways to participate in the company's profit. It is an attractive extension of basic insurance products, often of group pension insurance.
The payoff is part of the gain on investments after deduction of costs. Sources of the profit of an insurance company are the actuarial principles interest, mortality, costs and additional technical assumptions. The total result on these sources generates profit.

The main source of profit is the income on investments. The insurer invests the incoming premiums during the maturity of the contract and a part of the return on those investments pours to the policy holders. Most profit sharing contracts distribute the profit to the policyholders at the end of the year, but there are also contracts written that take predetermined profits into account, the expected profit in the year to come.

The most common forms of profit sharing are excess interest sharing and interest rate discount. This latter type guarantees in advance a higher return than the basic interest rate. This higher return will be guaranteed during a pre-determined period and the expected interest profit will be charged immediately.

In the sequel we will focus on excess interest sharing and price this kind of so-called embedded option with the two introduced interest rate models calibrated in various ways.

4.1 Excess interest sharing

The excess interest is based on a prescribed or published return. Excess interest is the difference between the scored interest on the reserve of the insurer and the basic interest rate. Kessels investigated the reserves at market value that

insurance companies should hold to cover future interest liabilities in an excess interest sharing system. She used several interest rate models to discount the liabilities and estimate u-returns and examined the differences in the reserves for the models. For these differences, the interest rates are important. The mentioned u-return is used as return on investments. The value of the u-return is published by the 'Verbond van Verzekeraars' in the Netherlands. The u-return is based on the average of effective returns of a selection of government loans with time to maturity between two and ten years.
The basic interest, which is fixed, usually lies between 3% and 4%. The excess interest is paid out in the form of a discount on the premium to the policyholders.

The total reserve that the insurer should have consists of a reserve for the expected pension payments and a reserve for the excess interest. The expected pension payments are estimated in advance, so these payments can be discounted with the term structure.
The excess reserve has to be computed by risk neutral valuation. The excess interest can be computed with a fictive Tranche System. How that calculation process exactly takes place is described in [Kes07].

The contract we regard obliges the insurance company to pay the excess interest for 5 years. If the excess interest is positive in year 5, the profit sharing goes on till the first year in which the excess interest gets negative, with a maximum of 15 years.

The value of the excess interest option can be written as

$$V = \mathbb{E}^{\mathbb{Q}} \left[\sum_{t=0}^{20} \frac{\text{Excess Interest Rate Payments(t)}}{\text{Cash Rollup(t)}} \right], \tag{4.1}$$

where $\mathbb{E}^{\mathbb{Q}}$ indicates we take expectations under the risk neutral measure, with corresponding numeraire the cash rollup.

The option is valued by Monte Carlo simulation via

$$V = \frac{1}{N} \sum_{i=1}^{N} \sum_{t=1}^{20} \frac{M[\text{u–return}-/-4\%]^+}{\text{Cash Rollup}_i(t)}, \tag{4.2}$$

where N is the number of simulations, n the number of years with option payments and M the notional amount.

It is not immediately clear how to simulate the u-returns. Several ways are possible, what the possibilities have in common is the idea to write the u-return as a linear combination of some swap rates. Some choices for the u-return are

7-year swap rate and a multiple of the 8-year swap rate. This last one is chosen by Kessels. The multiple is determined by regression analysis of the historical u-returns and historical 8-year swap rates. The results show that a factor .95 is appropriate. We are not going to repeat the regression for end June 2007, we assume that:

$$\text{u–return}(t) = 0.95 \text{ swaprate}(t, 8). \tag{4.3}$$

The value of the multiple is not important for the investigation on calibration risk, it does not affect the comparison of calibration techniques.

4.2 Calibration results

In chapter three, several options for calibration of the interest rate models have been presented, either in the form of selection of calibration instruments or in the form of the measure of fit. Collecting these options, the eight calibrations that are executed for both interest rate models (Two-Additive-Factor Gaussian and Libor Market Model) are presented in the following table.

Calibration instruments	**Squared differences in measure of fit**
Swaptions	Absolute Prices Relative Prices Absolute Implied Volatilities Relative Implied Volatilities
At-the-money Caps	Absolute Prices Relative Prices
All Caps	Absolute Prices Relative Prices

The measure of fits discussed in 4.3 are thus of these forms:

Absolute Prices (AP) $\sum_{i=1}^{n}\left(P_i^{\text{model}} - P_i^{\text{market}}\right)^2$

Relative Prices (RP) $\sum_{i=1}^{n}\left(\frac{P_i^{\text{model}} - P_i^{\text{market}}}{P_i^{\text{market}}}\right)^2$

Absolute Implied Volatilities (AV) $\sum_{i=1}^{n}\left(IV_i^{\text{model}} - IV_i^{\text{market}}\right)^2$

Relative Implied Volatilities (RV) $\sum_{i=1}^{n}\left(\frac{IV_i^{\text{model}} - IV_i^{\text{market}}}{IV_i^{\text{market}}}\right)^2$

Here P_i^{model} and P_i^{market} are the model respectively market prices of calibration instrument i. IV_i^{model} and IV_i^{market} are the Black implied volatilities according to the model respectively to the market for calibration instrument i.

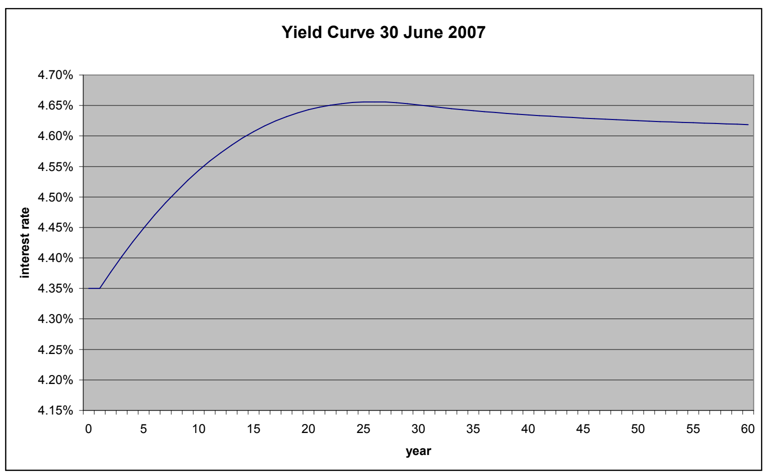
4.2.1 Market data

The market data used for calibration and for valuing the excess interest option is given in this section. Since it plays an important role and could have impact on the calibration outcomes, it is not moved to an appendix.

Interest rate

The continuous interest rates on an annual basis at 30 June 2007 for the Euro zone are the following:

year	yield
1	4.35%
2	4.38%
3	4.40%
4	4.43%
5	4.45%
6	4.47%
7	4.49%
8	4.51%
9	4.53%
15	4.61%
20	4.64%
25	4.66%
30	4.65%
40	4.63%
50	4.63%



Caps

The following table contains Black implied volatilitities for caps.

Cap volatility matrix as at 30 June 2007 on Euro market										
	1.75%	2.00%	2.25%	2.50%	3.00%	3.50%	4.00%	5.00%	6.00%	7.00%
1	16.16%	15.15%	13.85%	12.70%	10.70%	9.05%	7.75%	6.80%	7.55%	8.50%
2	20.88%	19.43%	18.18%	17.07%	15.25%	13.85%	12.50%	10.95%	11.53%	11.95%
3	20.73%	19.50%	18.43%	17.52%	15.92%	14.65%	13.45%	12.10%	12.50%	12.78%
4	20.45%	19.35%	18.35%	17.48%	16.05%	14.85%	13.80%	12.60%	12.85%	13.05%
5	20.55%	19.40%	18.40%	17.55%	16.10%	14.95%	14.00%	12.85%	12.95%	13.10%
6	20.10%	19.05%	18.10%	17.30%	15.95%	14.88%	14.00%	12.90%	12.95%	13.05%
7	19.75%	18.70%	17.80%	17.02%	15.75%	14.75%	13.90%	12.90%	12.85%	12.95%
8	19.55%	18.50%	17.65%	16.85%	15.60%	14.65%	13.85%	12.90%	12.85%	12.92%
9	19.20%	18.20%	17.35%	16.60%	15.40%	14.45%	13.70%	12.80%	12.70%	12.80%
10	19.00%	18.00%	17.15%	16.43%	15.22%	14.30%	13.60%	12.70%	12.55%	12.65%
12	18.55%	17.55%	16.73%	16.00%	14.85%	13.95%	13.25%	12.40%	12.15%	12.20%
15	18.28%	17.28%	16.43%	15.70%	14.55%	13.65%	12.97%	12.10%	11.85%	11.88%
20	17.90%	16.90%	16.05%	15.30%	14.12%	13.25%	12.60%	11.75%	11.50%	11.55%

The corresponding prices of the caps depend on the yield curve. This table presents the prices of caps with a principal amount of €100 at 30 June 2007.

Cap prices (€) as at 30 June 2007 on Euro market										
	1.75%	2.00%	2.25%	2.50%	3.00%	3.50%	4.00%	5.00%	6.00%	7.00%
1	1.27	1.15	1.03	0.91	0.67	0.43	0.19	0.00	0.00	0.00
2	3.77	3.42	3.07	2.72	2.02	1.33	0.69	0.05	0.00	0.00
3	6.21	5.64	5.07	4.50	3.36	2.27	1.28	0.20	0.03	0.00
4	8.58	7.80	7.02	6.24	4.71	3.24	1.93	0.43	0.09	0.02
5	10.89	9.91	8.93	7.95	6.04	4.23	2.63	0.71	0.19	0.05
6	13.12	11.95	10.78	9.63	7.36	5.22	3.34	1.02	0.31	0.10
7	15.28	13.93	12.59	11.26	8.65	6.21	4.06	1.37	0.46	0.16
8	17.37	15.85	14.34	12.84	9.92	7.19	4.79	1.74	0.64	0.25
9	19.39	17.70	16.03	14.37	11.16	8.15	5.50	2.11	0.82	0.34
10	21.33	19.49	17.66	15.86	12.36	9.09	6.22	2.49	1.01	0.44
12	24.98	22.86	20.75	18.67	14.64	10.88	7.58	3.22	1.38	0.64
15	29.93	27.43	24.96	22.52	17.80	13.41	9.54	4.32	2.00	1.00
20	36.83	33.82	30.85	27.93	22.28	17.03	12.41	6.03	3.03	1.66

The market values of at-the-money caps are quoted seperately, with the corresponding at-the-money cap rate for each term. The prices in the table are for a cap with principal amount of €100.

At-the-money caps market data as at 30 June 2007			
Term	ATM strikes	ATM volatility	ATM prices (€)
1	4.64%	6.90%	0.01
2	4.67%	10.70%	0.14
3	4.68%	12.00%	0.38
4	4.69%	12.60%	0.70
5	4.69%	12.90%	1.08
6	4.70%	13.00%	1.47
7	4.71%	13.00%	1.89
8	4.72%	13.00%	2.32
9	4.74%	12.90%	2.72
10	4.75%	12.80%	3.14
12	4.79%	12.45%	3.85
15	4.85%	12.20%	4.89
20	4.86%	11.85%	6.69

Swaptions

Swaptions are quoted in implied volatility matrices, in the vertical scale is the expiry of the option shown, on the horizontal scale the length of the swap, both in years.

Eurpean swaption implied volatilities at 30 June 2007 on Euro market											
	1	2	3	4	5	...	10	15	20	25	30
1	12.30%	12.80%	12.90%	12.90%	12.80%	...	12.40%	11.90%	11.60%	11.60%	11.55%
2	13.10%	13.10%	13.00%	12.90%	12.80%	...	12.30%	11.80%	11.55%	11.55%	11.45%
3	13.20%	13.10%	13.00%	12.80%	12.60%	...	12.10%	11.65%	11.35%	11.30%	11.25%
4	13.00%	12.90%	12.80%	12.60%	12.40%	...	11.90%	11.45%	11.20%	11.15%	11.05%
5	12.90%	12.70%	12.60%	12.40%	12.20%	...	11.70%	11.25%	11.00%	10.95%	10.90%
7	12.40%	12.30%	12.10%	11.90%	11.70%	...	11.30%	10.90%	10.65%	10.70%	10.50%
10	11.50%	11.40%	11.30%	11.10%	11.00%	...	10.80%	10.50%	10.15%	10.10%	10.00%
15	11.00%	10.60%	10.55%	10.40%	10.40%	...	10.25%	9.80%	9.55%	9.35%	9.35%
20	10.55%	10.25%	10.05%	9.95%	9.90%	...	9.90%	9.35%	9.00%	8.95%	8.90%
25	10.45%	10.05%	9.95%	9.90%	9.85%	...	9.70%	9.25%	8.85%	8.50%	8.80%
30	10.45%	10.05%	9.80%	9.65%	9.45%	...	9.55%	8.90%	8.60%	8.50%	8.80%

In this table, market data for swaptions with an underlying swap with length six year until nine year are omitted for layout reasons but are used for calibration.

The prices of these European swaption are given in next table. All swaptions have a principal amount of €100.

Eurpean swaption prices (€) as at 30 June 2007											
	1	2	3	4	5	...	10	15	20	25	30
1	0.20	0.41	0.62	0.81	0.98	...	1.74	2.28	2.69	3.05	3.32
2	0.29	0.58	0.85	1.10	1.34	...	2.35	3.06	3.62	4.11	4.45
3	0.35	0.68	1.00	1.29	1.56	...	2.72	3.55	4.18	4.71	5.13
4	0.38	0.75	1.09	1.41	1.70	...	2.96	3.86	4.55	5.13	5.56
5	0.41	0.79	1.16	1.49	1.80	...	3.12	4.06	4.78	5.39	5.85
7	0.43	0.84	1.21	1.56	1.88	...	3.26	4.25	4.99	5.68	6.08
10	0.42	0.82	1.19	1.52	1.85	...	3.25	4.25	4.94	5.56	6.01
15	0.39	0.74	1.08	1.38	1.69	...	2.97	3.81	4.46	4.95	5.41
20	0.34	0.65	0.93	1.20	1.45	...	2.58	3.28	3.80	4.29	4.67
25	0.29	0.55	0.80	1.03	1.25	...	2.21	2.85	3.29	3.59	4.07
30	0.25	0.47	0.67	0.87	1.04	...	1.88	2.37	2.77	3.11	3.53

4.2.2 Calibration outcomes

In chapter three, we discussed three minimization algorithms. The choice for the best algorithm for the calibration of the interest models is not clear in advance. Hence, we chose the minimization method 'on the job'.

In general, starting with a global minimizer is done. So in the calibration process performed for valuing the excess interest option, the measures of fit were first minimized by the simulated annealing method. Since this method requires a slow cooling process, see section 3.3, the computation time was quite long. To overcome this, several methods that combine local and global minimization algorithms are suggested in the literature, see for example [Loc02].

For the actual calibration we followed the global optimization method simulated annealing with a local optimization method. Due to the stochastics incorporated in simulated annealing and the slow cooling process, there is still some randomness in the final minimum found by the simulated annealing method. That randomness was eliminated by applying a local optimization algorithm. This specific approach is also used in [BM01].
In general, a danger involved in applying such hybrid methods is that one could end up in a local minimum near a global minimum since simulated annealing might be ended prematurely. This problem is overcome by keeping track of the iteration of simulated annealing. The best point seen in the iteration, that is the iteration step with minimal χ^2 value, is stored. If this point is not in a neighborhood of the final solution, simulated annealing was restarted with that specific point, until the "best seen" equals the final point.

After running these simulated annealing methods, both Levenberg-Marquardt and Downhill Simplex have been applied for some of the sixteen problems. There were only very slight differences in the optima that were found by Levenberg-Marquardt and Downhill Simplex. This can be explained by the fact that the simulated annealing already stuck up near a global minimum and hence the choice of the minimization function and the stopping criteria have little impact. Only if the measure of fit is quite rough, which might especially be the case for Two-Additive-Factor Gaussian for swaptions where numerical integration is incorporated in χ^2, differences can be expected since a local optimization algorithm can easily end up in a local minimum near the global minimum.

After these careful considerations and experiences, we applied simulated annealing followed by Downhill Simplex, which does not use derivatives of the function to minimize.

The results, the parameters that give the best fit and the corresponding χ^2-value, for the eight calibrations of the Two-Additive-Factor Gaussian model are presented below:

Two-Additive-Factor Gaussian calibrated parameters							
		a	b	σ	η	ρ	χ^2
Swaptions	AP	0.86248	0.02360	0.01027	0.00652	-0.36983	1.76E-04
	RP	0.98405	0.02441	0.00754	0.00641	-0.47577	1.69E-01
	AV	0.98408	0.02291	0.00462	0.00629	-0.47574	1.97E-03
	RV	0.65702	0.02343	0.00360	0.00634	-0.46405	1.78E-01
All Caps	AP	0.80027	0.09501	0.01296	0.01089	-0.96788	1.90E-04
	RP	0.99990	0.07724	0.01190	0.01078	-1.00000	4.54E+00
ATM Caps	AP	0.99990	0.04629	0.00439	0.00751	-0.96111	4.70E-07
	RP	0.97550	0.12051	0.01221	0.01124	-1.00000	1.67E-02

The column χ^2 gives the value of the measure of fit for the corresponding calibrated model parameters. These cannot be compared, since the function χ^2 is in fact different for all eight calibrations.

The calibrated parameters for the Two-Additive-Factor Gaussian model are for the four calibrations to swaptions quite similar. The parameter for correlation ρ differs significantly from -1 and thus there are really two stochastic factors driving the short rate.
The parameters for the calibration to caps are different from the calibration of Two-Additive-Factor Gaussian to swaptions. Striking is ρ which is almost equal to -1 for these four calibrations. Clearly, a one factor model could have been used. It seems that by the lack of information on correlation in the cap quotes a one factor model is really sufficient to explain the cap prices.

The results for the Libor Market Model are as follows:

Libor Market Model calibrated parameters							
		α_1	α_2	σ_1	σ_2	ρ	χ^2
Swaptions	AP	0.00482	0.02090	0.00013	0.13674	-0.29469	1.54E-04
	RP	0.16935	0.01694	0.02124	0.12390	0.74000	1.69E-01
	AV	0.00000	0.04845	0.05899	0.08415	0.88163	2.02E-03
	RV	0.04950	0.00000	0.08687	0.06056	0.77894	1.68E-01
All Caps	AP	0.03178	0.66732	0.14334	0.00000	0.82697	2.87E-04
	RP	0.05957	1.13196	0.16815	0.29224	-1.00000	2.06E+00
ATM Caps	AP	0.06932	0.68791	0.15918	0.00000	0.89849	1.85E-05
	RP	0.07215	1.11777	0.17897	0.31685	-1.00000	5.63E-04

The correlation parameter ρ is for the swaption calibration larger than -1. Two factors seem to be needed to fit the model swaption values to the market swaption values. But for the AP calibration, the corresponding σ_1 is nearly zero. The first factor has almost no volatility, hence a one factor model would have been appropriate.

For the caps we see the same: one factor would have been sufficient. For two calibrations, $\rho \approx -1$ and for the other two, $\sigma_2 \approx 0$. As for Two-Additive-Factor Gaussian, caps seems to require just one factor in stead of more factors.

If we compare the χ^2 values per calibration between the models, we see that for the swaption calibration they are of similar size. The Libor Market Model seems to fit slightly better. For all caps calibration, these values are also in line. For the calibration with respect to at-the-money caps, Two-Additive-Factor Gaussian fits better for absolute price differences and Libor Market Model for relative price differences.

4.3 Simulation of excess interest option

With both models we price the excess interest option by Monte Carlo simulation with the number of simulations set to one million. How Monte Carlo simulation is executed for both interest rate models is described in the appendix.

The following two tables contain the resulting prices for an excess interest option with principal $M = 100$.

Monte Carlo Two-Factor-Additive Gaussian			
		Option Value	Standard Error
Swaptions	AP	11.675	4.23E-03
	RP	11.512	4.12E-03
	AV	11.499	4.04E-03
	RV	11.488	4.03E-03
All Caps	AP	11.052	3.70E-03
	RP	12.160	4.09E-03
ATM Caps	AP	11.332	3.79E-03
	RP	10.328	3.07E-03

Monte Carlo Libor Market Model			
		Option Value	Standard Error
Swaptions	AP	11.471	7.48E-03
	RP	11.297	7.26E-03
	AV	11.281	7.22E-03
	RV	11.273	7.21E-03
All Caps	AP	11.070	6.73E-03
	RP	10.258	5.53E-03
ATM Caps	AP	9.693	4.44E-03
	RP	10.001	5.06E-03

The prices of the excess interest option by the two interest rate models are close to each other for the calibration to swaptions, regardless the measure of fit that is minimized. These eight values all lie between 11.273 and 11.675. The calibrations to caps however, result in option prices from 9.693 to 12.160. The calibrated parameters to caps of both models look different than the parameters obtained from calibration to swaptions. That seems to have effect on the price of the excess interest option.

Another aspect that attracts the eye is that both interest rate models when calibrated to swaptions assign a higher price to the option if the measure of fit consists of absolute price differences than the other three measures of fit.

4.3.1 Calibration risk

The calibration risk can be investigated by comparing price ratios. The price ratios for swaptions are as follows:

Price ratios of calibration to swaptions					
	AP/RP	AP/AV	RP/AV	RP/RV	AV/RV
G2	101.42%	101.53%	100.11%	100.20%	100.09%
LMM	101.54%	101.68%	100.14%	100.22%	100.08%

In this table, the resulting excess option prices are divided by each other, to give insight in the relative deviation from each other.

All values are close to each other. The range of option prices is somewhat larger for the Libor Market Model, but not significantly. That the RP, AV, and RV lie closely to each other and that the AP price is somewhat higher is clear from these ratios. The cause for the similar prices for calibration with respect to relative prices, absolute implied volatilities and relative implied volatilities is possibly the fact that these three measures of fit all stress relative differences. Implied volatilities in fact are a kind of relative difference as implied volatilities are of comparable size for different terms and times to maturity.

This analysis can be done for prices if the models have been calibrated to caps. 'All RP/ATM AP' means the price of the option by the model that was calibrated to all caps with the relative price differences in the measure of fit divided by the price of the option by the model that was calibrated to at-the-money caps with absolute price differences.

Price ratios of calibration to caps					
	All AP/ All RP	All AP ATM AP	All RP/ ATM AP	All RP/ ATM RP	ATM AP/ ATM RP
G2	90.89%	97.53%	107.31%	117.74%	109.72%
LMM	107.91%	114.20%	105.83%	102.57%	96.92%

The option prices all differ very much. The choice of the measure of fit and the choice whether to use all caps or only at-the-money caps has a lot of impact on the option price for a given interest rate model. The market cap prices range from 0.00 to 36.83. Relative price differences stress the small values very much: values for a strike of more than 5% and a term of less than ten years, refer to the table with cap prices.
This might not be representative for the excess interest option. Apparently, this leads to higher prices for the Two-Additive-Gaussian model and lower prices for the Libor Market Model. For swaption calibration, this is not a problem as swaption market prices lie more closely to each other.

Calibration to at-the-money caps uses thirteen market prices and five corresponding model values that should fit these market prices. Possibly, too few restrictions

and volatility information is incorporated in those thirteen data points.
The only cap values that give enough information is the set of all cap prices. Moreover, one would calibrate the models then to caps by minimizing absolute differences between market and model prices. Then we would compare calibration to all caps according to absolute price differences with swaption calibrations:

Price ratios of calibrations to swaptions and all caps				
	SAP/CAP	SRP/CAP	SAV/CAP	SRV/CAP
G2	105.64%	104.16%	104.04%	103.95%
LMM	103.63%	102.06%	101.91%	101.83%

In this tale, e.g. SRP means calibrated to swaption under relative price differences and CAP stands for calibrated to all caps under absolute price differences. The Libor Market Model prices are more in line with each other than the Two-Additive-Factor Gaussian calibration.
However, for both models, the difference between swaption calibration and all caps calibration is larger than the differences among the swaption calibrations. We already mentioned the problems with the at-the-money cap calibrations and the calibration to relative prices for caps.

There also can be an inconsistency in market data. Remember the market practice Black's swaption pricing formula (2.48). The market price of a swaption depends on the implied volatility and the rates S and K. Since market data consists of **at-the-money** swaption implied volatilities, the rates S and K are the same in the swaption pricing formula and both are retrieved from the market yield curve.
Compare this with market pricing of caps with Black's formula (2.49). Since the market quotes for caps are not all at-the-money, the K does not need to equal F_i. For the separate at-the-money cap market quotes, K is obtained from the market and F_i from the market yield curve.
Hence, a misalignment between the market cap data and market interest rates can result in cap prices that are not totally representative.
One might suggest that calculating cap implied volatilities from the model prices and fitting those volatilities to the market volatilities can overcome this possible inconsistency. But in calculating the model implied volatilities of the caps, market interest rate data is incorporated, and the model volatilities are then compared with the market implied volatilities that might be based on other market interest rates we do not know. Thus the problem is still present. Like for this thesis, the market data should be retrieved from the same provider, but the absence of inconsistency is then still not guaranteed.

4.3.2 Model risk

So far we compared the price ratios per model for several calibrations. This gives information about the impact of the calibration method. Interesting is how calibration risk is in proportion with model risk. To be able to compare these risk classes, the price ratios for the model prices are presented in the following table.

Model price ratios per calibration								
	Swaptions				Caps All		Caps ATM	
	AP	RP	AV	RV	AP	RP	AP	RP
G2/LMM	101.78%	101.90%	101.93%	101.91%	99.84%	118.53%	116.91%	103.27%

The smallest model ratio for swaptions is 101.78%. This is still larger than the largest swaption calibration ratio for the measure of fits: 101.68%. So if one has been convinced that swaptions gives the best information for pricing an excess interest option, then the model risk is larger than the calibration risk.

For calibration to caps, such a comparison does not make sense. The model risk ratios for calibration to caps vary from 99.84% to 118.53%, whereas the calibration risk ratios vary from 96.92% to 117.74%.
If one decides to calibrate to all caps with respect to absolute price differences, then the model risk is also small: 99.84%.

4.3.3 Additional observations

A remarkable fact observed during the calibration process is that different parameters for an interest rate model can lead to almost the same option price.
For example, the Two-Additive-Factor Gaussian parameters calibrated to swaptions AP have a ρ which differs from the other ρs calibrated to swaptions. However, the option price is similar for the four calibrations to swaptions.

A second point that caught the attention is that for the Libor Market Model calibration to swaptions by minimizing AP, simulated annealing came across two parameters that after running downhill simplex resulted in two different minima with χ^2 values that are similar. See this table:

	LMM parameter 1	LMM parameter 2
α_1	0.00482	0.01333
α_2	0.02090	0.02088
σ_1	0.00013	0.00000
σ_2	0.13674	0.13665
ρ	-0.29469	-0.69072
χ^2	1.54E-04	1.54E-04
Option value	11.471	11.454
Standard error	7.48E-03	5.00E-03

This shows that a well performed calibration, i.e. where the obtained parameters minimize the measure of fit on the constrained parameter space, results in a reliable, reconstructible, excess interest option price.

Chapter 5

Conclusions

5.1 Conclusions and recommendations

We have investigated the effect of the choice of the calibration method. We distinguished eight ways to calibrate an interest rate model to the Euro market. Our analysis is based on market data of 30 June 2007. The interest rate models we compared are two-factor models: the Additive-Factor Gaussian model and a two-factor version of Libor Market Model.

We have shown that different calibration instruments lead to significant differences in price of an excess interest option. We have performed the calibration by fitting the model values of swaptions, caps and at-the-money caps to the corresponding market quotes.

Besides the choice of the calibration instruments, we investigated how four different ways to measure the fit between model and market values in the calibration routine affect the price of the option. The measures of fit we regarded for calibration to swaptions are the sum of squared errors of absolute or relative differences between prices and implied volatilities of the calibration instruments in the market and the model. For the calibrations to caps, we focused on absolute and relative differences in price.

Calibration to swaptions leads to similar prices of the excess interest option. Within the calibration to swaptions, two groups can be distinguished. Calibration with respect to absolute prices on one side, and calibration with respect to relative prices, absolute implied volatilities or relative implied volatilities on the other side.
Calibration to swaptions with respect to absolute prices leads to a higher price of the embedded option than calibration to swaptions with respect to the other measures of fit.

However, the differences between these groups within the swaption calibration are negligible in relation to the differences between the swaption calibrations and the calibration to caps.

Calibrating to caps leads to a wide range of option prices. For calibration with respect to relative prices this is caused by the prices of far out-of-the money caps that are almost zero. The fitting errors for these out-of-the-money caps is stressed heavily, calibration of the interest models leads to deviant parameters and to deviant option prices. The at-the-money quotes are too few for calibration purposes.
In general, cap calibration leads to model parameters that imply that a single stochastic factor in the interest rate model would give the same quality of fit of the model prices to the market prices of caps.

If one decides to obtain market information from swaption data, then the model risk is larger than calibration risk. Since swaption data contain information about the correlation between multiple stochastic drives of interest rates, I recommend to use swaptions for calibrating interest rate models. The difference between calibration to swaptions and to caps with respect to absolute prices is relatively small for the value of the excess interest option. If the calibrated models are deployed to run sensitivity analyses on the value of the interest option or deployed to value other interest rate derivatives, the cap calibration could default.
Furthermore, I recommend to use absolute differences. As we saw for calibration to caps, small prices of calibration instruments causes the relative measure of fit to stress these prices more than others.
Finally, I recommend to calibrate to implied volatilities, since this gives approximately equal weight to each entry in the swaption matrix. If one would like to stress a particular part of that matrix, additional weights can be added to those market values.

5.2 Further research

The impact of the calibration routine on the price of an excess interest option has been investigated. The impact of calibration on the calculation of the sensitivity of the excess interest option to market factors like movement of the interest rates or a shock in the implied volatilities of swaptions or caps is not known. These sensitivity analyses become more and more important for insurance companies as those studies give information for the investment strategies that must be followed.

Nowadays, the market volatility skew of forward rates is more commonly given by providers of market data. The volatility skew consists of volatilities for at-the-money swaptions, but also volatilities for swaptions with strikes that are in-or out-of-the-money. Whether calibrating to all these volatilities has effect on the option price or would supply additional information is up till now not totally clear. As we saw for calibration to caps, the more information the better we can fit. Maybe a wide range of volatilities provides the possibility to calibrate three-factor interest rate models.

The functional form of the volatilities in the Libor Market Model we used is derived from Barrie and Hibbert. There are some other forms that are used. It would be interesting to investigate the impact of other functional forms on the price of options and whether another functional form is able to retrieve correlation information from cap market information.

Appendix

Formulas in Two-Additive-Factor Gaussian

For a derivation of these formulas, see [BM01].

Expression for $V(t,T)$

$$\begin{aligned} V(t,T) &= \frac{\sigma^2}{a^2}\left[T - t + \frac{2}{a}e^{-a(T-t)} - \frac{1}{2a}e^{-2a(T-t)} - \frac{3}{2a}\right] \\ &+\frac{\eta^2}{b^2}\left[T - t + \frac{2}{b}e^{-b(T-t)} - \frac{1}{2b}e^{-2b(T-t)} - \frac{3}{2b}\right] \\ &+2\rho\frac{\sigma\eta}{ab}\left[T - t + \frac{e^{-a(T-t)}5 - 1}{a} + \frac{e^{-b(T-t)} - 1}{b} + \frac{e^{-(a+b)(T-t)} - 1}{a+b}\right]. \end{aligned}$$

Expression for $\Sigma(t,T,S)$

$$\begin{aligned} \Sigma(t,T,S) &= \frac{\sigma^2}{2a^3}\left[1 - e^{-a(S-T)}\right]^2\left[1 - e^{-2a(T-t)}\right] \\ &+\frac{\eta^2}{2b^3}\left[1 - e^{-b(S-T)}\right]^2\left[1 - e^{-2b(T-t)}\right] \\ &+2\rho\frac{\sigma\eta}{ab(a+b)}\left[1 - e^{-a(S-T)}\right]\left[1 - e^{-b(S-T)}\right]\left[1 - e^{-(a+b)(T-t)}\right]. \end{aligned}$$

Swaption price in Two-Additive-Factor Gaussian

It can be shown that in the Two-Additive-Factor Gaussian model, the two factors x and y under the forward measure $\mathbb{Q}^T$ evolve according to the following differential equations

$$\begin{aligned} dx(t) &= \left(-ax(t) - \frac{\sigma^2}{a}(1 - e^{-a(T-t)}) - \rho\frac{\sigma\eta}{b}(1 - e^{-b(T-t)})\right)dt + \sigma dW_1^T(t), \\ dy(t) &= \left(-by(t) - \frac{\eta^2}{b}(1 - e^{-b(T-t)}) - \rho\frac{\sigma\eta}{a}(1 - e^{-a(T-t)})\right)dt + \eta dW_2^T(t). \end{aligned}$$

Here, W_1^T and W_2^T are two correlated Brownian motions under $\mathbb{Q}^T$ with $dW_1^T(t)dW_2^T(t) = \rho dt$.
But x and y can be explicitly written as

$$\begin{aligned} x(t) &= x(s)e^{-a(t-s)} - M_x^T(s,t) + \sigma \int_s^t e^{-a(t-u)} dW_1^T(u), \\ y(t) &= y(s)e^{-a(b-s)} - M_y^T(s,t) + \eta \int_s^t e^{-b(t-u)} dW_1^T(u), \end{aligned}$$

where

$$\begin{aligned} M_x^T(s,t) &= (\frac{\sigma^2}{a^2})(1-e^{-a(t-s)}) - \frac{\sigma^2}{2a^2}(e^{-a(T-t)} - e^{-a(T+t-2s)}) - \frac{\rho\sigma\eta}{b(a+b)}(e^{-b(T-t)} - e^{-bT-at+(a+b)s}), \\ M_y^T(s,t) &= (\frac{\eta^2}{b^2})(1-e^{-b(t-s)}) - \frac{\eta^2}{2b^2}(e^{-b(T-t)} - e^{-b(T+t-2s)}) - \frac{\rho\sigma\eta}{a(a+b)}(e^{-a(T-t)} - e^{-aT-bt+(a+b)s}). \end{aligned}$$

Now consider a European option with exercise time T on a swap with strike rate K and reset dates (tenor structure) $\{t_i\}$ with period length δ between the reset dates.
The following formula for the price of a payer swaption is proved in [BM01]:

$$V_{PS}^{G2} = P(0,T) \int_{-\infty}^{\infty} \frac{e^{-\frac{1}{2}\left(\frac{x-\mu_x}{\sigma_x}\right)^2}}{\sigma_x\sqrt{2\pi}} \left[\Phi(-h_1(x)) - \sum_{i=1}^{n} \lambda_i(x) e^{\kappa_i(x)} \Phi(-h_2(x)) \right] dx, \quad (1)$$

where

$$\begin{aligned} h_1(x) &:= \frac{\bar{y} - y}{\sigma_y\sqrt{1-\rho_{xy}^2}} - \frac{\rho_{xy}(x-\mu_x)}{\sigma_x\sqrt{1-\rho_{xy}^2}} \\ h_2(x) &:= h_1(x) + B(b,T,t_i) \\ \lambda_i(x) &:= \delta K A(T,t_i) B(t,T,t_i) \sigma_y \sqrt{1-\rho_{xy}^2} \\ \kappa_i(x) &:= -B(b,T,t_i) \left[\mu_y - \frac{1}{2}(1-\rho_{xy}^2)\sigma_y^2 B(b,T,t_i) + \rho_{xy}\sigma_y \frac{x-\mu_x}{\sigma_x} \right] \end{aligned}$$

in which $\bar{y}$ depends on x and is the unique solution of

$$\sum_{i=1}^{n} \delta K A(T,t_i) e^{-B(a,T,t_i)x - B(t,T,t_i)\bar{y}} = 1.$$

In finding this unique solution, numerical root finding is used, as mentioned in the main text.
Furthermore,

$$\mu_x := -M_x^T(0,T)$$

$$\begin{aligned}
\mu_y &:= -M_y^T(0,T) \\
\sigma_x &:= \sigma\sqrt{\frac{1-e^{-2aT}}{2a}} \\
\sigma_y &:= \eta\sqrt{\frac{1-e^{-2bT}}{2b}} \\
\rho_{xy} &:= \frac{\rho\sigma\eta}{(a+b)\sigma_x\sigma_y}\left(1-e^{-(a+b)T}\right) \\
A(T,t_i) &:= \frac{P^M(0,T)}{P^M(0,t_)}e^{\frac{1}{2}(V(t,T)-V(0,T)+V(0,t))} \\
B(z,t,T) &:= \frac{1-e^{-z(t-t)}}{z}.
\end{aligned}$$

Notice that the functions A and B are defined such that

$$P(t,T) = A(t,T)e^{-B(a,t,T)x(t)-B(b,t,T)y(t)}.$$

Monte Carlo Simulation

Suppose we have a general d-dimensional stochastic process x which satisfies the stochastic differential equation

$$dx(t) = \mu(t)dt + \sigma(t)dW(t), \tag{2}$$

under some measure $\mathbb{Q}$. Here μ and σ are functions of time t and possibly of $x(t)$. We discretize (2) to obtain the Euler forward scheme

$$x(t+\Delta t) - x(t) = \mu(t)\Delta t + \sigma(t)[W(t+\Delta t) - W(t)]. \tag{3}$$

This approximation only makes sense for small values of Δt.
The Brownian motion W has independent increments $[W(t+\Delta t) - W(t)]$ which are normal distributed with mean zero and variance Δt. One path of the stochastic process from time 0 to T can be simulated by splitting the intervap $[0,T]$ into small intervals $[t_i, t_{i+1}]$, where $t_i = i\Delta t$ for $i = 1, 2, \ldots$. Then the path of the Brownian motion W is simulated according

$$W(t_{i+1}) = W(t_i) + \sqrt{\Delta t}\epsilon_i, \tag{4}$$

where ϵ_i are independent standard normal variates.

With this simulation of W we can construct x. An expectation of the payoff $V(x)$ based on the path of x can be approximated by

$$\mathbb{E}^{\mathbb{Q}}[V] \approx \bar{V}, \tag{5}$$

where

$$\bar{V} = \frac{1}{N}\sum_{j=1}^{N} V_j, \tag{6}$$

with N the number of simulated paths and V_j the payoff corresponding to simulation j. From the central limit theorem, random variable $\bar{V}$ converges to a normal distribution with mean $\mathbb{E}^{\mathbb{Q}}[V]$ and variance $\mathrm{Var}^{\mathbb{Q}}(V)/N$ as N tends to infinity.

The *standard error* of the Monte Carlo simulation is a measure of the accuracy of the approximation. It is defined as:

$$\mathrm{SE} = \sqrt{\frac{\sum_{j=1}^{N} V_j^2 - N\bar{V}^2}{N(N-1)}}. \tag{7}$$

After a simulation the value of the standard error gives insight in the quality of the approximation.

Simulation of Two-Additive-Factor Gaussian

Risk neutral Monte Carlo simulation of the Two-Additive-Factor Gaussian model is straightforward. The differential equations for $x(t)$ and $y(t)$ are given by (2.36). The Euler scheme for annual time-steps of this stochastic equation is thus:

$$\begin{aligned} x_n &= x_{n-1} - a x_{n-1} + \sigma\epsilon_1 \\ y_n &= y_{n-1} - b y_{n-1} + \eta\epsilon_2 \end{aligned}$$

Here $\epsilon_1 \sim N(0,1)$ and $\epsilon_2 = \rho\epsilon_1 + \sqrt{(1-\rho^2)}\epsilon'$ with $\epsilon_2' \sim N(0,1)$, independent from ϵ_1.

During the generation of the process x and y, at each time-step i, the values $P(i,T)$ are calculated according (2.51), for relevant $T > i$.
The relevant Ts are $T = i+1, \ldots, i+8$, since $P(i, i+1)$ is needed for the cash-rollup ans the other bond values up to $P(i, i+8)$ are used for the u-returns.

Simulation of Libor Market Model

Monte Carlo simulation of the Libor Market Model to price interest rate derivatives is done under the risk neutral measure, according (2.44). With Ito's formula, we have for $\log F_n$:

$$d\log F_n(t) = \left[\sum_{j=m(t)}^{n} \frac{\delta F_j(t)\gamma_j(t)\gamma_n(t)}{1+\delta F_j(t)} - \frac{1}{2}\gamma_n^2(t)\right] dt + \gamma_n(t)dW(t). \tag{8}$$

To make things easier, the discretization is done at times $T_0, T_1, \ldots, T_N$. We get that

$$\begin{aligned} d\log F_n(T_{i+1}) &= d\log F_n(T_i) + \left[\sum_{j=i}^{n} \frac{\delta F_j(T_i)\gamma_j(T_i)\gamma_n(T_i)}{1+\delta F_j(T_i)} - \frac{1}{2}\gamma_n^2(T_i)\right]\delta \\ &\quad + \gamma_n(T_i)(W(T_{i+1}) - W(T_i)). \end{aligned}$$

This results in the following iteration scheme for the forward rates:

$$\begin{aligned} F_n(T_{i+1}) &= F_n(T_i)\exp\left\{\left(\sum_{j=i}^{n} \frac{\delta F_j(T_i)\sum_{q=1}^{2}\Lambda_{j-i+1}^{q}\Lambda_{n-i+1}^{q}}{1+\delta F_j(T_i)} - \frac{1}{2}\sum_{q=1}^{2}\Lambda_{n-i+1}^{q}\right)\delta\right. \\ &\quad \left. + \sum_{q=1}^{2}\Lambda_{n-i+1}^{q}{}^{2}\sqrt{\delta}\epsilon_q\right\}, \quad i = 1, \ldots, N-1,\ n = i+1, \ldots, N, \qquad (9) \end{aligned}$$

where ϵ_1 and ϵ_2 are independent standard normal random variables.
With these generated forward rates, we can generate the zero bond values for the cash-rollup and to determine the u-returns.

Bibliography

[BH00] Barrie and Hibbert. A Stochastic Model for the Term Structure of Interest Rates and Equity Returns, 2000.

[Bjö98] Tomas Björk. *Arbitrage Theory in Continuous Time*. Oxford University Press, New York, 1st edition, 1998.

[BM95] S.P. Brooks and B.J.T. Morgan. Optimization Using Simulated Annealing. *The Statistician*, 44(2):241–257, 1995.

[BM01] Damiano Brigo and Fabio Mercurio. *Interest Rate Models*. Springer, Berlin, 2001.

[DH07] K. Detlefsen and Wolfgang K. Härdle. Calibration Risk for Exotic Options. *The Journal of Derivatives*, 5:47–63, 2007.

[Haü83] W.M. Haüßler. A Local Convergence Analysis for the Gauss-Newton and Levenberg-Morrison-Marquardt Algorithms. *Computing*, 31:231–244, 1983.

[Hul00] John C. Hull. *Options, Futures & Other Derivatives*. Prentice-Hall International, Upper Saddle River, New Jersey, 4th edition, 2000.

[HW90] John Hull and Alan White. Pricing Interest-Rate-Derivative Securities. *The Review of Financial Studies*, 3(4):573–592, 1990.

[HW94] John Hull and Alan White. *The Journal of Derivatives*, 2:7–16, 1994.

[HW99] John Hull and Alan White. Forward rate volatilities, swap rate volatilities, and the implementation of the Libor Market Model, 1999.

[Ing93] Lester Ingber. Simulated Annealing: Practice versus Theory. *Mathematical and Computer Modelling*, 18(11):29–57, 1993.

[Jam89] Farshid Jamshidian. An Exact Bond Option Formula. *The Journal of Finance*, 44(1):205–209, 1989.

[Jam97] Farshid Jamshidian. LIBOR and swap market models and measures. *Finance and Stochastics*, 1:293–330, 1997.

[JDP01] Frank de Jong, Joost Driessen, and Antoon Pelsser. Libor Market Model versus Swap Market Models for Pricing Interest Rate Derivatives: An Emperical Analysis. *European Finance Review*, 5:201–237, 2001.

[Kes07] E.G.H. Kessels. Valuation of Profit Sharing Contracts under Stochastic Interest Rate Models. Master's thesis, Tilburg University, 2007.

[KGV83] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220:671–680, 1983.

[Loc00] M. Locatelli. Simulated Annealing Algorithms for Continuous Global Optimization: Convergence Conditions. *Journal of Optimization Theory and Applications*, 104(1):121–133, 2000.

[Loc02] M. Locatelli. Simulated Annealing Algorithms for Continuous Global Optimization, 2002.

[Mar63] D.W. Marquardt. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11(2):431–441, 1963.

[PCB02] C.J. Price, I.D. Coope, and D. Byatt. A convergent variant of the Nelder-Mead Algorithm. *Journal of Optimization Theory and Applications*, 113(1):5–19, 2002.

[PFTV92] William H. Press, Brian P. Flannery, Saul A. Teukolsky, and William T. Vetterling. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, 2nd edition, 1992.

[Vas77] O. Vasicek. An Equilibrium Characterization of the Term Structure. *Journal of Financial Economics*, 5:177–188, 1977.

[VPTZ01] L. Veláquez, G.N. Philips, R.A. Tapia, and Y. Zhang. Selective Search for Global Optimization of Zero or Small Residual Least-Squares Problems: A Numerical Study. *Computational Optimization and Applications*, 20:299–315, 2001.

[Wil05] D. Williams. *Probability with Martingales*. Cambridge University Press, 9th edition, 2005.