Univerzita Karlova v Praze
Fakulta sociálních věd

Institut ekonomických studií

Rigorózní práce

2007

Jan Doubrava

Univerzita Karlova v Praze
Fakulta sociálnich věd

Institut ekonomických studií

RIGORÓZNÍ PRÁCE

Pricing of Interest Rate Derivatives and Calibration Issues in a Multi-Factor LIBOR Market Model Framework

Vypracoval: Mgr. Jan Doubrava
Vedoucí: doc. Ing. Oldřich Dědek CSc.
Akademický rok: 2006/2007

Prohlášení

Prohlašuji, že jsem rigorózní práci vypracoval samostatně a použil pouze uvedené prameny a literaturu

V Praze dne podpis studenta

ABSTRAKT

Finanční deriváty jsou finanční instrumenty jež umožňují investorům i dlužníkům optimalizovat svá portfolia podle individuálních potřeb a míry akceptovatelného rizika. Jejich význam na finančních trzích v posledních deseti letech enormně stoupl a objemy zobchodovaných instrumentů nepřestávají růst. Deriváty úrokových měr tvoří velkou podskupinu, jejich oceňování tvoří ve finanční matematice díky zvláštním charakteristikám dynamiky výnosové a diskontní křivky samostatnou kapitolu. Tato práce se v první části zabývá základními principy oceňování derivátů úrokových měr vycházející z teorie o bezarbitráži a představením nejběžnějších modelů dynamiky výnosové křivky. Druhá část se zabývá otázkou kalibrace v rámci "LIBOR Market Modelu" s jedním až třemi faktory rizika. Tyto tři modely jsou použity k ocenění swapcí pomocí Monte Carlo simulace v rámci teorie o bezarbitráži představené v první části. Výsledkem práce je zjištění, že nejlépe jsou swapce oceněny pomocí modelu s pouze jedním faktorem rizika.

ABSTRACT

Financial derivatives are financial instruments which enable investor or a debtor to optimize his/her asset/debt portfolios according to individual needs and acceptable scale of risk. Their importance in financial markets rose enormously n past ten years as well as did their traded volumes. Interest rate derivatives form a large sub-group of financial derivatives, their valuation is a large self-contained chapter within financial mathematics thanks to the unique characteristics of yield- and discount-curve dynamics. In the first part of my thesis I derive the fundamental pricing principles stemming from no-arbitrage pricing theory and introduce the most common approaches in yield curve modeling. In the second part I discuss issues of calibration in a "LIBOR Market Model" with one to three risk factors. These models are used to price swaptions with Monte Carlo simulation within the no-arbitrage framework introduced in the first part. The result of the thesis is that one factor model performs the best in pricing swaptions.

Contents

1 **Introduction** **3**

2 **Financial Derivatives and No-Arbitrage Pricing** **5**
- 2.1 Dynamics of financial assets . . . 5
- 2.2 Deriving the no-arbitrage condition . . . 7
- 2.3 Deriving the fundemental Black-Scholes PDE . . . 8
- 2.4 Feynman-Kac stochastic representation formula . . . 9
- 2.5 Change of probability measure and fundamental pricing formula 10

3 **Definitions and Notations** **13**
- 3.1 Term Structure of Interest Rates . . . 13
  - 3.1.1 Money-market account . . . 13
  - 3.1.2 Zero-cupon bond . . . 13
  - 3.1.3 Spot interest rates . . . 13
  - 3.1.4 Interest rate curves . . . 15
  - 3.1.5 Forward interest rates . . . 15
- 3.2 Basic interest rates derivatives . . . 17
  - 3.2.1 Forward rate agreement - FRA . . . 17
  - 3.2.2 Interest rate swaps - IRS . . . 18
  - 3.2.3 Swap rate . . . 18
  - 3.2.4 Caps and Floors . . . 20
  - 3.2.5 Swaptions . . . 21

4 **Interest Rate Models** **23**
- 4.1 Short Rate Models . . . 24
  - 4.1.1 Dynamics under the Real Measure and Market Price of Risk . . . 24
  - 4.1.2 Dynamics under the Risk Neutral Measure . . . 25
  - 4.1.3 Examples of short rate dynamics . . . 25
  - 4.1.4 Pricing Derivatives and some Disadvantages of Short Rate Models . . . 27
- 4.2 Heath, Jarrow and Morton Framework . . . 28
  - 4.2.1 Dynamics under the Real Measure . . . 28
  - 4.2.2 Dynamics under the Risk-Neutral Measure . . . 29
  - 4.2.3 Model specification . . . 29
  - 4.2.4 Pricing Derivatives using HJM . . . 30
- 4.3 Market Models . . . 30
  - 4.3.1 LIBOR Market Model . . . 30
  - 4.3.2 Swap Market Model . . . 33
  - 4.3.3 Compatibility of LIBOR and Swap Market Models . . . 34

5 **Calibrating LIBOR Market Model** **35**
- 5.1 The Model . . . 35
  - 5.1.1 One Factor Case . . . 36
  - 5.1.2 N Factor case . . . 36
  - 5.1.3 Reducing Dimensionality of the Model . . . 37
- 5.2 Volatility Function of Forward Rates . . . 40
- 5.3 Correlations of Forward Rates . . . 47

5.3.1 Correlations in 2 Factor Model 47
5.3.2 Correlations in 3 Factor Model 49

6 Pricing Swaptions with Monte Carlo Simulation 53
6.1 Market Quotes of Swaptions 53
6.2 Discretization of the Forward Rate Dynamics 54
6.3 Simulations 55
6.4 Results 57

7 Conclusion 61

8 Literature 62

9 Appendix A - Ito's Lemma 64

10 Appendix B - Feynman-Kac's stochastic representation formula 66

11 Appendix C - Caplet's Vega 68

1 Introduction

The popularity of interest rate derivatives has dramaticaly increased in past couple of years. To a great extend this was possible because of big improvements in financial modelling in the last decade. New and revolutionary concepts that were developed (such as the Heath, Jarrow and Morton framework and especially the market models) provided new insights into understanding of yield curve dynamics, which enabled the banks to develop sophisticated evaluating, trading and hedging mechanisms. Thanks to this progress new product ideas could be realized and they met with big popularity among customers. Thanks to various structured swaps the customers from the liability sector can manage their debt more actively, which is the case of not only big and strong companies, but also of sovereign bodies like cities and states. On the other side of the customer spectrum, asset products enable big fonds, insurance companies and other investors to buy more interesting fixed income products where they can bet on certain market view about future developement of interest rates and thus enhance their profit.

To make an example, a company which has to pay fixed payments for it's debt can enter into so called range accrual swap with a third party. It pays a low fixed and receives a portion of higher rate, which (portion) depends on the amount of days that say a 3-month Euribor fixed on daily basis lies within a specified range. If the company is correct about the range, the portion is near one and it generates profit which helps lowering its debt costs.

Other example would be an investor, who buys a bond which can be called by the issuer if rates fall. Callable bond is also paying higher cupon as the investor basically sells right to cancel the trade if rates are developing in an unfavourable way for the issuer. Such trade involves three parties, an investor, an issuer and a counterparty for a swap, say an investment bank. Issuer receives investor's money and deposits them on an interbank market for a floating rate. These floating payments are swaped with an investment bank for a fixed payment which is passed on a regulary basis to an investor. In addition, the issuer sells an investment bank a swaption, or a right to enter into a swap which provides exactly the opposite flows of payments as the original swap, i.e. issuer would be obliged to exchange a fix rate for a floating rate if such swaption was exercised. If a swaption is exercised, the payments of the two swaps cancel out (bond is called). The sold option created a value to the issuer which translates into a higher fixed payments which he passes (with a discount) to an investor. In such scenario the issuer is basically an intermediary and it is the investment bank who is canceling the contract.

In both examples there are interest rate derivative products involved and the extend to which the deal is favourable for the buyer of the product depends on the derivative's price. Both example demonstrate that there is a big motivation and economic interest in a knowledge of pricing these derivative products. Specificaly, in the second example we would be interested to know what is the value of the sold swaption. In order to find out we have to make an assumption about the dynamics of the yield curve. A model has to be built which simulates possible evolutions of the interest rates so that we had an idea about the probabilities, with which such scenario occures where a swaption that we sold will be excercised. How exactly the model should be constructed and how they value of derivative product should be calculated is a topic of this thesis.

In the first section, no-arbitrage pricing mechanisms will be introduced to provide a framework for pricing derivatives. In the second section, various interest rates will be defined and notation will be introduced which will be used used throughout the thesis. The third section introduces various interest rate models that can be considered for yield curve simulations. It will be shown how the market prices (quotes) the most common interest rate derivatives, namely swaptions or options on a swap rate and caps or options on interest rate. It will be argued, that so called market models are very convenient for pricing products like caps and options because they aim to model directly the underlying quantitie, i.e. the swap rate in case of a swaption and a Libor rate in case of a cap.

Finally, in the last two sections a Libor market model will be built and calibrated with help of both market and historic data. It will be subsequently used for yield curve simulations which will allow to price swaptions on basis of the introduced no-arbitrage pricing framework. Furthermore, three models will be used to price swaptions, models with one, two and three risk factors. A hypothesis will be tested, wheather the three factor model which captures the evolution of a yield curve in the most realistic way, delivers the best results in pricing swaptions.

2 Financial Derivatives and No-Arbitrage Pricing

Derivatives are financial instruments, who's payoff at certain time in the future depends on an evolution of some underlying. Such underlying can be a tradeable asset such as stock or bond or it can be a non-tradable variable such as interst rate, swap rate but also amount of rain in a year. A typical exampe of a derivative instrument is a call option. It is defined by an underlying asset, say S, strike price K and expiry time T. The payoff of a call option at expiry time T is the following function of a value of it's underlying asset S at time T:

$$\max\{S(T) - K, 0\}. \tag{1}$$

In order to determine the price of such contract it is necessary to examine the charactaristics of the underlying asset's dynamics. It is intuitive, that at current time $t < T$ for a given strike price K (let's say $K >> S(t)$) the call option's value will be higher if the volatility of the underlying asset is high and therefore if chances are greater that the underlying's value will exceed the strike price at option's expiry time T. As will be shown later, the underlying's volatility plays an essential role in determining option's price.

2.1 Dynamics of financial assets

The following formula describes a general dynamics of an underlying, say stock S:

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

It can interpreted as follows: an infinitezimal change of stock S at time t consist of a deterministic term μdt called *drift* and a stochastic term $\sigma dW(t)$ called *diffusion*. $\sigma(t, S(t))$ is then *volatility function* of the process and both μ and σ are functions of time and stock itself (meaning the drift as well as volatility depend on the level of S). Finally, $dW(t)$ is an infinitezimal change of a stochastic process called *Wiener Process*[1] which is normally distributed with zero expected value and variance of dt under the real world probability measure - P. This is the risk factor entailed in the equation of underlying's dynamics.

The first step in pricing a derivative is to specify the form and parameteres of the in derivative's underlying, especially (as will be shown later on) those of the volatility function. The key point which leads to the famous Black and Scholes[2] formula is to assume a linear form of drift and volatility functions,

$$dS(t) = \mu S(t)dt + \sigma S(t)dW(t). \tag{3}$$

Here μ and σ are constant coeffitients. The equation can be solved for $S(T)$ by showing, that the process of logarithm of $S(t)$ can be written as[3]

$$dln(S(t)) = (\mu - \frac{1}{2}\sigma^2)dt + \sigma dW(t). \tag{4}$$

[1]See Björk (2004).
[2]See Black and Scholes (1973).
[3]See Ito's Lemma in text below.

Now there is no state variable $S(t)$ on the right side and the process can thus be integrated to obtain value of $ln(S)$ at time T,

$$ln(S(T)) = ln(S(t)) + \int_t^T \left(\mu - \frac{1}{2}\sigma^2\right)dt + \int_t^T \sigma dW(t). \tag{5}$$

Because the integrated functions are constants, the first integral is equal $(\mu - 0.5\sigma^2)(T-t)$ and the second is equal $\sigma W(T)$. Taking exponentials of both sides we obtain:

$$S(T) = S(t)e^{(\mu-0.5\sigma^2)(T-t)+\sigma W(T)} \tag{6}$$

where we assume that $W(t)$ is zero. Because the exponent of 6 follows a normal distribution ($W(T)$ is distributed normaly), by deffinition, stock $S(T)$ is distributed *log-normaly*, which is the crucial assumption for evaluating options in Black-Scholes scenario. Further, the process describing dynamics of stock - equation 3 - is called *Geometrical Brownian Motion* or simply *GBM*. The coeffitient μ can than be interpreted as a local rate of return. This can be easily understood dividing equation 3 by $S(T)$:

$$\frac{dS(t)}{S(t)} = \mu dt + \sigma dW(t).$$

Here it is obvious that the fraction $dS(t)/S(t)$ is to be interpreted as a percentage growth. The expected infinitezimal percentage growth of stock S is μdt, because the expected value of Wiener process is zero. A care must be however taken here, we cannot substitute $dS(t)/S(t)$ by $dlnS(t)$, for process $S(t)$ is stchochastic and Ito's Lemma must be used as shown below. Such process with only deterministic coeffitients is called *Aritmetical Brownian Motion* or *ABM* and it is normally distributed.

Here a question can be raised if indeed the underlying stock is log-normally distributed implying that the returns are distributed normally, but that is a subject of different discussion. At this point we can however ask a question of how to estimate the volatility parameter in equation 3 - σ.[4] Basicaly we have two choices, either we can estimate volatility from historical data or we can accept a guess of our own, of some expert or that of a market. In the second part of the thesis it will be shown how the markets' assumption about volatility can be extracted from market data and how it plays a key role in model calibration.

Now let's consider such capital market which consists of n traded securities with one common risk factor $W(t)$ and a possibility to deposit money in a bank at a constant risk-free interest rate r.[5] If we assume that all securities follow *GBM*, buying any one of these, our investment at time t will have the following dynamics:

$$dS_i(t) = \mu_i S_i(t)dt + \sigma_i S_i(t)dW(t), \quad \text{for} \quad i = 1, ..., n. \tag{7}$$

[4]Later in the thesis it will be shown that the drift parameter μ doesn't play in role in derivatives evaluation.

[5]This assumption is made for simplicity, risk-free interest rate can also be made stochastic which wouldn't change anything in the agumentation below.

If instead we decide to deposit $B(t)$ units of money into a bank account, our investment will accrue according to the differential equation

$$dB(t) = B(t)rdt, \tag{8}$$

where $B(t)$ stands for *money market account.* We can see that in equation 8 there appears no stochastic term which means that the dynamics of money market account is deterministic or risk-less. It is easy to see that if the initial investment at time t_0 is *one*, the balance of our deposit in $T > t_0$ will be:

$$B(T) = e^{r(T-t_0)}. \tag{9}$$

This can be obtained by solving the diffenrential equation 8 with initial condition $B(t_0) = 1$.

2.2 Deriving the no-arbitrage condition

Now I will show that assuming only one risk factor (the same Wiener process $W(t)$ is present in dynamics of all traded assets), we can take any two of the n securities and find the unique quantities needed to form a portfolio, which will be localy riskless. The dynamics of securities S_i and S_j at current time t will be:

$$dS_i(t) = \mu_i S_i(t)dt + \sigma_i S_i(t)dW(t)$$
$$dS_j(t) = \mu_j S_j(t)dt + \sigma_j S_j(t)dW(t)$$

and buying quantities $\alpha(t)$ and $\beta(t)$ of securities S_i and S_j the value of our portfolio $V(t)$ at current time will be:

$$V(t) = \alpha(t)S_i(t) + \beta(t)S_j(t). \tag{10}$$

Knowing dynamics of stochastic variables X_1 and X_2 one can derive dynamics of their function $Y(X_1, X_2)$ using the famous *Ito's Lemma*:

$$\begin{aligned} dY(X_1, X_2) &= \frac{\partial Y}{\partial X_1}dX_1 + \frac{\partial Y}{\partial X_2}dX_2 + \frac{1}{2}\frac{\partial^2 Y}{\partial X_1^2}vol(dX_1) + \frac{1}{2}\frac{\partial^2 Y}{\partial X_2^2}vol(dX_2) + \\ &\quad + \frac{\partial^2 Y}{\partial X_1 \partial X_2}\sqrt{vol(dX_1)}\sqrt{vol(dX_2)}. \end{aligned} \tag{11}$$

Appendix A proves Ito's Lemma for one variable case, two- and n-variable case is analogous.

We follow formula 11 to derive dynamics of our portfolio V in equation 10. Because V is linear in Ss, second derivations are zero and we obtain the following expression:

$$dV(t) = \alpha(t)dS_i(t) + \beta(t)dS_j(t).$$

Substituting for $dS_i(t)$ and $S_j(t)$ and putting dt and $dW(t)$ terms together we obtain:

$$\begin{aligned} dV(t) &= [\alpha(t)S_i(t)\mu_i + \beta(t)S_j(t)\mu_j]\,dt + \\ &\quad [\alpha(t)S_i(t)\sigma_i + \beta(t)S_j(t)\sigma_j]\,dW(t). \end{aligned} \tag{12}$$

This stochastic diferential equation says that the dynamics of our portfolio V is dependent on the level of S_i and S_j at time t and on the quantities α and β. We can however choose

$$\alpha(t) = \frac{-\sigma_j}{(\sigma_i - \sigma_j)S_i(t)} \quad \text{and} \quad \beta(t) = \frac{\sigma_i}{(\sigma_i - \sigma_j)S_j(t)}$$

and substituting these quantities into dynamics 12 we obtain:

$$dV(t) = \frac{\sigma_i \mu_j - \sigma_j \mu_i}{\sigma_i - \sigma_j} dt. \tag{13}$$

We can see that the above choice of quantities α and β causes that the $W(t)$ factor falls out which makes the above portfolio dynamics riskless just like the money market account 8 and to prevent presence of arbitrage, the drift parameters of dynamics 8 and 13 must be set equal. Rearranging we obtain:

$$\begin{aligned} \frac{\sigma_i \mu_j - \sigma_j \mu_i}{\sigma_i - \sigma_j} &= r \\ \sigma_i \mu_j - r\sigma_i &= \sigma_j \mu_i - r\sigma_j \\ \frac{\mu_j - r}{\sigma_j} &= \frac{\mu_i - r}{\sigma_i} = \lambda. \end{aligned} \tag{14}$$

This fundamental result can be interpreted as follows: higher excess return over risk-less interest rate must be compensated by higher volatility. The resulting ratio λ is called *market price of risk* and if the market is free of arbitrage, it must be equal for all traded securities.

2.3 Deriving the fundemental Black-Scholes PDE

Let us now return to our call-option C and express its dynamics at time t with the following general differential equation:

$$dC(t) = \mu_C(t, C(t))dt + \sigma_C(t, C(t))dW(t). \tag{15}$$

If we assume that this call-option is as well traded on the market, the above no arbitrage condition 14 must apply to our option's dynamics as well:

$$\frac{\frac{\mu_C(t,C(t))}{C} - r}{\frac{\sigma_C(t,C(t))}{C}} = \lambda. \tag{16}$$

Further we know that the payoff of the option C at the expiry time is a function of its underlying, say security S_i and therefore the options's current price must as well be a function of this underlying's current price - *C(t,S_i(t))*. Assuming the underlying's dynamics 7 we can rewrite options dynamics using Ito's Lemma as:

$$dC(t, S_i(t)) = \frac{\partial C}{\partial t} dt + \frac{\partial C}{\partial S_i} dS_i(t) + \frac{1}{2}\frac{\partial^2 C}{\partial S_i^2} vol(dS_i(t)).$$

This result is easily obtained substituting into equation 11 and realizing, that time variable t is deterministic and therefore has *zero* variance which causes the second derivatives with t to fall out. Now we can substitute for S_i and after rearranging we obtain:

$$dC(t, S_i(t)) = \left[\frac{\partial C}{\partial t} + \frac{\partial C}{\partial S_i}\mu_i S_i(t) + \frac{1}{2}\frac{\partial^2 C}{\partial S_i^2}\sigma_i^2 S_i^2(t)\right] dt + \frac{\partial C}{\partial S}\sigma_i S_i dW(t). \quad (17)$$

It can be seen that the drift function $\mu_C(t, C(t))$ from the general option's dynamics equation 15 is equal to the term in square brakets in the above equation 17 and the volatility function $\sigma_C(t, C(t))$ is equal to the expression before $dW(t)$ term in equation 17. If we substitute these into equation 16 which guarantees absence of arbitrage, after rearranging we obtain the following expression:

$$Cr = \frac{\partial C}{\partial t} + \frac{\partial C}{\partial S_i} S_i [\mu_i - \lambda \sigma_i] + \frac{1}{2}\frac{\partial^2 C}{\partial S_i^2}\sigma_i^2 S^2.$$

We can notice that in absence of arbitrage (equation 14) the term in square brackets can be substituted by r and adding the payoff function of a call option at maturity T (equation 1) as a boundary condition we obtain the famous Black-Scholes option pricing equation:

$$C(t, S(t))r = \frac{\partial C}{\partial t} + \frac{\partial C}{\partial S_i} r S_i(t) + \frac{1}{2}\frac{\partial^2 C}{\partial S_i^2}\sigma_i^2 S_i^2(t) \quad (18)$$

$$C(T, S_i(T)) = max\,[S_i(T) - K]\,. \quad (19)$$

We can be however quite general as far as the boundary condition 19 is concerned and replacing the call option payoff at its expiry time T we can price any derivative who's payoff is a function of its underlying at T:

$$C(T, S_i(T)) = \phi(S(T)). \quad (20)$$

2.4 Feynman-Kac stochastic representation formula

Solving the above differential equation 18 with boundary condition 19 or 20 gives us a price of a derivative at time t assuming lognormal distribution of its underlying S_i. One way of solving equation 18 is analytically by lengthy computations. The other way is employing a *Feynman-Kac stochastic representation formula*,[6] which says, that having a boundary value problem

$$F(t, X(t))r = \frac{\partial F}{\partial t}(t, X(t)) + \frac{\partial F}{\partial X}\mu(t, X(t)) + \frac{1}{2}\frac{\partial^2 F}{\partial X^2}\sigma^2(t, X(t)) \quad (21)$$

$$F(T, X(T)) = \phi(X(T)), \quad (22)$$

where some integrability conditions for $\frac{\partial F}{\partial X}\sigma(t, X(t))$ are fulfilled[7] and where dynamics of X is expressed by stochastic differential equation

$$dX(t) = \mu(t, X(t))ds + \sigma(t, X(t))dW(t),$$

then the solution to problem 21 - 22 is

$$F(t, X(t)) = e^{-r(T-t)} E\left[\phi(X(T)\right]. \quad (23)$$

The proof is presented in Appendix B.

[6] See for instance Björk (2004).

[7] $\int_0^t E\left[(\frac{\partial F}{\partial X}\sigma(s, X(s)))^2\right] ds < \infty$, see Björk (2004).

2.5 Change of probability measure and fundamental pricing formula

We can see that our task to solve derivative-pricing equation 18 is almost identical to the problem presented in Feynman-Kac's formula, except for the drift-function $\mu(t, S_i(t))$ in the underlying's dynamics 7, which doesn't correspond to $rS_i(t)$ from eqn. 18, but to $\mu_i S_i(t)$. A very handy fix is provided by *Girsanov Theorem*,[8] which allows us to change the probability measure which governs the process $S_i(t)$ (or $W(t)$ more precisely) in such way, that the drift changes according to our needs while the diffusion remains the same. We can thus define a probability measure Q such, that:

$$dS_i(t) = rS_i(t)dt + \sigma_i S_i(t)d\widehat{W}(t) \quad \text{for} \quad i = 1, ..., n, \tag{24}$$

where $d\widehat{W}(t)$ is a Wiener proces under newly defined measure Q, which is called *equivalent probability measure*[9] to measure P. Two important points should be made here. First it is necessary to realize that the new process 24 doesn't describe real behaviour of securitity S_i, instead it is an artificial auxiliary process to help us solve equation 18. Second, it can be shown, that the two Wiener processes $dW(t)$ and $\widehat{W}(t)$ are linked by the following symbolical relationship:[10]

$$dW(t) = d\widehat{W}(t) - \lambda dt = d\widehat{W}(t) - \frac{\mu_i - r}{\sigma_i}dt, \tag{25}$$

where λ is the market price of risk from equation 14. Therefore by substituting relationship 25 for $dW(t)$ into equation 7 for any security i, the drift term $\mu_i S_i dt$ falls out and we obtain a new drift term $rS_i dt$ as in equation 24. This can be interpreted so, that under the probability measure Q, all traded instruments have the same drift r if the market is free of arbitrage. The Q measure is also called a *risk-neutral measure* and the fact that its existence guarantees absence of arbitrage was shown by Harrison and Pliska.[11]

Having switched from measure P to measure Q our problem to solve the derivative-pricing equation becomes identical with the Feynman-Kac formula and we obtain the following result as a solution to equation 18 with boundary condition 20:

$$C(t, S(t)) = e^{-r(T-t)}\widehat{E}\left[\phi(S(T))\right], \tag{26}$$

or

$$\frac{C(t, S(t))}{B(t)} = B(T)^{-1}\,\widehat{E}\left[\phi(S(T))\right], \quad B(t_0) = 1,$$

where $\widehat{E}$ means expected value under the risk-neutral measure Q and $B(T)^{-1}$ is the inverse of money market account with initial investment of *one* (equation 9) which can be interpreted as a continuous deterministic discounting factor.

[8] See for instance Brigo and Mercurio (2001).

[9] Two measures are equivalent, if they share te same sets of null probability. See for instance Brigo and Mercurio (2001).

[10] See for instance Brigo and Mercurio (2001).

[11] See Harrison and Pliska (1981).

This important result says that if the market is free of arbitrage, the unique price of derivative C at any time $t < T$ is equal to the discounted risk-neutral expectations of the instrument's payoff at expirty T.

In order to achieve the solution 26 it was assumed, that the underlying of a derivative C follows a *GBM* and the risk-less rate of return r over time $T - t$ is constant. Harrison and Pliska have however prooved, that this result can be generalized to the case where we allow the risk-less interest rate r to be stochastic and at the same time loosen the lognormality-of-underlying assumption. The price of a derivative is than expressed by:

$$\frac{C(t, S(t))}{B(t)} = \widehat{E}\left[\frac{\phi(S(T))}{B(T)}\right] \quad , \quad B(t_0) = 1. \tag{27}$$

This result says that using money market account as a nummeraire, the discounted price of a derivative C at time t is equal to the risk-neutral expectation of its discounted payoff at its expiry T. We can say that the discounted process (by money market account) of a derivative C is a *martingale* under Q or that it is equal to it's Q-expectations.[12]

Let's repeat at this point, that the Q-probability measure was defined so, that any traded instrument S_i had a local rate of return equal r. We can equivalently define this Q-measure by requiring, that the process of any discounted security (by money market account) which is traded on market is Q-martingale. As mentioned before, the existence of such measure Q is a guarantee, that the market is free of arbitrage. This definition will be used in the following paragraph.

As will be shown later in the thesis, it is often handy to use different nummeraire than the money market account $B(t)$. We can define a Q_X measure so, that the process of any discounted (traded) security is a martingale under Q_X, when security X is used as a nummeraire.[13] Additionaly, we have to require that this nummeraire X is a non-negative asset with no intermediate payments during the life of evaluated derivative. This is expressed by the following fundamental condition:

$$\frac{Y(t)}{X(t)} = E_X\left[\frac{Y(T)}{X(T)}\right],$$

where Y is any traded instrument and E_X denotes expected value under Q_X-measure. In case we derive a price of a derivative, we obtain a derivative-pricing condition

$$\frac{C(t, S(t))}{X(t)} = E_X\left[\frac{\phi(S(T))}{X(T)}\right]. \tag{28}$$

Analogous to 27 the interpretation of this result is, that under absence of arbitrage on the market, a unique price of a derivative $C(t)$ discounted be the known value at t of nummeraire X can be assessed as discounted expectations under measure Q_X of the derivative's payoff at its maturity T. Furthermore,

[12]Simplified definition of martingale is $X(t_0) = E[X(T)]$, under the relevant probability measure. For full definition including necessary assumptions see Bjork (2004).

[13]See Brigo and Mercurio (2004).

the process of any tradable security discounted by a nummeraire X must be a martingale under Q_X should the market be free of arbitrage. This result will be extensively used in the remainder of the thesis and will be refered to as a *fundamental pricing formula.*

3 Definitions and Notations

Having introduced the fundementals of stochastics of financial assets and no-arbitrage pricing in the previous section, this section will go on with basic definitions, showing different ways of defining the term structure of interest rates and the most basic interest rate derivatives.

3.1 Term Structure of Interest Rates

3.1.1 Money-market account

First definition is that of a *money-market account* or a *bank account* mentioned already in the first section, repeated here for completeness. The investment of one unit of currency at time 0 will at time t have the dynamics

$$dB(t) = r(t)B(t)dt. \tag{29}$$

where $r(t)$ can be either deterministic or stochastic function of time. In both cases it can be solved for B(t) by dividing 29 by B(t), realizing that $dB(t)/B(t)$ is equal to $d\ln B(t)$[14] and integrating the resulting equation we have:

$$\ln B(t) = \int_0^t r(s)ds.$$

Taking exponential of the expression yields

$$B(t) = e^{\int_0^t r(s)ds}.$$

This is the value of investment assuming stochastic interst rate at time t. In case r is deterministic we obtain expression 9 from the previous section.

3.1.2 Zero-cupon bond

A T-maturity *zero-cupon bond* is a contract that guarantees its holder a payment of one unit of currency at contract's expiry time T with no intermediate payments. The value of zero-cupon bond at time $t < T$ is $P(t,T)$ with $P(T,T) = 1$. It is worth noting that zero-cupon bond price is not directly observable on the market, it can be either stripped from traded government bonds or calculated from products traded in the interbank sector such as swaps or futures. Zero-cupon bonds are however basic quantities in interest rate theory, they are used to define and derive all interest rates as shown below. To do this one only has to specify the *compounding type*, the point in the future when accruing begins and when it ends.

3.1.3 Spot interest rates

For the following definitions, t will denote the current time instant. In order to specify the time fraction between the current time instant and the future expirty time T, the notation $\tau(t,T)$ will be used meaning the *year fraction* between t

[14] Ito's Lemma doesn't have to be used on this place as no stochastic term $dW(t)$ appears in dynamics of 29. Even though $r(t)$ might be stochastic, its value at time t is known, the process is localy deterministic and $dB(t)/B(t)$ can thus be substituted by $d\ln B(t)$.

and T using chosen day-count convention.[15]

A *continuously-compounded spot interest rate* is than defined as the constant rate $R(t,T)$ which fulfilles the following equation given a price of a T-maturity zero-cupon bond:

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

and solving for $R(t,T)$ yields

$$R(t,T) = -\frac{lnP(t,T))}{\tau(t,T)}. \tag{30}$$

Similarily, *simply-compounded spot interest rate* is the constant rate $L(t,T)$ which solves

$$P(t,T) = \frac{1}{1+\tau(t,T)L(t,T)},$$

with the solution given by

$$L(t,T) = \frac{1-P(t,T)}{P(t,T)}\frac{1}{\tau(t,T)}. \tag{31}$$

This way of compounding is used by banks to set the daily inter-bank reference rates which are used on the inter-bank deposit market. They are called LIBOR *rates* which explains the letter L in notation.

Finally, the *annually-compounded spot interest rate* is the constant rate $Y(t,T)$ which solves

$$P(t,T) = \frac{1}{(1+Y(t,T))^{\tau(t,T)}},$$

with the solution given by

$$Y(t,T) = \left(\frac{1}{P(t,T)}\right)^{\frac{1}{\tau(t,T)}} - 1. \tag{32}$$

Subsequently we can define a *short rate* $r(t)$ as the annual rate at which we can deposit money at current time for an infinitezimally short period ($\tau(t,T) \rightarrow 0$). This rate cannot be observed on the market, it is however a handy variable to model in order to price interest rate derivatives as will be shown later. The following can be shown:[16]

$$r(t) = \lim_{T\rightarrow t^+} R(t,T) = \lim_{T\rightarrow t^+} L(t,T) = \lim_{T\rightarrow t^+} Y(t,T). \tag{33}$$

[15] Depending on the day-count convention, the fraction of time measured in years between t and T can vary. For details, see for instance Deutsch (2004).

[16] See for instance Brigo and Mercurio (2004).

3.1.4 Interest rate curves

Collecting the data of iterest rates for all maturities, we can plot a cuve which informs us about the *term structure of interest rates.* The most used such curve is indeed the *zero-cupon curve* or the *yield curve*, which is defined as the following real function of maturity T:

$$T \to \begin{cases} L(t,T), & t < T \leq t+1 \text{ (years)} \\ Y(t,T), & T > t+1 \text{ (years)} . \end{cases} \tag{34}$$

It can be noticed, that the very left point of the yield curve is the short rate defined above.

The other very important curve is the*zero-bond curve* or the *discounting function* which is defined as the following function of maturity T:

$$T \to P(t,T) \quad , \quad T > t.$$

It gives us the value of certain payment of *one* at given maturity T and is thus used to discount certain payments in the future. These two curves contain the same information about the structure of interest rates as can be seen in the previous sub-section where the interest rates are defined. Thus knowing one of them the other can be easily derived. It is the uncertainty of evolution of these curves in time which gives motivation to the existence of interest rates derivatives.

3.1.5 Forward interest rates

As opposed to spot rates, which are used to accrue deposits at current time t, forward rates are interest rates that can be locked in today for an investment in the future. They can be defined using the *forward rate agreement* contract or simply FRA, which enables the investor to lock his/her interest payments between T and S at rate K. $\text{FRA}(t,T,S)$ is a traded intrument defined by its strike price K, maturity T and a time instant $S > T$, at which it has the following payoff on a unit nominal:

$$\tau(T,S)[K - L(T,S)].$$

We can substitute for $L(T,S)$ from definition 31 which delivers

$$\tau(T,S)\left[K - \frac{1-P(T,S)}{P(T,S)\tau(T,S)}\right].$$

We can notice that all these values are known at time T and the payoff is thus certain at this moment. Thus we can discount it by $P(T,S)$ to obtain the value of this payoff at time T:

$$P(T,S)\tau(T,S)K - 1 + P(T,S).$$

Now the value of zero cupon bond $P(T,S)$ at current time t is clearly $P(t,S)$ and the current t-value of certain payment -1 at time T is $-P(t,T)$. Thus the contract's value today is

$$\text{FRA}(t,T,S) = P(t,S)\tau(T,S)K - P(t,T) + P(t,S). \tag{35}$$

If we look for the rate K which makes this contract fair, setting it equal to zero and rearranging we obtain the following expression:

$$K = \frac{P(t,T) - P(t,S)}{P(t,S)} \frac{1}{\tau(T,S)} = F(t,T,S), \tag{36}$$

where $F(t,T,S)$ denotes the *simply compounded forward rate* between T and S at current time. Looking back at equation 36 we can see that a portflio consisting of an amount $1/\tau(T,S)$ of zero cupon bond $P(t,T)$ long and the same amount of $P(t,S)$ short is a traded instrument. Thus using $P(t,S)$ as a nummeraire, the discounted portfolio must be a martingale under measure $Q_{P(t,S)}$, or simply Q^S, should the market be free of arbitrage, as argued at the end of the previous section. It is clear from equation 36, that the discounted portfolio is exactly equal to our forward rate $F(t,T,S)$ which means that $F(t,T,S)$ is a martingale under measure Q^S. This result will be used in the next section when talking about the LIBOR market model. The last point to be made here is that the evolution of $F(t,T,S)$ stops at time $t = T$ when the forward rate $F(T,T,S)$ becomes $L(T,S)$ which means that

$$E^S[L(T,S)|t] = E^S[F(T,T,S)|t] = F(t,T,S).$$

In line with the above definitions of spot interest rates, starting with the known price of zero-cupon bonds $P(t,T)$ and $P(t,S)$ we can equivalently define simply compounded forward rate as the constant rate $F(t,T,S)$ which solves equation

$$\frac{P(t,S)}{P(t,T)} = \frac{1}{1 + F(t,T,S)\tau(T,S)}. \tag{37}$$

The solution is expression 36. Similarily we can define a *continuously-compounded forward rate* as the constant rate $G(t,T,S)$ which solves equation

$$\frac{P(t,S)}{P(t,T)} = e^{-G(t,T,S)\tau(T,S)}$$

and the solution is

$$G(t,T,S) = \frac{\ln P(t,T) - \ln P(t,S)}{\tau(T,S)}. \tag{38}$$

Now we can define *instantaneous forward rate* $f(t,T)$ as:

$$f(t,T) = \lim_{S \to T^+} G(t,T,S) = \lim_{S \to T^+} F(t,T,S).$$

Substituting for $G(t,T,S)$ from 38 we can see that the obtained result is exactly the definition of derivation of logarithm of zero-bond curve $\ln P(t,X)$ around $X = T$:

$$f(t,T) = \lim_{S \to T^+} \frac{\ln P(t,T) - \ln P(t,S)}{\tau(T,S)} = \frac{\partial \ln P(t,T)}{\partial T}. \tag{39}$$

An important point to be repeated here is, that if we wish to model evolution of the term structure of interest rates, we have the choose which quantity we want to model. Among the many posibilities we have are different forward rates, instantaneous forward rates, spot rates or a discount function. All of these rates contain the same information about the markets' anticipation about the future rates. The different approaches of pricing derivatives with interest rates as an underlying choose different quantities as their starting point. Each of the choices has certain advantages and disadvantages. These will be briefly presented in the next section. Before doing so, some basic interest rate producs will be defined.

3.2 Basic interest rates derivatives

3.2.1 Forward rate agreement - FRA

The first derivative product was already defined in the previous section - the *forward rate agreement* or FRA. It is defined by its strike price K and by time-instants T and S with $T < S$. FRA contract obliges its holder to exchange a payment of at time t uncertain simply compounded interest rate $L(T, S)$ for a fix rate K at time S. The contract's payoff at time S is

$$FRA(S, T, S, K) = \tau(T, S)\left[K - L(T, S)\right],$$

where the first argument denotes time instant of FRA's evaluation, second argument denotes FRA's expiry, third denotes time when exchange of payments takes place and fourth is FRA's strike price.[17] The value of this payoff is certain at time T and can be thus discounted to obtain:

$$FRA(T, T, S, K) = P(T, S)\tau(T, S)\left[K - L(T, S)\right]$$

If we try to evaluate this uncertain time T FRA value, this time using the fundamental pricing formula from section one and using zero-cupon bond $P(T, S)$ as nummeraire, according to 28 we obtain

$$\frac{FRA(t, T, S, K)}{P(t, S)} = E^S\left[\frac{P(T, S)\tau(T, S)\left[K - L(T, S)\right]}{P(T, S)}\right]$$

and the no-arbitrage value of this contract is thus

$$FRA(t, T, S, K) = P(t, S)\tau(S, T)E^S\left[K - L(T, S)\right]. \tag{40}$$

As stated above, $L(T, S)$ is a martingale under measure Q^S and $E[L(T, S)]$ is thus equal $F(t, T, S)$:

$$FRA(t, T, S) = P(t, S)\tau(S, T)\left[K - F(t, T, S)\right]. \tag{41}$$

Substituting now for $F(t, T, S)$ from definition 31, we obtain the expression 35 which is the FRA value at time t. Further it was said, that the forward rate

[17] Clearly if first argument is equal to the second, it denotes discounted payoff, if it is equal to the third, it denotes actual payoff.

$F(t,T,S,K)$ is such, that when substituted into FRA for K the contract is equal zero. This is the case only if

$$\begin{aligned} 0 &= E^S[K - L(T,S)] \\ K &= E^S[L(T,S)] \\ F(t,T,S) &= E^S[L(T,S)]. \end{aligned}$$

This result is often misleadingly interpreted, that the forward rates are expectations of the spot rates in the future. We can see that the true interpretation is that forward rate $F(t,T,S)$ is Q^S-expectation of the future spot rate $L(T,S)$. These expectations are the same under the assumption of risk-neutrality of the market, which is however not the case in the real world where risk-aversion is present.

3.2.2 Interest rate swaps - IRS

There are different types of *interest rate swap* contracts. Let's consider a set of fixed time instants $t_0, t_1, ..., t_n$ in the future. Swap is a contract, which provides a payoff at times $t_1, t_2, ..., t_n$ and each of these payoffs dependends on a LIBOR rate reset a times $t_0, t_1, ..., t_{n-1}$ respectively. We assume that year fractions $\tau(t_0,t_1), \tau(t_1,t_2), ..., \tau(t_{n-1},t_n)$ are all equal τ. Then a *forward start receiver swap* or RFS obliges its holder to exchange a floating payment (pay float) $L(t_{i-1},t_i)\cdot\tau$ for a fix payment (receive fix) $K\cdot\tau$ at times t_i, where $i = 1, ..., n$. RFS provides the same payoff as a portfolio of n FRA contracts long with the same strike price K and its value at current time t can thus be written as:

$$\begin{aligned} RFS(t,t_0,t_n,K) &= \sum_{i=1}^{n} FRA(t,t_{i-1},t_i,K) \qquad (42) \\ &= \sum_{i=1}^{n}\left[P(t,t_i)\tau K - P(t,t_{i-1}) + P(t,t_i)\right] \\ &= P(t,t_n) - P(t,t_0) + K\tau\sum_{i=1}^{n} P(t,t_i). \qquad (43) \end{aligned}$$

A *forward rate payer swap* or PFS is a similar contract, its holder is however obliged to exchange fix payment (pay fix) $K\cdot\tau$ for a floating payment (receive float) $L(t_{i-1},t_i)\cdot\tau$ at times t_i. This contract is equivalent to a portfolio of n FRA contracts short with the same strike price K and its value at current time t is therefore:

$$PFS(t,t_0,t_n,K) = P(t,t_0) - P(t,t_n) - K\tau\sum_{i=1}^{n} P(t,t_i). \qquad (44)$$

3.2.3 Swap rate

Setting any of RFS or PFS equal to *zero* and solving for K, we obtain the *forward swap rate* $S(t,t_0,t_n)$:

$$S(t,t_0,t_n) = \frac{P(t,t_0) - P(t,t_n)}{\tau\sum_{i=1}^{n} P(t,t_i)}. \qquad (45)$$

Dividing the by $P(t,t_0)$ we get

$$S(t,t_0,t_n) = \frac{1 - \frac{P(t,t_n)}{P(t,t_0)}}{\tau \sum_{i=1}^{n} \frac{P(t,t_i)}{P(t,t_0)}}, \tag{46}$$

which can be rewritten in terms of simply compounded forward rates as follows (se equation 37):

$$S(t,t_0,t_n) = \frac{1 - \prod_{j=1}^{n} \frac{1}{1+\tau F(t,t_0,t_j)}}{\tau \sum_{i=1}^{n} \prod_{j=1}^{i} \frac{1}{1+\tau F(t,t_0,t_j)}}. \tag{47}$$

This expression will be used when talking about modelling swaption payoff in section four.

Similarily as in the forward rate case, portfolio, say A, consisting of $1/\tau$ portion of $P(t,t_0)$ bond long and the same portion of $P(t,t_n)$ bond short is traded. If we use another portfolio, say B, consisting of bonds $P(t,t_i)$ for $i = 1,2,...,n$, as a nummeraire, the discounted portfolio A/B has to be a martingale under the measure Q_B. Again, it is straightforward that this discounted portfolio is exactly equal to the swap rate, which means, that the swap rate is a martingale under the measure Q_B. This result will be used in the next section when talking about the swap market model.

Another very useful expression for the swap rate can be derived when we substitute 41 into equation 42, set the value of the swap equal zero as done in the second step in the derivation below and subsequently solve for $K = S(t,t_0,t_n)$:

$$RFS(t,t_0,t_n,K) = \sum_{i=1}^{n} P(t,t_i)\tau [K - F(t,t_{i-1},t_i)] \tag{48}$$

$$S(t,t_0,t_n) = \frac{\sum_{i=1}^{n} P(t,t_i)F(t,t_{i-1},t_i)}{\sum_{k=1}^{n} P(t,t_k)}$$

$$S(t,t_0,t_n) = \sum_{i=1}^{n} \frac{P(t,t_i)}{\sum_{k=1}^{n} P(t,t_k)} F(t,t_{i-1},t_i)$$

$$S(t,t_0,t_n) = \sum_{i=1}^{n} w_i F(t,t_{i-1},t_i). \tag{49}$$

In the last step we have substituted the fraction with w_i. Clearly the sum of all ws is equal to one and they can be therefore interpreted as weights. In light of this restatement the forward swap rate spanning between t_0 and t_n can be interprated as a weighted average of forward rates F over this period of time.

Finally, we can also express the interest rate swap in terms of a swap rate which will come handy when talking about swaptions. Starting with equation 48 we can subtract a swap (which spannes over the same period of time) who's fix leg K is equal to the swap rate and which is thus by definition zero:

$$PFS(t,t_0,t_n,K) = \sum_{i=1}^{n} P(t,t_i)\tau [F(t,t_{i-1},t_i) - K] -$$

$$
- \sum_{i=1}^{n} P(t,t_i)\tau\left[F(t,t_{i-1},t_i) - S(t,t_0,t_n)\right]
$$

$$
= (S(t,t_0,t_n) - K)\tau\sum_{i=1}^{n} P(t,t_i). \tag{50}
$$

From the expression 50 it is nicely seen, that the the value of a payer swap is equal zero if $K = S(t,t_0,t_n)$ it is positive if $K < S(t,t_0,t_n)$ and it is negative if $K > S(t,t_0,t_n)$.

Having $t < t_0$ the first forward rate resetting point is a time instant in the future. If we set $t_0 = t$, the first resetting date will be the current time instant t and thus the first payment at time t_1 will be known and there is uncertainty only about the value of the next payments which depend on the future value of spot rates. Furthermore, if we enter into a reciever swap with $t_0 = t$ and $K = S(t,t,t_n)$, it can be shown that this first payment will be positive if the yield curve is upward sloping and negative if the yield curve is downward sloping (inverted) and vice-versa for a payer swap. The reason for this is, that the received fixed leg is a weighted average of the one-period forward rates $F(t,t,t_1) = L(t,t_1)$,$F(t,t_1,t_2)$,$F(t,t_2,t_3)\ldots$ and these are increasing if the yield curve is upward sloping.

3.2.4 Caps and Floors

A call option with an interest rate as an underlying is called a *caplet*. It is defined by the underlying forward rate $F(t,T,S)$ and the strike price K. Caplet expires at time T and as opposed to an ordinary stock option, it provides a payoff at time S:

$$
Cpl(S,T,S,K) = \tau(T,S)\max\left[F(T,T,S) - K, 0\right],
$$

where the first argument denotes time instant of caplet's evaluation, second argument denotes option's expiry, third denotes time when payment takes place and fourth caplet's strike price.[18] $F(T,T,S)$ is equal to $L(T,S)$. Clearly, the value of caplet's payoff is known at time T:

$$
Cpl(T,T,S,K) = P(T,S)\tau(T,S)\max\left[F(T,T,S) - K, 0\right]
$$

and the current market price of a caplet can be denoted as $Cpl(t,T,S,K)$. Furthermore, a caplet is called at-the-money (ATM) if $F(t,T,S) = K$ or $Cpl = Cpl(t,T,S,F(t,T,S))$, in-the-money (ITM) if $F(t,T,S) > K$ and out-of-the-money (OTM) if $F(t,T,S) < K$.

A put option on an interest rate is called a *floorlet* an it provides the following payoff at time S:

$$
Frl(S,T,S,K) = \tau(T,S)\max\left[K - F(T,T,S), 0\right].
$$

A floorlet is called at-the-money (ATM) if $F(t,T,S) = K$, in-the-money (ITM) if $F(t,T,S) < K$ and out-of-the-money (OTM) if $F(t,T,S) > K$.

[18] Clearly if first argument is equal to the second, it denotes discounted payoff, if it is equal to the third, it denotes actual payoff.

Let's consider a set of fixed time instants $t_0, t_1, ..., t_n$ in the future. A *cap* is defined as a portfolio of n caplets with the same strike price K, each having a forward LIBOR rate spanning between the fixed time instants as an underlying. Thus the cap provides a payoff at times $t_1, t_2, ..., t_n$ which is known already at times $t_0, t_1, ..., t_{n-1}$. The time period which spans between t_0 and t_n is called a *tenor* of a cap. Cap's current time t-market value is:

$$Cap(t, t_0, t_n, K) = \sum_{i=1}^{n} Cpl(t, t_{i-1}, t_i, K).$$

Similarily a *floor* is defined as a series of n floorlets with the same strike price K and the following market value:

$$Floor(t, t_0, t_n, K) = \sum_{i=1}^{n} Frl(t, t_{i-1}, t_i, K).$$

In reality, however, one observes prices of caps and floors and the prices of caplets and floorlets must be stripped. It can be noticed that the payoff of a cap is the same as a payoff of a PFS where the payment takes place only if it is positive.

Cap as well as floor are called ATM when $S(t, t_0, t_n) = K$, ITM when $S_{\alpha,n}(t_0) > K$ for cap and $S(t, t_0, t_n) < K$ for floor and opposite inequalities hold for OTM definition. Furthermore, the first resetting time can also be the current time instant, or $t_0 = t$.

The derivation of the closed formulas for FRA and swap values above was a result of an assumption of absence of arbitrage on the market. No assumption was taken about the dynamics of the yield curve. This is not the case when we want to find a value of a caplet. An option's value generaly depends on our view of dynamics of its underlying and thus on our view of probability distribution of the underlying at option's expiry time. There can be as many different views about the concrete shape of this distribution as the number of market-participants and thus no single fair value of option is available. Thus no closed formula of a caplet/cap or floorlet/floor can be stated unless an assumptions on a yield curve dynamics are made.

3.2.5 Swaptions

A call option who's underlying is a swap rate is called a *swaption*. It is basically a right but not an obligation to enter into a swap contract at time t_0 with a fixed leg of K. This right will be executed if the value of a swap is positive. Let's start with the expression 50 for a payer swap. The payoff of a swaption at at its expiry time t_0 is the following in case of a *payer swaption* (payer swap is an underlying) or a *call swaption*:

$$CSwaption(t_0, t_0, t_n, K) = \max\left[(S(t, t_0, t_n) - K)\tau \sum_{i=1}^{n} P(t_0, t_i), 0\right], \quad (51)$$

whereas the payoff of a *receiver swaption* (receiver swap is an underlying) or a *put swaption* is:

$$PSwaption(t_0, t_0, t_n, K) = \max\left[(K - S(t, t_0, t_n))\tau \sum_{i=1}^{n} P(t_0, t_i)\, ,\, 0\right]. \quad (52)$$

The market prices are denoted $CSwaption(t, t_0, t_n, K)$ and $PSwaption(t, t_0, t_n, K)$. Similarily to a cap, receiver as well as payer swaptions are ATM when $S(t, t_0, t_n) = K$, swaption is ITM when $S(t, t_0, t_n) > K$ for payer swaption and $S(t, t_0, t_n) < K$ for a receiver swaption. Opposite inequalities hold for OTM definitions.

4 Interest Rate Models

In the previous section it was shown, that the same term structure of interest rates can be described using different quantities. In order to price an interest rate derivative, we have to build a model which describes dynamics of one of these quantities. Some of the important aspects one has to consider when choosing one particular model are presented below.

Choice of quantity to model

The first choice is about the variable which is to be modeled. One can either try to model dynamics of a traded instrument like a bond, or an evolution of certain interest rate. More complex way is to model the whole yield curve.

Discrete versus continuous dynamics

Designing the dynamics of chosen variable, one can describe the model using continuous notification in the form of equation 2, or a discrete setup specifying the length of each time period (one hour, one day, one week etc.). However, if one specifics dynamics of some variable in a continuous way and if a simulation of a future developement is to be done according to this specification, one has to resort to discretizing the model anyway. In this brief introduction to only continuous model specification will be used.

Fitting the observed term structure

There are models where the observed term structure is set exogenously as a starting point. Other models try to capture dynamics of one state variable and the term structure at the starting point comes endogeneously from the concrete drift and volatility specification.

Number of factors

By number of factors the sources of risk or soruces of volatility are meant. Volatility of yield curve can be decomposed into three most significant factors. First factor contributes to parallel shifts of the whole curve, second one contributes to change of slope and the third one contributes to change of curvature.[19]. These factors can be expressed by adding extra non-correlated Wiener processes in the diffusion equation (2). This choice is very important. Adding the third factor, which might account for only 1-2 percent of total volatility[20], might seem senceless when pricing a caplet. There are however positions who's value is insensitive to a parallel shift in the yield curve, who's value is however sensitive to change of curvature (such as sale of medium-term zero cupon bond and a duration weighted investment into cash and a long term zero-cupon bond).

There are three groups of models that are mostly used in practice: the short rate models, Heath Jarrow Morton (HJM) models and the market models. Each

[19] See Gibson, Lhabitant and Talay (2001).
[20] See Gibson, Lhabitant and Talay (2001).

group of models has strengths and weaknesses. In order to decide which model to use we have to state the purpose of the model and look at the features which the models have. For instance our goal might be to price a simple derivative, who's payoff is contingent upon the state of only one point of the yield curve in the future (like a single caplet). Or we are interested in pricing a complex product, which provides multiple payoffs in the future, each depending on a different point of the yield curve and each entailing some other condition like for instance an activation (knock-in) or deactivation (knock-out) mechanism (like a knock-in swaption). It is clear that in the first case we will choose a model which realisticaly captures the evolution of the single relevant rate. Such model can however imply an unrealistic evolution of the whole yield curve and might therefore be completely unsuitable for pricing the other complex product. The following text will briefly introduce the above mentioned models.

4.1 Short Rate Models

One of the oldest approaches to express interest rate dynamics is via dynamics of the short rate. In the most general form it can be written as:

$$dr(t) = \mu(t, r(t))dt + \sigma(t, r(t))dW(t), \tag{53}$$

where the drift function $\mu(t, r(t))$ as well as the volatility function $\sigma(t, r(t))$ can be any function of time and state variable and $dW(t)$ is a Wiener process under the real probability measure. The concrete specification of these functions determines the shape of the whole yield curve as well as the probability distribution of interest rates which will be shown below.

4.1.1 Dynamics under the Real Measure and Market Price of Risk

The problem with this framework is that it assumes only one traded instrument which is the money market account from which stochasticity of the short rate stems.[21] The basic assumption about no-arbitrage pricing is that we can take any two traded instruments and create such self-financing trading strategy (no additional payments or withdrawal of money during the strategy life) that an arbitrary payoff of a derivative can be achieved. The initial investment that we need for setting this strategy forth is then equal to the value of the derivative. Assuming only one risk factor, such model that provides dynamics of at least two traded instruments is called *complete*.[22] Our model provides dynamics of only one instrument and therefore additional information is needed to make the model complete. This information is the market price of risk λ. Taking relationship 25 into account, equation 53 can than be rewritten as follows:

$$dr(t) = [\mu(t, r(t)) - \lambda\sigma(t, r(t))]dt + \sigma(t, r(t))d\widehat{W}(t).$$

Knowing dynamics of short rate under the risk neutral measure, contingent claims can be prized using the fundemantal pricing equation. The most basic derivative in the short rate model framework is the zero-cupon bond, which

[21] The deposit is riskless, though only for an infinitesimal time period. The offer and demand of deposits sets the equilibrium interest rate for the next infinitesimal period.

[22] See Branger and Schlag (2004).

provides a payoff of *one* at its maturity T and can thus be priced as follows using MMA as a nummeraire:

$$P(t,T) = \widehat{E}\left[\frac{1}{e^{\int_t^T r(s)ds}}\middle| t\right]. \tag{54}$$

Given the dynamics of $r(t)$ the integral in the formula above can be computed for well behaved drift and volatility functions of the short rate and calculation of expected value is than straighforward employing basic stochastic calculus. This obtained price is the short-rate-model-implied no-arbirage price of zero cupon bond which is a (theoreticaly) tradable asset. Our model is now complete and self-financing strategies can be created.

Furthermore, if prices of all zero-cupon bonds are computed we obtain the discount function from which all spot rates as well as forward rates can be derived. This model-implied term structure of interest rates is not necessarily the one which is observed on the market. To obtain the real term structure, the resulting expression for zero-cupon bond has to be set equal to the observed value and inverting the equation one can solve for the parameters of the drift and volatility functions and for λ. It is obvious that we can only fit as many observed points of the yield curve as the number of parameters we have in our short rate dynamics (equation 53). The solution to this problem will be shown later on.

4.1.2 Dynamics under the Risk Neutral Measure

An alternative and for derivative-pricing purposes more convenient way is to model dynamics of short rate directly under the risk-neutral probability measure:

$$dr(t) = \theta(t, r(t))dt + \sigma(t, r(t))d\widehat{W}(t),$$

where the volatility function stays the same [23] while the drift function changes. Within such specification no additional information is needed to derive prices of zero-cupon bonds. These can be then inverted in order to calibrate to the observed term strcture of interest rates.

4.1.3 Examples of short rate dynamics

The big advantage of short rate models is that they are very flexible as to the exact formulation of drift and volatility functions. There is a big choice of functions one can employ and thus there are big chances one can capture the dynamics of short rate in a realistic way. One can include features like jumps, stochastic volatility, shift factors which allow for exact calibration etc.[24] The most known models will now be introduced.

Vasicek Model

$$dr(t) = k[\theta - r(t)]dt + \sigma d\widehat{W}(t) \quad , \quad r(t) = r_0, \tag{55}$$

[23] Girsanov's Theorem, see the first section.

[24] See for instance Musiela and Rutkwski (2005).

r_0 is the current short rate level. This form of dynamics first used by Vasicek[25] captures the desired feature of interest rate to revert to certain equilibrium level θ by the pace k, which is a positive constant. The mechanism is easily seen realizing, that if the short rate $r(t)$ at time t exceeds the level θ, the term in the brackets of equation 55 is negative over the next time instant of short rate's dynamics. The higher the k term, the faster the short rate reverts to its equilibrium level.

Integrating equation 55 over a time period say $< t, T >$ we obtain the following value for the short rate at time T:

$$r(T) = r(t)e^{-k(T-t)} + \theta(1 - e^{-k(T-t)}) + \sigma \int_t^T e^{-k(T-u)} d\widehat{W}(u).$$

The last term is a stochastic integral which is normally distributed with zero expected value and variance of $\sigma^2 \int_t^T e^{-2k(T-u)}dt$ [26] and thus the short rate in Vasicek model is also normally distributed. This is obviously an undesired feature as the short rate can get into a negative teritory. This creates an arbitrage in case we have the opportunity to hold money at home with no iterest, which is the case of the real world.

Dothan Model

$$dr(t) = \sigma r(t)dW(t), r(t) = r_0, \tag{56}$$

where $dW(t)$ is Wiener process under real probability measure. If again we substitute from relationship 25 for $dW(t)$, the following dynamics under risk neutral probability measure is obtained:

$$dr(t) = \theta r(t)dt + \sigma r(t)d\widehat{W}(t) \quad , \quad r(t) = r_0,$$

where $\theta = -\lambda\sigma$. Integrating equation 57 over period $< t, T >$ we obtain the following value for the short rate at time T:

$$r(T) = r(t) \exp\left\{ \left(\theta - \frac{1}{2}\sigma^2\right)(T-t) + \sigma \int_t^T d\widehat{W}(s) \right\}.$$

Because the stochastic integral (which itegrates simply to $W(T)$ starting with $W(t) = 0$) in the equation above is in the exponent, the short rate $r(T)$ is log-normally distributed. This is more realistic assumption because the short rate cannot be below zero at any point of time. Because this model has $k = 0$, the process reverts to *zero* which is neither realistic. A combination of Vacicek and Dothan model can be taken with the following dynamics:

$$dr(t) = k[\theta - r(t)]dt + \sigma r(t)d\widehat{W}(t) \quad , \quad r(t) = r_0, \tag{57}$$

which is also log-normally distributed or a Cox, Ingersoll and Ross process (CIR):

$$dr(t) = k[\theta - r(t)]dt + \sigma\sqrt{r(t)}d\widehat{W}(t) \quad , \quad r(t) = r_0, \tag{58}$$

[25] See Vasicek (1977).

[26] see Bjork (2004) on moments of stochastic integrals.

which follows noncentral χ^2 distribution.

Hull and White Model

As already noted earlier, the common problem of the above specifications of short rate processes is that if we want to fit the model to the observed term structure, only as many points of the yield curve can be recovered as the number of parameters in the model, whereas in reality the yield curve is described by infinity of maturities. Hull and White (1990) came with a solution to this problem, they introduced time dependence of one of the parameters. Thus the whole yield curve can be fitted when inverting the model-implied dicsount curve to solve for the time dependent parameter. The remaining parameters can than be adjusted as to recover prices of some liquid derivatives like ATM caps or swaptions. As a starting point they used the Vasicek Model, this feature can however be used by any other specification and choosing any parameter. The Hull and White short rate dynamics has the following form:

$$dr(t) = [\vartheta(t) - ar(t)]dt + \sigma d\widehat{W}(t) \quad , \quad r(t) = r_0. \qquad (59)$$

4.1.4 Pricing Derivatives and some Disadvantages of Short Rate Models

Short rate models are suitable for pricing complex exotic instruments using *Monte-Carlo simulation*. One carries out a very large number of simulations (tens of tousands, millions) of a possible short rate developement using calibrated short rate model. The values of a short rate at time instants, which influence the instrument's payoff, are recorded, the payoff is evaluated and the average of these payoffs resulting from all simulations is discounted to the presence using the stochastic money market account $\exp\left\{\int_t^T r(s)ds\right\}$ which is a funciton of the by us modelled short rate. This is called the risk neutral valuation.[27] To make clear how the payoff is evaluated, let's assume that the derivative's payoff is a function of future spot rate $L(T,S)$. If we carry out a single simulation of a short rate, we will get its time T value, which implies a value of a zero-cupon bond $P(T,S)$ (see equation 54) which in turn implies a value for $L(T,S)$ (see equation 31).

In cases of simpler drift and volatility specifications (like the Hull and White model)[28] the fundamental pricing formula implies a closed form analytical formulas for value of simple interest rate derivatives such as an option on a zero cupon bond for instance. There is an easy relationship between a put option on a zero-cupon bond and a caplet.[29] Using formula for a bond option an option on interest rate option can thus be evaluated.

The main disadvantage of short rate models is that if we choose more complicated drift and volatility functions (which might imply a very realistic evolution

[27] Note that the phrase 'risk neutral valuation' doesn't imply anything about the degree of risk aversion of the market, it simply means that the fundamental pricing formula is used with MMA as a nummeraire.

[28] See Brigo and Mercurio (2001).

[29] See for instance Branger and Schlag (2004).

of short rate) we will be most likely unable to find analytical formulas for simple derivatives. The result is that it is often difficult to calibrate such models to the current yield curve as there is no formula for zero-cupon bonds available which can be inverted. If we find a realistic dynamics for a short rate, even thougth we might be able to achive calibration to the observed yield curve, the model might imply an undesired evolution of the whole yield curve. Such model will therefore be unsuitable for pricing say swaptions, whos payoff is dependent on a swap rate which is a function of the yield curve. Furthermore the resulting calibrated parameters often don't have a straightforward meaning which makes the model somewhat obscure and difficult to interpret.

4.2 Heath, Jarrow and Morton Framework

In 1992 a paper from Heath, Jarrow and Morton was published where a general framework for interest rate dynamics was introduced. The dynamics of the whole yield curve is modeled with instantaneous forward rates as the fundamental quantities. Here I will breefly show the main idea.

4.2.1 Dynamics under the Real Measure

In HJM model the following instantaneous forward rate dynamics with general drift and volatility functions is considered:

$$df(t,T) = \mu_f(t,T)dt + \sigma_f(t,T)dW(t), \tag{60}$$

where $dW(t)$ is a Wiener process dynamics under the real probability measure and $f(t,t)$ is the short rate. In the original paper $dW(t)$ is a vector of n Wiener processes, for simplicity only scalar is considered here. The observed instantaneous forward rate curve $T \to f(t,T)$ is taken as the starting point and the model is thus automaticaly adjusted to the current market term structure. From equation 39 we obtain the following relationship for zero cupon bond:

$$P(t,T) = e^{\int_t^T f(t,s)ds} \tag{61}$$

and using Ito's Lemma we obtain the following dynamics for zero-cupon bond:[30]

$$dP(t,T) = \mu_P(t,T)P(t,T)dt - \sigma_P(t,T)P(t,T)dW(t), \tag{62}$$

where

$$\mu_P(t,T) = f(t,t) - \int_t^T \mu_f(t,s)ds + \frac{1}{2}\left(\int_t^T \sigma_f(t,s)ds\right)^2 \tag{63}$$

$$\sigma_P(t,T) = -\int_t^T \sigma_f(t,s)ds. \tag{64}$$

[30] See Heath, Jarrow and Morton (1992) for proof or Branger and Schlag (2004).

4.2.2 Dynamics under the Risk-Neutral Measure

For an arbitrary T we now have dynamics of traded zero-cupon bond with maturity T and in order to prevent arbitrage, the no-arbitrage relationship 14 from the first section must hold true for all T's:

$$\frac{\mu_P(t,T) - f(t,t)}{\sigma_P(t,T)} = \lambda(t), \tag{65}$$

where the dependence of λ on time should ilustrate that the risk aversion of markets can change in time. Multiplying 65 by σ_P, substituting for μ_P and σ_P from expressions 63 and 64, and deriving over T the no-arbitrage condition becomes:

$$\mu_f(t,T) + \sigma_P(t,T) \cdot \sigma_f(t,T) = \lambda(t)\sigma_f(t,T),$$

and rearranging we obtain the following restriction for drift of instantaneous forward rates:

$$\mu_f(t,T) = \sigma_P(t,T) \cdot \sigma_f(t,T) - \lambda(t)\sigma_f(t,T). \tag{66}$$

If we substitute this no-arbitrage drift restriction into our original inst. forward dynamics 60 we have

$$df(t,T) = [\sigma_P(t,T) \cdot \sigma_f(t,T) - \lambda(t)\sigma_f(t,T)]dt + \sigma_f(t,T)dW(t),$$

and switching to a Wiener process und risk neutral probability measure using expression 25 we obtain the following no-arbitrage risk-neutral dynamics of instantaneous forward rates:

$$df(t,T) = \sigma_P(t,T) \cdot \sigma_f(t,T)dt + \sigma_f(t,T)d\widehat{W}(t). \tag{67}$$

The dynamics of zero-cupon bond under the risk-neutral measure will be clearly

$$dP(t,T) = f(t,t)P(t,T)dt - \sigma_P(t,T)P(t,T)d\widehat{W}(t),$$

as the diffusion term doesn't change switching between equivalent probability measures and since the local rate of return of any traded instrument must be equal to the riskless rate of return, or $f(t,t)$ in our case, to prevent arbitrage (see section 2.5). The same result can be obtained applying Ito's Lemma to eq. 61 using directly dynamics 67.

4.2.3 Model specification

The equation 67 of risk-neutral dynamics of instantaneous forward rates tells us that only the volatility function of instantaneous forward rate σ_f is needed in order to fully specify the model (note that bond volatility function σ_P is a function of forward rate volatility σ_f). In principle any functional form for forward rate volatility can be chosen.

The most simple way is to assume constant volatility for all inst. forward rates - $\sigma_f(t,T) = \bar{\sigma}$. It can be shown that this volatility specification implies short rate dynamics of a Hull and White model - equation 59 - with constant a equal *zero*.[31] The probability distribution of such specified short rate is normal and thus probability of negative short rate is positive. Furthermore, the expected level of yield curve is strictly rising with time as well as yield curve's slope, which are not very desired properties.

A little more realistic is to assume a parametric form where the volatility of inst. forward rate is decreasing with forward maturity T - $\sigma_f(t,T) = \bar{\sigma} e^{-a(T-t)}$. This specification also implies Hull and White dynamics of a short rate, however the mean-reversion feature of short rate is preserved which implies that the expected level of the yield curve is not rising to infinity. The short rates are however also normaly distributed and negative rates are thus possible.

4.2.4 Pricing Derivatives using HJM

Pricing derivatives within HJM framework similar to pricing with short rate models. One can either resort to pricing via Monte-Carlo simulations, where those forward rates are simulated (after being first discretized) which are necessary for evaluating the derivative's payoff. Using the fundamental pricing formula one can derive closed form formulas for simple derivatives like futures, options on bonds and interest rates. HJM models are very complex and which prevented them from coming into more common use.

4.3 Market Models

HJM framework as well as short rate models describes evolution of instantaneous interest rates that are not directly observed on the market. A derivative, who's payoff is dependent on an interest rate directly observed on the market (say a LIBOR or a swap rate) has to be evaluated by deriving dynamics of the observed rate via Ito's Lemma. The obtained dynamics will imply a probability distribution which is then used when evaluating the expected payoff of the derivative via fundamental pricing formula. The problem is that the obtained probability distribution will not allways be "nice" and not allways shall we obtain closed analytical formula for our derivative for all model (volatility) specifications. Pricing a derivative using Monte Carlo simulation might neither be optimal within the HJM framework. In this case the derivative's payoff will be derived for each single simulation from the instantaneous rate and this will make the simulation process very slow.

An alternative is to model directly the market observed interest rates - thus the name "Market Models". This concept was first introduced by Brace, Gatarek and Musiela (1997), Miltersen, Sandmann and Sondermann (1997) and Jamshidian (1997).

4.3.1 LIBOR Market Model

The starting point is the result from the previous section which shows, that the simply compounded forward rate $F(t,T,S)$ is a martingale under the Q^S measure. At time T rate $F(T,T,S)$ becomes $L(T,S)$ which is the interbank

[31] See Branger and Schlag (2004) for derivation.

reference spot interest rate set daily by a group of banks with largest volumes of trades on the interbank money market. As a result of the martingale property, it is assumed, that the forward rate follows a driftless GBM process under the Q^S measure:

$$dF(t,T,S) = \sigma_F(t,S,T)F(t,T,S)dW(t)^S. \qquad (68)$$

The volatility function in this case is linear and the terminal probability distribution Q^S of LIBOR rate $L(T,S)$ is thus *lognormal*[32] with Q^S-expected value of $F(t,T,S)$. Under this specificatin the forward rate can never become negative.

Pricing a Caplet in LIBOR Market Model Framework

Let's now consider a caplet $Cpl(t,T,S,K)$ with a payoff $Cpl(S,T,S,K) = \tau(T,S)\max[L(T,S)-K,0]$ at time S. The value of this payoff is known at time T - $Cpl(T,T,S,K) = P(T,S)Cpl(S,T,S,K)$. According to the fundamental pricing formula 28 the following must hold in absence of arbitrage:

$$\begin{aligned}
\frac{Cpl(t,T,S,K)}{P(t,S)} &= E^S\left[\frac{Cpl(T,T,S,K)}{P(T,S)}\right] \\
Cpl(t,T,S,K) &= P(t,S)E^S\left[Cpl(S,T,S,K)\right] \\
Cpl(t,T,S,K) &= P(t,S)\tau(T,S)E^S\big[\max[L(T,S)-K,0]\big]. \qquad (69)
\end{aligned}$$

To evaluate an expected value of a maximum of difference of two log-normaly distributed variables and a *zero*, the following formula can be used:[33]

$$E\big[\max[e^X - e^Y],\, 0\big] = E\big[e^X\big]N(d) - E\big[e^Y\big]N(d-s), \qquad (70)$$

where $N(.)$ is a standard normal distribution function and

$$\begin{aligned}
d &= \frac{1}{s}\left(\ln\frac{E\big[e^X\big]}{E\big[e^Y\big]} + \frac{1}{2}s^2\right) \\
s^2 &= var[X-Y].
\end{aligned}$$

Because $L(T,S)$ is distributed log-normaly and K can be thought of as a variable distributed log-normaly with *zero* variance, this formula can be used for the expectation under Q^S measure $E^S[...]$ in equation 69. The value of our caplet thus becomes

$$Cpl(t,T,S,K) = P(t,S)\tau(T,S)\Big(E^S\big[L(T,S)\big]N(d) - E^S\big[K\big]N(d-s)\Big).$$

Now we can realize, that $L(T,S)$ is a martingale under Q^S and therefore the expectation under Q^S is equal to the today's forward rate $F(t,T,S)$:

$$Cpl(t,T,S,K) = P(t,S)\tau(T,S)\big(F(t,T,S)\cdot N(d) - K\cdot N(d-s)\big), \qquad (71)$$

with

$$\begin{aligned}
d &= \frac{1}{s}\left(\ln\frac{F(t,T,S)}{K} + \frac{1}{2}s^2\right) \\
s^2 &= \int_t^T \sigma_F(s,T,S)^2 ds.
\end{aligned}$$

[32]See section 2.1.

[33]See Branger and Schlag (2004).

Finally, the expression for s^2 is obtained as follows:

$$s^2 = var[\ln L(T,S) - K] = var[\ln L(T,S)].$$

To derive variance of $\ln L(T,S)$, we first derive dynamics of $\ln F(t,T,S)$ using Ito's Lemma:

$$\begin{aligned} d\ln F(t,T,S) &= \frac{1}{F(t,T,S)}\sigma_F(t,T,S)F(t,T,S)dW(t)^S - \\ &- \frac{1}{2}\frac{1}{F(t,T,S)^2}\sigma_F(t,T,S)^2F(t,T,S)^2dt \\ &= -\frac{1}{2}\sigma_F(t,T,S)^2dt + \sigma_F(t,T,S)dW(t)^S, \end{aligned}$$

and integrating over the forward's life (t,T) we obtain

$$\ln F(T,T,S) = F(t,T,S) - \frac{1}{2}\int_t^T \sigma_F(s,T,S)^2ds + \int_t^T \sigma_F(s,T,S)dW(s)^S.$$

Only the last term is a stochastic and it can be shown[34] that this integral is normaly distributed with *zero* mean and variance of $\int_t^T \sigma_F(s,T,S)^2ds$, which is the desired expression.

Black's Formula for a Caplet and Cap

We can see that all variables in our caplet pricing equation 71 are given or observed on the market except of the volatility function (meant $\sigma_F(t,S,T)$) of forward rate over time. An assumption has to be made about the future volatility in order to price a caplet. As stated in the first section, we can either base our assumption on the past developement or on our own view. If we assume, that the volatility function is constant over time - $\sigma_F(t,T,S) = \bar{\sigma}_F(T,S)$, than s^2 in the caplet pricing formula 71 becomes $\sigma_F(T,S)^2(T-t)$. This result is known as the *Black's formula*[35] and the caplet pricing formula can be written as:

$$Cpl^{Black}(t,T,S,K,\sigma) = P(t,S)\tau(T,S)Bl(K,F(t,T,S),\sigma), \tag{72}$$

where

$$\begin{aligned} Bl(K,F(t,T,S),\sigma) = &\quad F(t,T,S)N(d) - K\cdot N(d-s) \\ d = &\quad \tfrac{1}{s}\left(\ln\tfrac{F(t,T,S)}{K} + \tfrac{1}{2}s^2\right) \\ s = &\quad \sigma\sqrt{T-t}. \end{aligned}$$

If we are given a price of a caplet $Cpl(t,T,S,K)$, we can solve for parameter σ, which implies this price when substituted as a constant into the forward rate dynamics 68.

Caplets and floorlets are in reality not traded on the market, instead, the caps and floors are traded instruments. It is market practice to quote caps and floors

[34]See Björk (2004).
[35]See Black 1976.

directly in *implied volatility* - the constant σ that recovers the market price of a cap $Cap(t, t_0, t_n, K)$ when used as a constant parameter in the forward rate dynamics 68:

$$Cap^{Black}(t, t_0, t_n, K, \sigma) = \sum_{i=1}^{n} Cpl^{Black}(t, t_{i-1}, t_i, \bar{\sigma}) \tag{73}$$

This practice is very handy as the forward rates move frequently with the movement of the swap market and thus the caplet prices fluctuate as well. Oposite to this the volatilities might stay the same maybe the whole day or more.

4.3.2 Swap Market Model

Similarily as LIBOR market model, the Swap market model aims to model a rate directly observed on the market, namely the swap rate. From the previous sectin we know, that the swap rate is a martingale under the measure Q_B, where B is a portfolio consisting of n bonds, namely $P(t, t_i)$ for $i = 1, 2, ..., n$. It is assumed, that the swap rate follows a driftless GBM process under the Q_B measure:

$$dS(t, t_0, t_n) = \sigma_S(t, t_0, t_n) S(t, t_0, t_n) dW(t)^B. \tag{74}$$

Again, the volatility function is linear which implies that the terminal Q_B distribution of the swap rate $S(t_0, t_0, t_n)$ is log-normal with Q_B-expected value equal $S(t, t_0, t_n)$.

Pricing a Swaption in the Swap Market Model Framework

Deriving the value of a swaption within the swap market model framework is very similar as deriving the caplet formula above. The task is to evaluate the uncertain payoff of a payer swaption at t_0, formula 51. Using the fundamental pricing formula, its today value is:

$$\frac{CSwaption(t, t_0, t_n, K)}{\sum_{i=1}^{n} P(t, t_i)} = E_B\left[\frac{CSwaption(t_0, t_0, t_n, K)}{\sum_{i=1}^{n} P(t_0, t_i)}\right]$$

$$CSwaption(t, t_0, t_n, K) = \tau \sum_{i=1}^{n} P(t, t_i) E_B\left[\max[S(t_0, t_0, t_n) - K,\ 0\,]\right].$$

The terminal distribution of the swap rate is log-normal and thus the formula 70 can be used for the expected value in the equation above, which becomes:

$$CSwaption(t, t_0, t_n, K) = \tau \sum_{i=1}^{n} P(t, t_i) \Big(E_B[S(t_0, t_0, t_n)] N(d) - K \cdot N(d - s) \Big)$$

We know that the swap rate is a martingale under the measure Q_B and thus the Q_B-expectation about the future swap rate $S(t_0, t_0, t_n)$ is equal to the today's swap rate $S(t, t_0, t_n)$ and the swaption's price becomes:

$$CSwaption(t, t_0, t_n, K) = \tau \sum_{i=1}^{n} P(t, t_i) \Big(S(t, t_0, t_n) \cdot N(d) - K \cdot N(d - s) \Big), \tag{75}$$

with

$$
\begin{aligned}
d &= \frac{1}{s}\left(\ln\frac{S(t,t_0,t_n)}{K} + \frac{1}{2}s^2\right),\\
s^2 &= \int_t^{t_0}\sigma_S(s,t_0,t_n)^2 ds.
\end{aligned}
$$

All variables in the above swaption pricing formula 75 are given or observed on the market, except of the volatility function of the swap rate, which has to be again estimated.

Black's Formula for Swaption

As in the case of LIBOR market model above, we can assume constant volatility for the swap rate volatility function $\sigma_S(t,t_0,t_n) = \bar{\sigma}_S(t_0,t_n)$. Doing this we get the Black's formula for a payer swaption:

$$
CSwaption^{Black} = \tau\sum_{i=1}^{n} P(t,t_i)Bl(K,S(t,t_0,t_n),\bar{\sigma}_S(t_0,t_n)), \tag{76}
$$

where

$$
\begin{aligned}
Bl(K,S(t,t_0,t_n),\bar{\sigma}_S(t_0,t_n)) =& \quad S(t,t_0,t_n)N(d) - K\cdot N(d-s)\\
d =& \quad \tfrac{1}{s}\left(\ln\tfrac{S(t,t_0,t_n)}{K} + \tfrac{1}{2}s^2\right)\\
s =& \quad \bar{\sigma}_S(t_0,t_n)\sqrt{T-t}.
\end{aligned}
$$

Market practice is to quote swaptions directly in implied volatilities. It is important to mention, that the market participants don't have to beleive that the Black formula is correct (and thus the assumption of lognormality of distribution of the underlying). The Black formula is just a way of expressing the price. The fact that different prices of swaptions are available reflects that there is no single dynamics deemed as universaly true. Features such as jumps, stochastic volatility etc. can be added to forward rate dynamics, Black's formula is however still used to express the option's price.

4.3.3 Compatibility of LIBOR and Swap Market Models

The main drawback of the two presented frameworks is that they are not mutualy compatible. We can suppose for instance that indeed the dynamics of forward rate follow equation 68. As shown in the previous section, the swap rate is a function of forward rates (equation 49) and thus the swap rate's dynamics can be calculated with help of Ito's Lemma knowing dynamics of forward rates. Such derived dynamics will however imply a nontrivial terminal swap rate distribution which is contradiction with the log-normal distribution implied by the swap-market model in the form presented here. Therefore if we truly beleive we have managed to capture the true dynamics of forward rates, to obtain a value of a swaption we would rather want to rely on this LIBOR rate model than on a swap market model. The next section will deal with this topic.

5 Calibrating LIBOR Market Model

In the previos chapter it was shown, that the linear form of volatility function in the state variable $F(t,T,S)$ (implying its lognormal terminal distribution) leads to a Black formula of a caplet/floorlet. The formula was derived under the no-arbitrage conditions using fundamental pricing formula, where the expectations about the instrument's payoff under the Q^S measure in equation 69 was substituted by a formula for expected value of difference of a maximum of two log-normaly distributed variables (formula 70). The same result would be obtained if we decide to calculate the expected value via Monte Carlo simlation under the relevant probability measure. If we want to price an instrument, who's payoff function is more complicated, Monte Carlo simulation will be the only way of calculating the expectation in fundamental pricing formula. Furthermore if the evaluated instrument's payoff depends on more than only one variable, we will need to simulate evolution of all of these relevant underlyings. This is the case of instruments who's payoff depends for instance on a spread between two rates. Other example can be say a 1 year to 2 years swaption, an option to enter into a swap contract lasting two years starting one year in the future. Because the value of such swap one year in the future depends on the value of 6-month, 12-month, 18-month and 24-month forward LIBOR rate in one year, all four rates have to be simulated so that the swaption's payoff can be evaluated. To go ahead with this task one needs to know not only the exact shape of volatility functions ($\sigma(t,S,T)$) of all four forward LIBOR rates over the period of one year, but also their expected correlation structure.

Besides the fact, that there is a great variety of products who's payoff depends directly on evolution of the market rates (LIBOR rate, swap rate), the main reason, why the market models became very popular is the relative ease of the model's calibration (finding the right volatility and correlation structures) to the prices of caps and floors (in case of LIBOR market model) and swaptions (in case of swap market model). These are the most liquid interest rate derivatives and their prices should therefore reflect well the market expectation of the future volatility of the relevant rates.

In the two sections ahead a LIBOR market model will be build and calibrated to market data from 11^{th}November 2005. Subsequently various swaptions will be priced for which market data are available and the model's performance can thus be tested. The following chapter will start with deriving the model for six month forward rates (six months is the frequency with which the exchange of payments takes place in swap contracts underlying the traded swaptions in EUR) and the next two chapters in this section will elaborate on the issue of calculating the volatility and correlation functions.

5.1 The Model

The task is to price a set of swaptions with Monte Carlo simulation. To do that we need to simulate the evolution of forward rates which determine the value of a swap rate in one year of time - see equation 47. In general, each of the forward rates can be driven by a different Wiener process, so that we can have as many sources of risk as forward rates that are modelled. It is intuitive that these Wiener processes will be correlated, if 6-month forward rate maturing in one year moves up it is very likely that a 6-month forward rate maturing in two years moves up as well. It might however be argued, that using more than let's say three Wiener processes to simulate an evolution of a set of forward rates doesn't bring additional benefit in comparison with the increased complexity. Later in the chapter it will be shown how the model with n forward rates can be compressed to have only a limited number of risk factors included in the model.

5.1.1 One Factor Case

Compression of the model to one factor is straightforward. One simply assumes that the same Wiener process drives innovation of all modelled forward rates, while the magnitude of the change is given by forward rate's individual volatility function. The correlation between any two forward rate will then obviously be one. First step to be done is to choose a probability measure under which all rates will be modeled so that they were all in the same probability space. Under the chosen measure at most one forward rate can however be a martingale. If we model the evolution of all relevant rates under the measure $Q^{1.0}$, only forward rate $F(t, 0.5, 1.0)$ will be driftless. To simplify notation, let's have $F_{1.0}(t) = F(t, 0.5, 1.0)$ and the same for σs and μs (which also depend on a state variable $F_{t_i}(t)$). The dynamics of n rates will than look as follows:

$$\begin{aligned} dF_{1.0}(t) &= 0dt + \sigma_{1.0}(t)F_{1.0}(t)dW(t)^{1.0} \\ dF_{1.5}(t) &= \mu_{1.5}(t)dt + \sigma_{1.5}(t)F_{1.5}(t)dW(t)^{1.0} \\ dF_{2.0}(t) &= \mu_{2.0}(t)dt + \sigma_{2.0}(t)F_{2.0}(t)dW(t)^{1.0} \\ &\vdots \\ dF_{t_n}(t) &= \mu_{t_n}(t)dt + \sigma_{t_n}(t)F_{t_n}(t)dW(t)^{1.0} \end{aligned} \tag{77}$$

Note that in line with the notation used in the previous chapters, $t_0 = 0.5$, 1/2 year from current moment, $t_1 = 1.0$, 1 year from current moment and so on.[36] As mentioned in the first chapter, if we employ the Girsanov Theorem to switch the probability measures, the volatility function remains the same while only the drift function changes. It can be shown that the arbitrage-free drift function of the forward rates $F_{1.0}, ..., F_n$ under the probability measure $Q^{1.0}$ with one risk factor is:[37]

$$\mu_i(t) = \sigma_{t_i}(t)F(t, t_i, t_{i+1}) \sum_{j=2}^{i} \frac{\tau\sigma_{t_j}(t)F_{t_j}(t)}{1 + \tau F_{t_j}(t)}, \quad \text{for } i = 2, ..., n \tag{78}$$

5.1.2 N Factor case

If we want to be more general, we can allow each of the Wiener process to be different from each other. The general dynamics will than be:

$$\begin{aligned} dF_{1.0}(t) &= 0dt + \sigma_{1.0}(t)F_{1.0}(t)dW_{1.0}(t)^{1.0} \\ dF_{1.5}(t) &= \mu_{1.5}(t)dt + \sigma_{1.5}(t)F_{1.5}(t)dW_{1.5}(t)^{1.0} \end{aligned}$$

[36] The forward rate $F_{1.0}$ will not be simulated, it is included for completness. Thus t_0 is actually equal to one year, which is the first expirty of the swaptions evaluated.

[37] See Branger and Schlag (2004) for derivation.

$$dF_{2.0}(t) = \mu_{2.0}(t)dt + \sigma_{2.0}(t)F_{2.0}(t)dW_{2.0}(t)^{1\,0}$$
$$\vdots \qquad (79)$$
$$dF_{t_n}(t) = \mu_{t_n}(t)dt + \sigma_{t_n}(t)\,F_{t_n}(t)\,dW_{t_n}(t)^{1\,0}$$

Let's denote a vector $dW(t)^T = \{dW_{1\,0}(t), dW_{1\,5}(t), ..., dW_{t_n}(t)\}$. Then

$$dW(t)dW(t)^T = \begin{pmatrix} dt & \rho_{1.0,1.5}dt & \cdots & \rho_{1.0,t_n}dt \\ \rho_{1.5,1.0}dt & dt & \cdots & \rho_{1.5,t_n}dt \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{t_n,1.0}dt & \rho_{t_n,1.5} & \cdots & dt \end{pmatrix} \qquad (80)$$

is the transition correlation matrix of the system. All ρs in the above matrix can be functions of time (and also state variable), for simplicity only constants will be considered here.

Similarily as in the one factor case above, it can be shown, that the arbitrage-free drift function of the forward rates $F_{1.0}, ..., F_n$ is:[38]

$$\mu_i(t) = \sigma_{t_i}(t)F(t, t_i, t_{i+1})\sum_{j=2}^{i} \frac{\rho_{t_i,t_j}\tau\sigma_{t_j}(t)F_{t_j}(t)}{1+\tau F_{t_j}(t)}, \quad \text{for } i = 2, ..., n\,. \qquad (81)$$

From the above dynamics functions it is clear, that the whole model is completely defined if we know the transition correlation matrix and the volatility functions of all simulated forward rates.

5.1.3 Reducing Dimensionality of the Model

The above formulation of the model is not very handy for simulation purposes for couple of reasons. First it is akward to have to generate a set of random numbers which are correlated according to a prescribed correlation matrix (equation 80). Second it can be argued that n risk factors are too many and that such complexity of the model doesn't bring additional benefit.[39] Rebonato shows a very handy way, how the model can be reformulated using independent Wiener processes.[40] Furthermore, this formulation enables to easily reduce the dimensionality of the model.

First it can be shown, that the system of n equations above can be rewritten as follows using n independent Wiener processes denoted $\bar{W}_i(t)$, all under the measure $Q^{1\,0}$ (appropriate drift term is to be substituted on the begining of each forward's dynamics):

$$\frac{dF_{1.0}(t)}{F_{1.0}(t)} = 0 + \sigma_{1,1}(t)d\bar{W}_1(t) + \sigma_{1,2}(t)d\bar{W}_2(t) + ... + \sigma_{1,n}(t)d\bar{W}_n(t)$$
$$\frac{dF_{1.5}(t)}{F_{1.5}(t)} = ... + \sigma_{2,1}(t)d\bar{W}_1(t) + \sigma_{2,2}(t)d\bar{W}_2(t) + ... + \sigma_{2,n}(t)d\bar{W}_n(t)$$

[38] See Brigo and Mercurio (2001).
[39] See for instance Fan et al. (2002).
[40] See Rebonato (2004).

$$
\begin{aligned}
\frac{dF_{2.0}(t)}{F_{2.0}(t)} &= \ldots + \sigma_{3,1}(t)dW_1(t) + \sigma_{3,2}(t)d\bar{W}_2(t) + \ldots + \sigma_{3,n}(t)dW_n(t) \\
&\vdots \\
\frac{dF_{t_n}(t)}{F_{t_n}(t)} &= \ldots + \sigma_{n,1}(t)dW_1(t) + \sigma_{n,2}(t)dW_2(t) + \ldots + \sigma_{n,n}(t)dW_n(t).
\end{aligned} \tag{82}
$$

The goal now is to show that the sigmas in the set of equations 82 can be specified in such way, that the correlation structure as well as the volatility functions of equations 80 is preserved. In order for the volatility to be retrieved we use the relationship for addition of two independent stochastic variables - $var(aX+bY) = a^2Var(X)+b^2Var(Y)$. Because our new Wiener processes are independent from definition, the following relationship has to hold:

$$
\sigma_{t_i}(t)^2 = \sum_{j=1}^{n} \sigma_{i,j}(t)^2, \tag{83}
$$

where $\sigma_{t_i}(t)$ is from the set of equations 80. This way the volatility of each forward from the set of equations 82 is

$$
Var(dF_{t_i}) = \sum_{j=1}^{n} \sigma_{i,j}(t)^2 F_{t_i}^2 dt = \sigma_{t_i}(t)^2 F_{t_i}^2 dt. \tag{84}
$$

Clearly, $\sigma_{i,j}$s can be chosen in an infinity of ways so that condition 83 was satisfied. Additional restrictions has to be imposed so that correlation matrix 80 was recovered as well. Instead of trying to recover conditions for individual $\sigma_{i,j}$s, we can restate our problem as follows. First we can multiply and divide each diffusion term in each equation from the set 82 by the forward's desired sigma function:

$$
\begin{aligned}
\frac{dF_{1.0}(t)}{F_{1.0}(t)} &= \ldots + \sigma_{1.0}(t)\left(\frac{\sigma_{1,1}(t)}{\sigma_{1.0}(t)}d\bar{W}_1(t) + \frac{\sigma_{1,2}(t)}{\sigma_{1.0}(t)}d\bar{W}_1(t) + \ldots + \frac{\sigma_{1,n}(t)}{\sigma_{1.0}(t)}d\bar{W}_1(t)\right) \\
\frac{dF_{1.5}(t)}{F_{1.5}(t)} &= \ldots + \sigma_{1.5}(t)\left(\frac{\sigma_{2,1}(t)}{\sigma_{1.5}(t)}d\bar{W}_1(t) + \frac{\sigma_{2,2}(t)}{\sigma_{1.5}(t)}d\bar{W}_1(t) + \ldots + \frac{\sigma_{2,n}(t)}{\sigma_{1.5}(t)}d\bar{W}_1(t)\right) \\
\frac{dF_{2.0}(t)}{F_{2.0}(t)} &= \ldots + \sigma_{2.0}(t)\left(\frac{\sigma_{3,1}(t)}{\sigma_{2.0}(t)}d\bar{W}_1(t) + \frac{\sigma_{3,2}(t)}{\sigma_{2.0}(t)}d\bar{W}_1(t) + \ldots + \frac{\sigma_{3,n}(t)}{\sigma_{2.0}(t)}d\bar{W}_1(t)\right) \\
&\vdots \\
\frac{dF_{t_n}(t)}{F_{t_n}(t)} &= \ldots + \sigma_{t_n}(t)\left(\frac{\sigma_{n,1}(t)}{\sigma_{t_n}(t)}d\bar{W}_1(t) + \frac{\sigma_{n,2}(t)}{\sigma_{t_n}(t)}d\bar{W}_1(t) + \ldots + \frac{\sigma_{n,n}(t)}{\sigma_{t_n}(t)}d\bar{W}_1(t)\right).
\end{aligned}
$$

Now we can take a square root of the condition 83 and substitute in each of the equations above for $\sigma_{t_i}(t)$ in denominator of each of the terms in brakets. Those terms will then take to following form:

$$
b_{i,j} = \frac{\sigma_{i,j}(t)}{\sigma_{t_i}(t)} = \frac{\sigma_{i,j}(t)}{\sqrt{\sum_{j=1}^{n}\sigma_{i,j}(t)^2}} \tag{85}
$$

and the set of equation can be rewritten as follows:

$$
\frac{dF_{1.0}(t)}{F_{1.0}(t)} = \ldots + \sigma_{1.0}(t)\left(b_{1,1}d\bar{W}_1(t) + b_{1,2}d\bar{W}_1(t) + \ldots + b_{1,n}d\bar{W}_1(t)\right)
$$

$$\frac{dF_{1.5}(t)}{F_{1.5}(t)} = ... + \sigma_{1.5}(t)\left(b_{2,1}d\bar{W}_1(t) + b_{2,2}d\bar{W}_1(t) + ... + b_{2,n}dW_1(t)\right)$$
$$\frac{dF_{2.0}(t)}{F_{2.0}(t)} = ... + \sigma_{2.0}(t)\left(b_{3,1}d\bar{W}_1(t) + b_{3,2}d\bar{W}_1(t) + ... + b_{3,n}dW_1(t)\right)$$
$$\vdots \tag{86}$$
$$\frac{dF_{t_n}(t)}{F_{t_n}(t)} = ... + \sigma_{t_n}(t)\left(b_{n,1}dW_1(t) + b_{n,2}d\bar{W}_1(t) + ... + b_{n,n}(t)d\bar{W}_1(t)\right).$$

Looking back at the volatility condition 83, we can restate it by dividing the equation by $\sigma^2_{t_i}(t)$:

$$1 = \sum_{j=1}^{n} \frac{\sigma_{i,j}(t)^2}{\sigma_{t_i}(t)^2}. \tag{87}$$

Comparing the result with equation 85 we obtain a refrased volatility condition

$$1 = \sum_{j=1}^{n} b^2_{i,j}. \tag{88}$$

which ensures, that the volatilities $\sigma_{t_i}(t)$ of diffusion terms in set of equations 86 are preserved.

This formulation of the problem (equations 86) is particulary handy, because it decomposes the dynamics into volatility and correlation. It can be appreatiated when we write the formula of instantaneous percantage covariance of any two forward rates $F_{t_i}(t)$ and $F_{t_j}(t)$ and substitute from the above dynamics 86:

$$Cov\left(\frac{dF_{t_i}}{F_{t_i}}, \frac{dF_{t_j}}{F_{t_j}}\right) = E\left[\frac{dF_{t_i}}{F_{t_i}}\frac{dF_{t_j}}{F_{t_j}}\right] - E\left[\frac{dF_{t_i}}{F_{t_i}}\right]E\left[\frac{dF_{t_j}}{F_{t_j}}\right] = E\left[\frac{dF_{t_i}}{F_{t_i}}\frac{dF_{t_j}}{F_{t_j}}\right] =$$
$$= \sigma_{t_i}(t)\sigma_{t_j}(t)E\left[b_{i,1}b_{j,1}dW_1(t)^2 + b_{i,1}b_{j,2}W_1(t)W_2(t) + ... + b_{i,1}b_{j,n}(t)W_1(t)\bar{W}_n(t) + \right.$$
$$\left. ... + ... + ... + b_{i,n}(t)b_{j,1}(t)\bar{W}_n(t)\bar{W}_1(t) + ... + b_{i,n}b_{j,n}\bar{W}_n(n)^2\right].$$

The expected value of sums is equal to the sum of expected values and therefore the expected values can be evaluated term by term. The terms with dt^2 or $dtdW_i(t)$ tend to *zero*. All the terms with $E\left[...\bar{W}_i(t)\bar{W}_j(t)\right]$ where $i \neq j$ are covariances of independent Wiener processes and are therefore also *zero*. The terms where $i = j$ are variances of two Wiener processes and are thus equal dt.[41] The transition covariance can thus be rewritten as

$$Cov\left(\frac{dF_{t_i}}{F_{t_i}}, \frac{dF_{t_j}}{F_{t_j}}\right) = \sigma_{t_i}(t)\sigma_{t_j}(t)\left[b_{i,1}b_{j,1} + ... + b_{i,n}b_{j,n}\right]dt, \tag{89}$$

which implies that

$$b_{i,1}b_{j,1} + ... + b_{i,n}b_{j,n} = \rho_{i,j}, \tag{90}$$

where ρ_{t_i,t_j} is a correlation between rates $F_{t_i}(t)$ and $F_{t_j}(t)$ - (see the correlation matrix 80).

[41] See Bjork (2004) for moments of independent and dependent stochastic variables or also Appendix A.

The formulation of dynamics of forward rates with $b_{i,j}$s allows to model separately the volatility functions $\sigma_{t_i}(t)$ and correlation matrix by defining the $b_{i,j}$s in such way, that the volatility condition 88 was satisfied and the desired correlation matrix 80 was recovered if n factors are kept. If we decide to reduce the number of factors, we have to accept the fact that the implied correlation matrix will differ from the desired one. Let's say that we want to build a model with only two factors. The set of equations will than look as follows:

$$\begin{aligned}
\frac{dF_{1.0}(t)}{F_{1.0}(t)} &= \; ... + \sigma_{1.0}(t)\left(b_{1,1}dW_1(t) + b_{1,2}dW_2(t)\right) \\
\frac{dF_{1.5}(t)}{F_{1.5}(t)} &= \; ... + \sigma_{1.5}(t)\left(b_{2,1}dW_1(t) + b_{2,2}dW_2(t)\right) \\
\frac{dF_{2.0}(t)}{F_{2.0}(t)} &= \; ... + \sigma_{2.0}(t)\left(b_{3,1}dW_1(t) + b_{3,2}dW_2(t)\right) \\
&\vdots \\
\frac{dF_{t_n}(t)}{F_{t_n}(t)} &= \; ... + \sigma_{t_n}(t)\left(b_{n,1}dW_1(t) + b_{n,2}dW_2(t)\right)
\end{aligned} \tag{91}$$

One has to then choose the parameters b in such way, that the volatility condition 88 held and that the difference between the desired correlation matrix was minimal. How to specify the volatility functions and how the estimate the b will be the subject of next chapters.

5.2 Volatility Function of Forward Rates

When determining the volatility function of the forward rates (meant $\sigma_F(t,S,T)$),[42] As mentioned earlier in the thesis, one has a choice to either make an own guess or look at what market implies about volatility in the future. This information can be obtained from instruments which are traded on the market, namely caps and floors. The quotes of caps in EUR are given directly in terms of implied volatilities (see section 4.3.1) of the ATM Caps with the expiry of first caplet equal 6 months. They are called *flat volatilities* and will be denoted $\sigma_F(t_n)$, where t_n is the expiry of last caplet contained in the cap. There are caps with following tenors quoted on the market in EUR:

$Cap_1(0, 0.5, \mathbf{1}, S(0, 0.5, \mathbf{1}))$, $Cap_2(0, 0.5, \mathbf{2}, S(0, 0.5, \mathbf{2}))$, $Cap_3(0, 0.5, \mathbf{3}, S(0, 0.5, \mathbf{3}))$, $Cap_4(0, 0.5, \mathbf{4}, S(0, 0.5, \mathbf{4}))$, $Cap_5(t, 0.5, \mathbf{5}, S(0, 0.5, \mathbf{5}))$, $Cap_7(0, 0.5, \mathbf{7}, S(0, 0.5, \mathbf{7}))$, $Cap_{10}(0, 0.5, \mathbf{10}, S(0, 0.5, \mathbf{10}))$.

where the strike is a swap rate with relevant tenor one 1/2 year forward. The market observed implied volatilities on 11th November 2005 were:

[42] To be exact one should say percentage volatility function when talking about $\sigma_F(t,S,T)$. To see why, see first section for dynamics of a logarithm of a financial asset, which is basically asset's percentage dynamics. For brevity, by volatility function $\sigma_F(t,S,T)$ will be meant from now on as a concrete function of time.

Market flat volatilities in %, 11-11-2005						
$\sigma_F(1)$	$\sigma_F(2)$	$\sigma_F(3)$	$\bar{\sigma}_F(4)$	$\sigma_F(5)$	$\bar{\sigma}_F(7)$	$\sigma_F(10)$
18.13	20.74	21.49	21.64	21.47	20.79	19.68

The cap Cap_1 consist of one caplet who's underlying is a 6-month forward rate $F(0, 0.5, 1)$. Cap Cap_2 consists of three caplets with underlying forward rates $F(0, 0.5, 1)$, $F(0, 1, 1.5)$ and $F(0, 1.5, 2)$, cap Cap_3 consists of five caplets and so on. The quote $\sigma_F(\mathbf{2})$ is such constant, which recovers the cap price when substituted into Black formula for cap, equation 73 - $Cap_2 = Cap(0, 0.5, 2, S(0, 0.5, 2), \sigma_F(\mathbf{2}))$. In other words, $\sigma_F(2)$ is the market-implied constant volatility function of all three forward rates, which are underlyings of the caplets contained in Cap_2.

One can make an objection, that the market quotes of flat volatilities are contradictory, when the quote $\bar{\sigma}_F(\mathbf{1})$ implies a constant volatility of underlying forward rate $F(0, 0.5, 1)$ of 18.13% and at the same time quote $\bar{\sigma}_F(\mathbf{2})$ implies constant volatility of underlying forward rate $F(0, 0.5, 1)$ (and that of $F(0, 1, 1.5)$) of 20.74%. This objection is justified. The Black formula is however just a tool to convert the premium (price) in EUR of a cap into flat volatility, which is more informative quantity than the price.

Consistent (constant) market-implied volatilites of forward rates - those implied by caplets when substituted in Black pricing formula for a caplet (formula 72) - $\sigma_F(t_{i-1}, t_i)$ for $i = 0, ..., n$ can however be recovered from the quoted flat volatilities. The first cap $Cap_1(0, 0.5, \mathbf{1}, K)$ with strike $K = S(0, 0.5, \mathbf{1})$ consists of only one caplet and thus $\sigma_F(1) = \bar{\sigma}_F(0.5, 1)$. This $\bar{\sigma}_F(0.5, 1)$ can be substituted into a Black formula of a caplet $Cpl(0, 0.5, 1, K, \sigma)$ with strike $K = S(0, 0.5, \mathbf{1})$ to recover the caplet premium in EUR. The price of a $Cap_{1.5}$ consists of premiums for two caplets:

$$Cap_{1.5}(0, 0.5, \mathbf{1.5}, K) = Cpl(0, 0.5, 1, K) + Cpl(0, 1, 1.5, K),$$

where strike $K = S(0, 0.5, \mathbf{1.5})$. If we make an assumption, that there are no smiles[43] in the market, we can use the implied volatility $\bar{\sigma}_F(0.5, 1)$ to be substituted into Black's formula to back up the premium of $Cpl(0, 0.5, \mathbf{1}, S(0, 0.5, \mathbf{1.5}))$. The premium $Cpl(0, 0.5, \mathbf{1}, S(0, 0.5, \mathbf{1}))$ is known and if there was a quote $\sigma_F(\mathbf{1.5})$ available, it could be substituted into a Black formula for cap which would deliver the price of cap $Cap_{1.5}$ and the premium $Cpl(0, 1, \mathbf{1.5}, S(0, 0.5, \mathbf{1.5}))$ could be easily backed up. This is unfortunately not the case as the "nearest" quote is $\sigma_F(\mathbf{2})$. One can however interpolate the missing flat volatilities, either lineary or more optimal using a cubic spline.[44] The following table shows the complete flat volatilities recovered using natural cubic spline interpolation:[45]

[43]Smile is a plot of option's strike prices on X-axes and implied volatilities on Y-axes. See Hull (1993). If the market really beleives, that the underlying asset (forward rate in case of a caplet) is log-normaly distributed, than the smile will be a horizontal line. Usually one can observe smile-like shape, which implies that the expected underlying's terminal distribution has "fatter tails" - more mass on extremes.

[44]see Nummerical Recipies (1988 - 1992).

[45]Natural cubic spline assumes zero second derivation of the obtained function at the outside points.

Interpolated flat volatilities in %, 11-11-2005	
$\bar{\sigma}_F(1.0)$	18.13
$\bar{\sigma}_F(1.5)$	19.61
$\sigma_F(2.0)$	20.74
$\sigma_F(2.5)$	21.29
$\sigma_F(3.0)$	21.49
$\bar{\sigma}_F(3.5)$	21.60
$\sigma_F(4.0)$	21.64
$\sigma_F(4.5)$	21.59
$\sigma_F(5.0)$	21.47
$\sigma_F(5.5)$	21.32
$\sigma_F(6.0)$	21.15
$\sigma_F(6.5)$	20.98
$\sigma_F(7.0)$	20.79
$\sigma_F(7.5)$	20.60
$\bar{\sigma}_F(8.0)$	20.42
$\bar{\sigma}_F(8.5)$	20.23
$\sigma_F(9.0)$	20.05
$\sigma_F(9.5)$	19.86
$\sigma_F(10.0)$	19.68

The quote for a cap $Cap_{1.5}$ is now available and the premium $Cpl(0, 1, 1.5, S(0, 0.5, 1.5))$ can be calculated. This method can be used to recover the premiums of the remaining caplets. Now using the Black formula for a caplet, one could calculate the consistent market implied constant volatilities of forward rates. Unfortunately the Black formula cannot be inverted to solve for σ and an iterative numerical method has to be employed to solve for relevant caplet's σ.

When writing a programm to recover these volatilities quickly, an alternative mehtod can be used which offers an analitical recursive formula which can be automated in a typical "for-cycle" code. One can write cap and the relevant caplets in dependence of only the underlying forward's volatility as follows:

$$Cap_n(\bar{\sigma}_F(t_n)) = \sum_{i=1}^{n} Cpl_i(\sigma_F(t_{i-1}, t_i)).$$

One approximates the right side of the above equation using first order Taylor expansion around point $\bar{\sigma}_F(t_n)$ for each of the caplets in the sum:

$$\sum_{i=1}^{n} Cpl_i(\bar{\sigma}_F(t_{i-1}, t_i)) \approx \sum_{i=1}^{n} \left(Cpl_i(\bar{\sigma}_F(t_n)) + (\bar{\sigma}_F(t_{i-1}, t_i) - \bar{\sigma}_F(t_n)) \frac{\partial Cpl_i(\bar{\sigma}_F(t_{i-1}, t_i))}{\partial \bar{\sigma}_F(t_{i-1}, t_i)} \right). \tag{92}$$

The partial derivations, all taken at points $\sigma_F(t_{i-1}, t_i) = \bar{\sigma}_F(t_n)$, are denoted *vega*$-\nu_i$. Because by definition (see relationship 73) $Cap_n(\sigma_F(t_n)) = \sum_{i=1}^{n} Cpl_i(\sigma_F(t_n))$, the sum of the terms with *vegas* in 92 has to be approximately *zero* and from here

$$\sum_{i=1}^{n} (\bar{\sigma}_F(t_{i-1}, t_i) - \bar{\sigma}_F(t_n)) \frac{\partial Cpl_i(\sigma_F(t_{i-1}, t_i))}{\partial \bar{\sigma}_F(t_{i-1}, t_i)} \approx 0$$

$$\sum_{i=1}^{n} \bar{\sigma}_F(t_{i-1}, t_i)\nu_i \approx \bar{\sigma}_F(t_n) \sum_{i=1}^{n} \nu_i$$
$$\sigma_F(t_n) \approx \frac{\sum_{i=1}^{n} \bar{\sigma}_F(t_{i-1}, t_i)\nu_i}{\sum_{i=1}^{n} \nu_i}. \tag{93}$$

The relationship 93 can be easily used to recover the constant volatilities of forward rates $\bar{\sigma}_F(t_{i-1}, t_i)$ by first setting $\sigma_F(0.5, 1) = \bar{\sigma}_F(1)$, then solving for $\sigma_F(1, 1.5)$ from

$$\sigma_F(1.5) = \frac{\nu_1 \bar{\sigma}_F(0.5, 1) + \nu_2 \bar{\sigma}_F(1, 1.5)}{\nu_1 + \nu_2}$$

and so on.

The formula for caplet's vega can is obtained by deriving Black's formula of caplet (equaiton 72) with respect to σ (see Appendix C for derivation):

$$vega = \frac{\partial Cpl(t, t_{i-1}, t_i, K, \sigma)}{\partial \sigma} = P(t, t_i) F(t, t_{i-1}, t_i) \sqrt{t_{i-1} - t} \frac{\tau(t_{i-1}, t_i)}{\sqrt{2\pi}} \exp\left\{\frac{d^2}{-2}\right\},$$

where d is the same as in Black's formula for caplet (equation 72).

In order to derive vegas one needs to derive the EUR term structure of interest rates. It will also be needed later on when simulating the forwards evolution. The EURIBOR rates are published every day in the morning by the European Banking Federation (FBE), however only maturities up to 1 year are available. If we wish to calculate vegas to derive Black volatilities of forward rates maturing in more than one year (as well as simulate the evolution of forward rates), we will need much longer yield curve than the one published by FBE. The common practice is to derive this term structure from the money market, future and swap markets, which is very liquid and thus best reflects the shape of a yield curve. The formula for a swap rate (equation 45) shows its relationship with zero bond prices. The swaps which are quoted on the swap market start today, or $t = t_0$ in formula 45, and quotes are available for $t_n = 1, 2, ..., 9, 10, 12, 15, 20, 25, 30$ years. Starting with swap rate $S(0, 0, 0.5)$ and substituting for ts we have

$$S(0, 0, 0.5) = \frac{1 - P(0, 0.5)}{\tau P(0, 0.5)} = L(0, 0.5).$$

The EURIBOR rate L(0,0.5) quoted daily can thus be taken as S(0,0,0.5) and P(0,0.5) can be derived. Next we take the market quote for S(0,0,1) to back up zero bond P(0,1). To recover zero bond P(0,1.5), we need a quote for S(0,0,1.5), which is not available. Here again we can interpolate to recover the missing swap rates 1.5, 2.5, ... Table below shows the interpolated swap rates (again using natural cubic spline method) for up to 10 years, the recovered zero bond prices and corresponding forward rates:

Term structure of interest rates, 11-11-2005			
Year	Swap Rate	Zero Bond Price	Forward Rate
0.5	2.4760	0.9878	0.0248
1.0	2.7010	0.9735	0.0293
1.5	2.8360	0.9586	0.0311
2.0	2.9140	0.9437	0.0316
2.5	2.9864	0.9284	0.0329
3.0	3.0560	0.9128	0.0342
3.5	3.1156	0.8972	0.0349
4.0	3.1680	0.8815	0.0356
4.5	3.2173	0.8657	0.0364
5.0	3.2640	0.8499	0.0372
5.5	3.3080	0.8341	0.0379
6.0	3.3500	0.8182	0.0387
6.5	3.3908	0.8024	0.0394
7.0	3.4300	0.7866	0.0401
7.5	3.4672	0.7710	0.0407
8.0	3.5030	0.7554	0.0413
8.5	3.5379	0.7398	0.0420
9.0	3.5710	0.7245	0.0424
9.5	3.6013	0.7094	0.0426
10.0	3.6300	0.6944	0.0430

Substituting the missing variables into the formulas for *vegas*, the constant percentage volatilities of forward rates can be calculated as described above. The result for the volatilities on 11th November 2005 is:

Black volatilities of forwards, 11-11-2005	
$\sigma_F(0.5, 1.0)$	18.13
$\sigma_F(1.0, 1.5)$	20.63
$\sigma_F(1.5, 2.0)$	22.28
$\sigma_F(2.0, 2.5)$	22.38
$\sigma_F(2.5, 3.0)$	21.94
$\sigma_F(3.0, 3.5)$	21.89
$\sigma_F(3.5, 4.0)$	21.74
$\sigma_F(4.0, 4.5)$	21.26
$\sigma_F(4.5, 5.0)$	20.75
$\sigma_F(5.0, 5.5)$	20.33
$\sigma_F(5.5, 6.0)$	19.94
$\sigma_F(6.0, 6.5)$	19.55
$\sigma_F(6.5, 7.0)$	19.20
$\sigma_F(7.0, 7.5)$	18.88
$\sigma_F(7.5, 8.0)$	18.57
$\sigma_F(8.0, 8.5)$	18.26
$\sigma_F(8.5, 9.0)$	17.94
$\sigma_F(9.0, 9.5)$	17.61
$\sigma_F(9.5, 10.0)$	17.28

The above constant volatilities are now consistent. It can be observed, that the volatilities are first rising with the expiry, they reach maximum for the forward maturing in two years and than sink for the remaining rates. This fenomenon can be explained by the activities of monetary authorities. On one hand the central bank tries to send clear signals about its monetary policy which is translated into its activity on the money market. Thus the short end of the forward rate curve is not very volatile as surprises come rather rarely. On the other hand, the volatility of the long end of the forward rate curve is given by the market's changing expectations about the future inflation. If the monetary authority is credible and the markets trust central bank's explicit or implicit inflation targets, the volatility of the distant forward rates will be low as well. The greatest uncertainty is about the forward rates in the middle maturity spectrum, where the expectations about loose or tight monetary policy might change fast as new information about the state of economy reach the market.

When we simulate a forward rate which matures in say five years - $F(0,5,5.5)$, it would not be very realistic if we used the constant cap-implied volatility $\sigma_F(5.0,5.5)$ for the whole time of the forward's life. As discused above, we would rather expect that in about two years the volatility of the forward ($F(2,5,5.5)$), will be on its maximum and will than gradualy decrease as the forward matures in five years from now - $F(5,5,5.5)$. It should be pointed out, that the market doesn't say anything about the exact shape of the volatility function of a forward rate, it only says that the constant $\sigma_F(5.0,5.5)$ plugged into the Black formula for a caplet recovers the relevant caplet's premium in EUR. If we look back at the relationship for a no-arbitrage price of a caplet in the previous section (equation 71), we can see that we will get exactly the same result if we allow the volatility function to dependend on time as long as

$$\int_0^5 \sigma_F(s,5,5.5)^2 ds = \bar{\sigma}_F(5.0,5.5)^2 \cdot 5.$$

By rearranging

$$\sigma_F(5.0,5.5)^2 = \frac{1}{5}\int_0^5 \sigma_F(s,5,5.5)^2 ds,$$

we can see that the market-implied constant volatility $\sigma_F(5.0,5.5)$ can be interpreted as a square root of an average percentage variance of a forward rate $F(0,5,5.5)$. Thus a much more realistic approach would be to choose a suitable parametrical form for a forward volatility in line with the above discussion and demand that when integrated it satisfies condition

$$\int_0^{t_i} \sigma_F(s,t_i,t_{i+1})^2 ds = \bar{\sigma}_F(t_i,t_{i+1})^2 \cdot t_i. \qquad (94)$$

One such form which allows for a hump in its shape is the following function of time to maturity $(t_i - t)$:

$$\sigma_F(t,t_i,t_{i+1}) = [a \cdot (t_i - t) + b] \cdot e^{-c(t_i - t)} + d. \qquad (95)$$

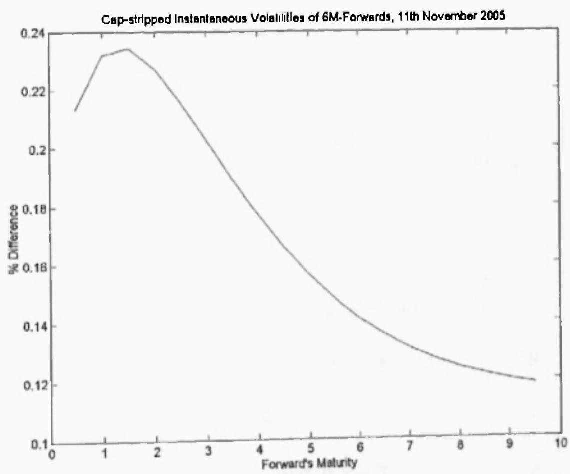
Because exactly four parameters are available, only four caplet prices can be recovered exactly. In our case 19 caplet prices are to be recovered and therefore the obtained volatility function will not price the market observed cap prices exactly. As will be shown later on the discrepancies are however quite small. In order to recover the cap prices as well as possible, one has to find such paramters a, b, c and d, that the condition 94 is satisfied as well as possible for all caplet expiries t_i in question. One possibility is to minimize the following sum of squared residuals:

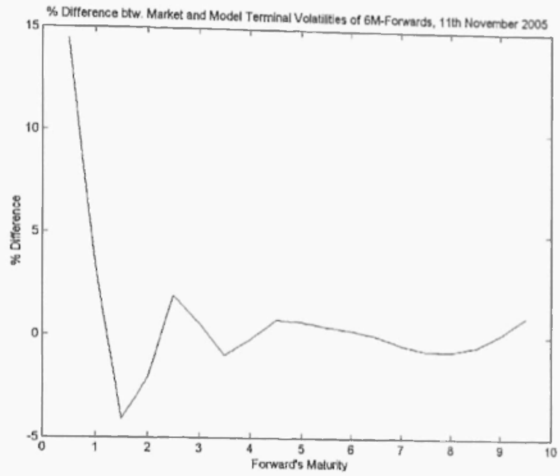
$$SSR = \sum_{i=0}^{18} \left(\sigma_F(t_i, t_{i+1})^2 t_i - \int_0^{t_i} \left([a \cdot (t_i - t) + b] \cdot e^{-c(t_i - t)} + d \right)^2 dt \right)^2 .$$

The integral was calculated with help of MatLab and iterative algorithm was created to minimize SSR with respect to a, b, c and d. The table below presents the results with exactness of four decimal places.

Parameters of volatility function, 11th November 2005	
a	0.1619
b	0.0549
c	0.5987
c	0.1128
SSR	0.0000541321

The first graph below shows the exact shape of the volatility function and the second graph shows percentage difference between market-implied terminal forward volatilities and those implied by the model (% difference between left and right side of condition 94). Except of the volatility of forward $F(0.0, 0.5, 1.0)$ no volatilities differ by more than 5%, most of them by no more than 1%. The differences are minimal and thus employing the parametrical form of forward volatility 95 with the above estimated coeffitiens we are able to recover the observed prices of caplet (caps) on the market.





5.3 Correlations of Forward Rates

Having estimated the volatility function of the forward rates, one can go ahead with Monte-Carlo simulation in a one factor model, where no other parameters need to be estimated (see set of equations 77 and expression for drift 78). In the next section, three models will be used to model the evolution of a yield curve - one factor model, two factor model and three factor model. So that we were able to go ahead with this task, correlation matrix has to be modelled to provide the missing information in the drift of forward rate dynamics (see expresion 81).

5.3.1 Correlations in 2 Factor Model

As shown in chapter 5.1.3, the problem of modeling correlation matrix under the condition, that the cap prices implied by calibrating the volatility function to the market quotes are preserved, reduces to estimating the coefitients b in set of equation 91. This will be automatically truth as long as we choose the bs in line with condition 88. In our case of two factors expression

$$b_{i,1}^2 + b_{i,2}^2 = 1 \ , \ \text{for } i = 1, ..., 19 \tag{96}$$

must hold. Rebonato[46] has pointed out, that the condition 96 is a definition of coordinates of points laying on a circle in a 2D-plane with radius 1. Setting

$$b_{i,2} = \cos(\theta_i) \quad \text{and} \quad b_{i,1} = \sin(\theta_i) \quad \text{for} \quad i = 1, ..., 19 \tag{97}$$

the condition 96 can be rewritten as

$$\sin^2(\theta_i) + \cos^2(\theta_i) = 1 \tag{98}$$

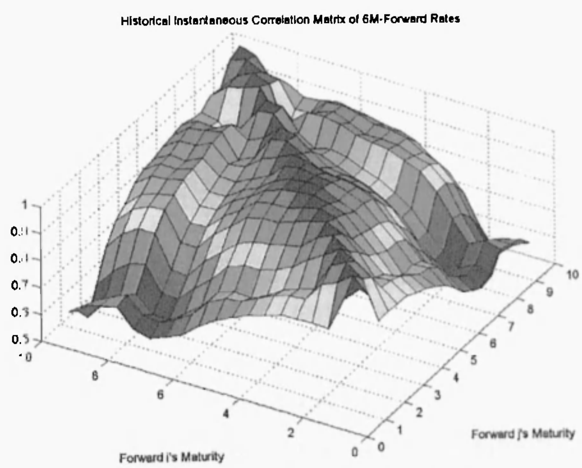
and thus the problem of finding 19x2 arguments $b_{i,j}$ for $i, j = 1, ..., 19$ reduces to finding 19 arguments θ_i for $i = 1, ..., 19$ which can be interpreted as a set of

[46]See Rebonato (2004).

angles. Each set of 19 angles will correspond to a unique correlation matrix and the task is to take a *desired* correlation matrix and try to find a set of angles so that the resulting correlation matrix was very similar to the desired one.

The issue of finding the desired matrix is very complex. The matrix we will try to mimic should capture an expected evolution of correlations among modelled forward rates in the future. It can (and very likely will be) dependent on time, it can be dependent on the actual level/slope/curvature of the yield curve. This shows that to determine the desired form of correlation matrix is a very delicate task. Some studies are arguing, that the desired matrix should be estmated on basis of historical data,[47] other studies show methods of estimating correlation from the swaption prices.[48] The main argument against implied correlation matrix is that the swaption and cap markets may not be integrated that well.[49] Also Rebonato[50] shows that the information that swaption prices imply about correlations might be ambiguous.

Based on the arguments above, the desired correlation matrix was calculated from historical data of swap rates, constant form is considered for the resulting matrix. The time period was chosen as one year and the same method was used to calculate the daily forward rates as when calculating vega above (interpolated daily swap rates were used to calculate zero bond prices based on which the forward rates were computed. The graph below shows the resulting historical correlation matrix which will be taken as the desired correlation matrix with elements denoted $\rho_{i,j}^{des}$.



[47]See for instnace Fan et al. (2003).
[48]See for instance Rebonato (2002).
[49]See Longstaff et al. (2001) or Collin-Dufresne and Goldstein (2002).
[50]See Rebonato (2004).

To construct the model correlation matrix one has to find such paramteth-ers θ_i for $i = 1, ..., 19$, that the difereces between the desired matrix and the model matrix were minimal. Using definitions 90 and 97 and some basic rules for addition of trigonometric functions, the model correlation between forward rates i and j takes the form

$$\begin{aligned} \rho_{i,j}^{mod} &= \cos(\theta_i)\cos(\theta_j) + \sin(\theta_i)\sin(\theta_j) \\ &= \cos(\theta_i - \theta_j) \end{aligned} \tag{99}$$

and to find an optimal set of angels θ_i one has to minimize some objective function such as the following sum of squared residuals:

$$SSR = \sum_{i=1}^{18}\sum_{j=1}^{i}(\rho_{i,j}^{des} - \rho_{i,j}^{mod})^2.$$

The sum above containes 19 parameters that have to be optimized, which makes the task numerically very demanding. Looking at the desired correlation matrix in the above figure one can observe that the correlations with $F_{8.0}$ (6-month forward maturing in 7.5 years from now) are minimal. Thus a grid of thetas pertaining to forward rates $F_{1.0}, F_{3.0}, F_{8.0}, F_{10}$ was created which translates into only four nested-in for-cycles and the remaining forward rates were lineary interpolated. The comlexity of the problem is thus significantly decreased. The following table shows the results for thetas.

Parameters of Correlation Function, 11th November 2005									
θ_1	θ_2	θ_3	θ_4	θ_5	θ_6	θ_7	θ_8	θ_9	θ_{10}
1.963	1.865	1.767	1.669	1.571	1.492	1.414	1.335	1.257	1.178
θ_{11}	θ_{12}	θ_{13}	θ_{14}	θ_{15}	θ_{16}	θ_{17}	θ_{18}	θ_{19}	
1.100	1.021	0.942	0.864	0.785	0.844	0.903	0.962	1.021	
SSR									0.429298

The first graph shows the model correlation matrix and the second graph shows the differences in percentage from the desired correlation matrix.

Comparing the desired and model correlation matrix it can be noticed that the latter is basically smoothing the former one, the parametrical fit is quite reasonable.

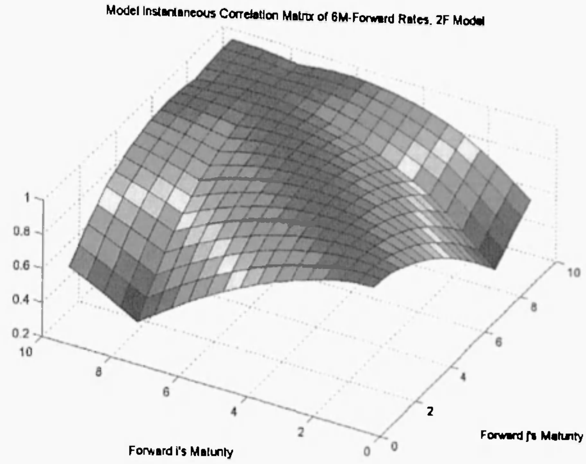
5.3.2 Correlations in 3 Factor Model

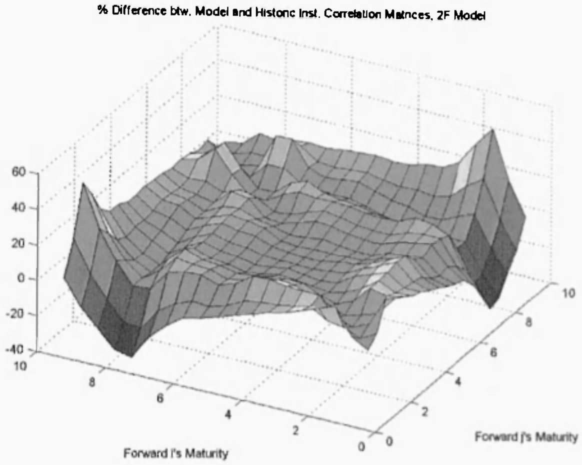
The procedure of determining a correlation matrix in a three factor model is the same as in case of two factors. Condition 96 becomes

$$b_{i,1}^2 + b_{i,2}^2 + b_{i,3}^2 = 1\ , \text{ for } i = 1, ..., 19 \tag{100}$$

which is a definition of coordinates of points laying on a hypersphere in a 3D-plane with radius 1. Setting

$$b_{i,1} = \cos(\theta_{i,1}),\quad b_{i,2} = \cos(\theta_{i,2})\sin(\theta_{i,1}) \quad \text{and} \quad b_{i,3} = \sin(\theta_{i,1})\sin(\theta_{i,2}) \quad \text{for } i = 1, ..., 19 \tag{101}$$





condition 100 simplifies to

$$\sin^2(\theta_{i,1}) + \cos^2(\theta_{i,1}) = 1$$

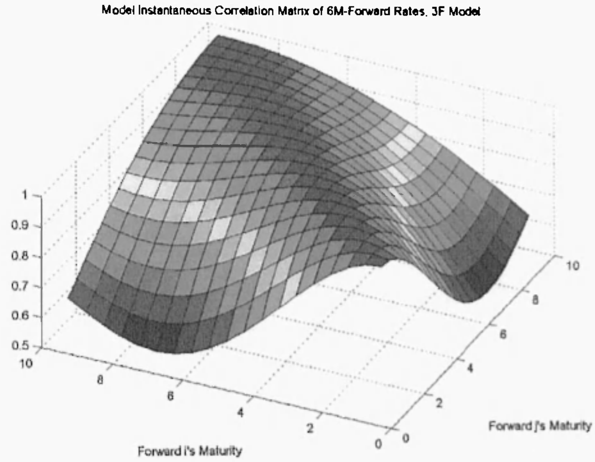
which always holds similarily as 98. Using relationships 90 and and 100, it follows that the model correlation is

$$\rho_{i,j}^{mod} = \cos(\theta_{i,1})\cos(\theta_{j,1}) + \sin(\theta_{i,1})\sin(\theta_{j\,1})\cos(\theta_{i,2} - \theta_{j,2}).$$

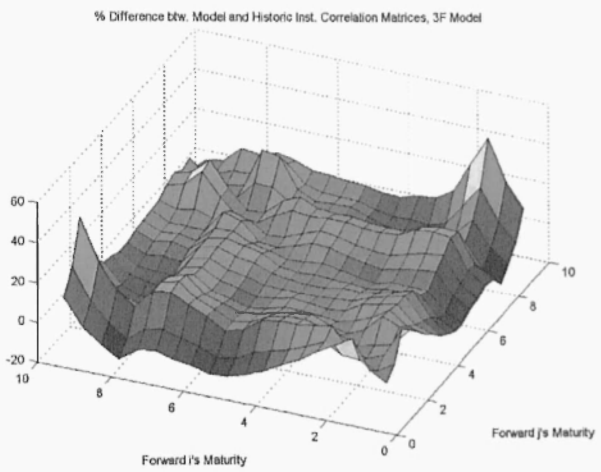
In order to specify the model correlations one has to find a set of 19x2 thetas which minimizes the difference between the desired and resulting correlation matrix. It is obvious that minimizing an objective function with respect to 38 parameters is very complex issue, for computational ease a grid from only four thetas pertaining to forward rates $F_{1.0}$ and $F_{10.0}$ was created and the remaining thetas were lineary interpolated. The same objective function as in the two-factor case was used. The table below shows the results for thetas.

Parameters of Correlation Function, 11th November 2005									
$\theta_{1,1}$	$\theta_{1,2}$	$\theta_{1,3}$	$\theta_{1,4}$	$\theta_{1,5}$	$\theta_{1,6}$	$\theta_{1,7}$	$\theta_{1,8}$	$\theta_{1,9}$	$\theta_{1,10}$
0.864	0.816	0.768	0.720	0.672	0.624	0.576	0.528	0.480	0.432
$\theta_{1,11}$	$\theta_{1,12}$	$\theta_{1,13}$	$\theta_{1,14}$	$\theta_{1,15}$	$\theta_{1,16}$	$\theta_{1,17}$	$\theta_{1,18}$	$\theta_{1,19}$	
0.384	0.336	0.288	0.240	0.192	0.144	0.096	0.048	0.000	
$\theta_{2,1}$	$\theta_{2,2}$	$\theta_{2,3}$	$\theta_{2,4}$	$\theta_{2,5}$	$\theta_{2,6}$	$\theta_{2,7}$	$\theta_{2,8}$	$\theta_{2,9}$	$\theta_{2,10}$
0.000	0.161	0.323	0.484	0.646	0.807	0.969	1.130	1.292	1.453
$\theta_{2,11}$	$\theta_{2,12}$	$\theta_{2,13}$	$\theta_{2,14}$	$\theta_{2,15}$	$\theta_{2,16}$	$\theta_{2,17}$	$\theta_{2,18}$	$\theta_{2,19}$	
1.614	1.776	1.937	2.099	2.260	2.422	2.583	2.745	2.906	
SSR									0.354567

The first graph shows the model correlation matrix and the second graph shows the differences in percentage from the desired correlation matrix.



Because the parametrization of thetas was very rough, the obtained fit is obviously not the most ideal and a better one could be achieved provided a most sophisticated minimization algorithm was available. Yet comparing the squared sum of residuals, the parametrization achieved is better than the one in two factor case. The next section will show the pricing performance of all three models when simulating the swaption prices.


% Difference btw. Model and Historic Inst. Correlation Matrices, 3F Model
60
40
20
0
-20
10
8
6
4
2
0
Forward i's Maturity
0
2
4
6
8
10
Forward j's Maturity

6 Pricing Swaptions with Monte Carlo Simulation

The previous section talked about possible ways of calibrating the LIBOR market model. It was shown, that a reasonable way is to use a combination of market and historical data. Furthermore it was explained that correlations of forward rates have to be employed when the model includes more than one risk factor and how these can be modelled. This section will use the calibration results from the previous section and show how swaptions prices can be calculated using Monte Carlo simulation.

6.1 Market Quotes of Swaptions

Swaptions are not traded on the market like swaps, there is no pool of supply and demand quotes and no mechanism that would pair the matching offers. However there are many brokers who offer indications of swaption prices and these do change as the swap market moves.

The standard quotes that are available as historical time series are those of at-the-money swaptions with expiries of 1, 3, 6 months, 1, 2, 3, 4 and 5 years (the price of ATM receiver swaption is equal to the price of ATM payer swaption). To each expiry 1 to 10 years of underlying swap tenors are available. The quotes are given in implied volatilities (see section 3) and are measured as the average of the last quotations of different brokers in a particular day, middle between bid and ask is taken. The fact that these average quotations are calculated and made published enables us to evaluate the pricing performance of the established models with "market" prices.

The price of a swaption depends on the exact shape of probability distribution of underying swap rate. Thus when pricing a swaption $Swaption(t, t_0, t_n)$ with Libor market model, evolution of all forward rates spanning between time instants t_0 and t_n have to be simulated which are than used to calculate the swap rate at swaption's expiry time. If we want to calculate a price of swaption $Swaption(0, 5.0, 15.0)$ (right to enter into a ten year swap in five years years from now), we have to simulate forward rates $F(0, 5.0, 5.5)$, $F(0, 5.5, 5.0)$, ..., $F(0, 14.5, 15.0)$. This requires further exrapolation of the cap volatility curve to determine the forward volatility function between 10 and 15 years as well as extrapolation of the swap curve to determine the levels of forward rate. This is a very delicate task as both of the quantities are very sensitive to the change of convexity of extrapolated curves and requires very good market experience and especially a market view. Thus only those swaptions will be simulated who's underlying is a swap with last exchange of payments in ten years from now. The following table shows the market volatilities of those swaptions that will be simulated further in this section, quoted on 11^{th} November 2005.

"Market" Implied Swaption Volatilities in %, 11th November 2005									
Expiry	Tenor in Years								
	1	2	3	4	5	6	7	8	9
1	20.7	21.3	21.3	21.1	20.4	19.5	18.9	18.3	17.8
2	21.3	21.2	20.6	19.9	19.3	18.7	18.2	17.8	-
3	20.5	20.3	19.5	18.8	18.3	17.8	17.4	-	-
4	19.5	19.3	18.5	17.8	17.4	17	-	-	-
5	18.5	18.2	17.6	17	16.7	-	-	-	-

Substituting these volatilities into the Black's formula for swaptions (equation 76 in section 3), the market prices can be calculated. The following table shows these prices with respect to a nominal of EUR 1000.

Average Swaption Prices in EUR, Nominal EUR 1000, 11th November 2005									
Expiry	Tenor in Years								
	1	2	3	4	5	6	7	8	9
1	2.00	4.27	6.60	8.94	11.05	12.93	14.85	16.61	18.29
2	3.12	6.37	9.51	12.50	15.43	18.19	20.85	23.40	-
3	3.86	7.82	11.49	15.01	18.50	21.75	24.87	-	-
4	4.43	8.92	13.01	16.88	20.74	24.32	-	-	-
5	4.86	9.68	14.17	18.31	22.45	-	-	-	-

6.2 Discretization of the Forward Rate Dynamics

The whole theoretical framework of forward rate evolution introduced in previous chapters was based on an assumption that the underlying set of forward rates follow a continuous stochastic process. In reality evolution of all financial assets is allways discrete. For instance the Euribor rates are set only once a day in the morning and thus the shortest possible time step in Euribor evolution is one day. On the other hand if one derives forward rates from future contracts on Euribor rate and/or swap contracts (as in our case), these are very liquid and the high frequency of realized trades could resemble continuous processes. Yet even these processes are discrete. The introduced scheme of continuos evolution is used because it is easy to operate with especially when deriving closed formulas for simple contracts as shown in section 3. When simulating the evolution of underying asset according to an available continuous time prescribtion, one has to resort to discretizing the processes.

Instead of simulating the forward rates in the form given by the equation 80 where the state variable is present in the volatility function, the process of logarithm of forward rate was calculated employing Ito's Lemma (see equation 5). The state variable after one time step in a three factor case takes the form

$$\begin{aligned} \ln(F_{t_i}(t+\Delta t)) &= \ln(F_{t_i}(t)) + \sigma_{t_i}(t)\sum_{j-2}^{i}\frac{\rho^{mod}_{t_i,t_j}\tau\sigma_{t_j}(t)F_{t_j}(t)}{1+\tau F_{t_j}(t)} - \frac{\sigma_{t_i}(t)^2}{2}\Delta t + \\ &+ \sigma_{t_i}(t)(b_{i,1}\epsilon_1 + b_{i,2}\epsilon_2 + b_{i,3}\epsilon_3), \end{aligned} \qquad (102)$$

where $\sigma_{t_i}(t)$ takes the parametrical form derived in the previous section, $\rho^{mod}_{t_i,t_j}$

is a relevant model correlation derived in the previous section just like the bs and ϵs are independent normally distributed random variables under a relevant terminal measure Q^{t_i} (see next subsection) with *zero* expected value and variance Δt. The dynamics of two factor model is analogous, dynamics of one factor model as well except of the fact that all ρs are equal *one*.

6.3 Simulations

To calculate prices of swaptions with expiry of one year, 18 forward rates have to be simulated, namely $F_{1.5}, ..., F_{10}$. These forward rates were simulated under the measure $Q^{1.5}$, under which only the first rate is a martingale. To calculate prices of swaptions with expiry of two years, 16 forward rates have to be simulated, namely $F_{2.5}, ..., F_{10}$. These rates were simulated under the measure $Q^{2.5}$, and so on. An alternative would be to simulate all forward rates under the measure $Q^{10.0}$. The only difference would be that the calculated payoff would be discounted by different zero bond price. Both ways would deliver the same results.

The time-step was chosen 1 day and 260 days were taken for one year. The real number of business days in each year is smaller, but this number differs from year to year and thus a fix number of 260 for all years was chosen. This time step is fairly small, in practise usually one week is taken as a minimal time step, so the simulation results should be quite precise.

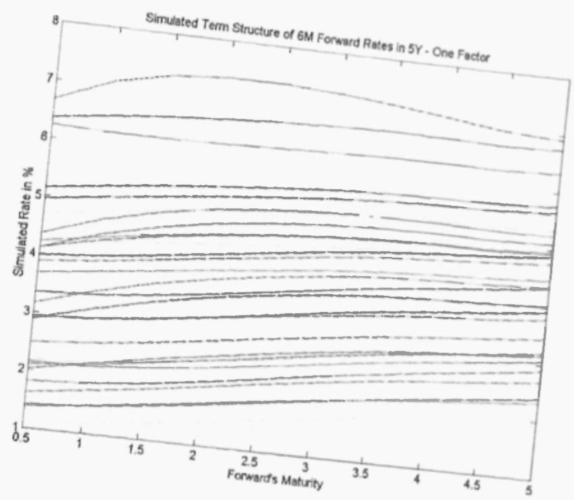
At the end of years 1, ... , 5, the swap rate was calculated from the simulated forward rates (exponential had to be taken first to transpose the simulated logarithm) and a swaption's payoff was evaluated according to a swaption's payoff function in equation 51 and discounted by the relevant zero bond price which was as well calculated from the resulting forward rates. The discounted payoff was recorded and an average was taken at the end. The number of simulations made is 100 000, for each simulation 260 x 5 independent standard-normally distributed random numbers had to be drawn for the case of one factor model, two times and three times as much for two- and three-factor models. In order to prevent serial correlations in the rundom numbers drawn, a powerful random number generator of Park and Miller with Bays-Durham schuffle mechanism was used.[51]

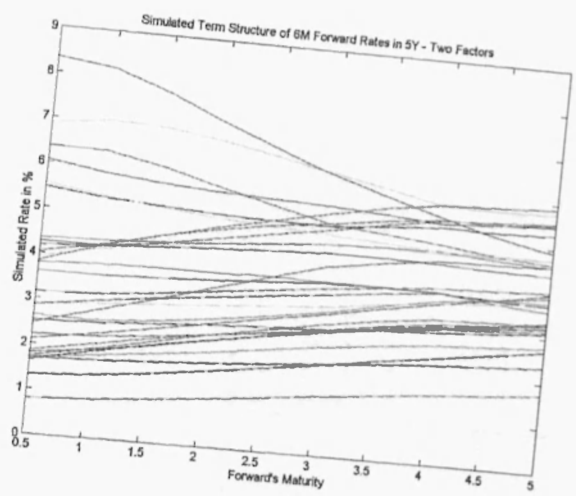
The three graphs below show thirty forward curves that were simulated by the one-, two- and three- factor models. It can be observed, that the one-factor model allows the forward curve to move more or less only horizontaly. Slight changes of shape are caused by the fact, that each forward is subject to different volatility (see the humped-shaped volatility function from the previous section).

The second graph shows a sample of simulated forward curves resulting from two-factor model. Tt can be noticed that this model allows for more significant change of slope and thus models the evolution in more realistic way.

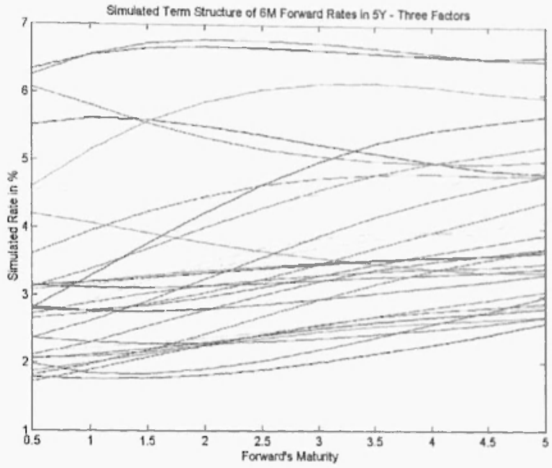
Additional risk factor added provide even more realistic behaviour. A sample of

[51]See Numerical Recipies.


Simulated Term Structure of 6M Forward Rates in 5Y - One Factor
Simulated Rate in %
Forward's Maturity


Simulated Term Structure of 6M Forward Rates in 5Y - Two Factors
Simulated Rate in %
Forward's Maturity

simulations presented in the last graph shows that there is a greater variation of convexity present in the modelled forward curves.



6.4 Results

The graphs above show how the forward curve is modelled by including one to three risk factors. It can be seen, that the two latter models capture the evolution of term structure in more realistic way. One would therefore expect these models to perform better in pricing swaptions. The results presented in the following tables show however exactly the opposite. The first three tables show the Monte-Carlo prices of swaptions using one-, two- and three-factor models in nominal of EUR 1000 on 11^{th} November 2005. All results presented are rounded to second deciaml place.

One Factor Model									
MC Swaption Prices in EUR, Nominal EUR 1000, 11^{th} November 2005									
Expiry	Tenor in Years								
	1	2	3	4	5	6	7	8	9
1	2.08	4.39	6.61	8.64	10.48	12.16	13.71	15.15	16.50
2	3.21	6.49	9.57	12.37	14.90	17.21	19.33	21.30	-
3	4.04	8.04	11.73	15.07	18.08	20.81	23.31	-	-
4	4.66	9.16	13.29	17.00	20.33	23.33	-	-	-
5	5.13	10.03	14.47	18.43	21.96	-	-	-	-

Two Factor Model									
MC Swaption Prices in EUR, Nominal EUR 1000, 11th November 2005									
Expiry	Tenor in Years								
	1	2	3	4	5	6	7	8	9
1	2.09	4.39	6.57	8.52	10.24	11.75	13.07	14.38	15.68
2	3.19	6.43	9.43	12.09	14.42	16.48	18.35	20.19	-
3	4.00	7.93	11.51	14.68	17.46	19.94	22.23	-	-
4	4.62	9.06	13.08	16.61	19.73	22.51	-	-	-
5	5.08	9.89	14.20	17.99	21.32	-	-	-	-

Three Factor Model									
MC Swaption Prices in EUR, Nominal EUR 1000, 11th November 2005									
Expiry	Tenor in Years								
	1	2	3	4	5	6	7	8	9
1	2.09	4.38	6.54	8.47	10.17	11.69	13.09	14.39	15.62
2	3.20	6.45	9.43	12.08	14.43	16.54	18.46	20.25	-
3	4.02	7.95	11.51	14.67	17.47	19.97	22.25	-	-
4	4.62	9.05	13.05	16.59	19.72	22.51	-	-	-
5	5.13	9.98	14.33	18.13	21.47	-	-	-	-

The next three tables show the percentage difference of the simulated Monte-Carlo prices from the market prices on 11th November 2005.

One Factor Model									
% Difference of MC Swaption Prices from Market Prices, 11th November 2005									
Expiry	Tenor in Years								
	1	2	3	4	5	6	7	8	9
1	-3.29	2.83	2.62	0.44	-1.83	-4.49	-6.19	-7.80	-8.78
2	3.27	4.88	3.68	1.44	-0.89	-3.35	-5.20	-7.43	-
3	5.15	4.83	3.67	1.42	-1.18	-3.77	-5.70	-	-
4	3.67	3.30	2.68	0.18	-2.52	-5.17	-	-	-
5	2.97	2.50	1.02	-1.07	-3.86	-	-	-	-

Two Factor Model									
% Difference of MC Swaption Prices from Market Prices, 11th November 2005									
Expiry	Tenor in Years								
	1	2	3	4	5	6	7	8	9
1	-2.82	2.76	1.94	-0.96	-4.09	-7.71	-10.54	-12.49	-13.32
2	2.77	3.88	2.08	-0.86	-4.04	-7.46	-10.00	-12.27	-
3	4.08	3.38	1.70	-1.16	-4.54	-7.79	-10.10	-	-
4	2.86	2.14	1.02	-2.11	-5.41	-8.52	-	-	-
5	1.96	1.11	-0.86	-3.44	-6.69	-	-	-	-

Three Factor Model									
% Difference of MC Swaption Prices from Market Prices, 11th November 2005									
Expiry	Tenor in Years								
	1	2	3	4	5	6	7	8	9
1	-2.84	2.63	1.55	-1.56	-4.73	-8.13	-10.43	-12.43	-13.62
2	3.21	4.18	2.16	-0.91	-4.01	-7.13	-9.47	-12.00	-
3	4.51	3.63	1.73	-1.25	-4.53	-7.67	-9.99	-	-
4	2.83	2.04	0.85	-2.22	-5.45	-8.54	-	-	-
5	2.88	2.05	0.02	-2.68	-6.03	-	-	-	-

From the presented results above the following comments can be made. The one factor model performed the best in general. It managed to recover the market swaption prices within a difference of 5 percent in 80 percent of the simulated swaption contracts and misspriced never by more than 9 percent. It slightly overpriced all swaptions with underlying swaps' tenor from one to four years (except of the first option in the table), while it underpriced the swaptions with underlying swaps' tenor of up to 9 years. Similar holds true for the two- and three-factor models. On the other hand, the two-factor model performed slightly better than the other two models in pricing swaptions with short underlying swaps' tenors, while it underpriced all swaptions with underlying tenors of 7, 8 and 9 years by more than 10 percent.

The three factor model didn't perform significantly worse. Even though it captures the evolution of the yield curve in the most realistic fashion by allowing for a change in slope as well as for the change in convexity, it didn't deliver the best results as would be expected. One reason could be, that the libor-market-model-implied probability distribution might be a different to that considered by the swap market participants. Swaptions are mostly priced by swap market models where the swap rate is modelled directly assuming a closed form probability distribution like a lognormal one. As already mentioned in the third section, these two market models are thus not compatible. In other words if prices of swaptions are computed by brokers say with a swaption market model with lognormal distribution, we will not be able to recover these prices with a libor market model where the implied swap's probability distribution is different.

One could indeed achieve better results than the ones presented by using different calibration method. The prices would have been recovered better if we chose a volatility function of forward rates such, that the model-implied volatility of swap rates was lower for the swaps with short tenors and higher for the swaps with long tenors. We would however be than calibrating to swaption prices and would not recover the market prices of caps.

It is important to note, that calibrating a libor market model to swaption prices can completely make sence even-though we won't recover the prices of caps. The purpose of market models is to price much more complicated instruments whos payoff depends on Euribor rate in our case or on swap rate in case of swap market model. Especially in the past couple of years instruments path-dependent instruments like various ladder swaps, or instruments with callable features like callable structured swaps or bonds became very popular. If we beleive, that the swaptions provide in fact more realistic view about the forward rate volatilities in the future, we can calibrate to swaption prices and use the model to price these complicated structures.

In that perspective, the presented results shouldn't be interpreted right away by concluding that adding more risk factors doesn't bring additional benefits in general. One could certainly calibrate the presented three factor model so that the results delivered were better in pricing swaptions than the one factor model by choosing proper correlation and volatility functions. To choose the right calibration method is a very complex and sensitive issue. A model calibrated in a certain way (say according to cap prices as in our case), which delivers good results in pricing certain type of instruments might perform very poorly pricing different instruments. In our case if we made a statement that one factor model is the best because it is the best in recovering swaption prices, we would probably be very disappointed if we used this model to price instruments who's payoff depends on say a spread between two swap rates, as it doesn't allow the forward curve to change slope much which translates into a smaller volatility between two swap rates. It is important to allways keep on mind what the model is built for, what kind of instruments it aims to price.

7 Conclusion

In the beginning of the thesis a no-arbitrage approach to pricing financial derivatives was introduced. It was shown how the stochastic calculus can be used in expressing dynamics of financial assets and how Ito's formula plays a crucial role in deriving more complicated dynamics. With help of the fundamental pricing formula presented at the end of the first chapter, prices of basic interest rate derivatives were derived later in the thesis, especialy those of caps and swaptions. Both these instruments are very frequent vanilla contracts which are extremely important in modern interest rate financial engineering both as basic building blocks when creating products with more complicated payoff structres, and as hedging instruments. It was described how these instruments are quoted using implied volatilities which stem from the famous Black formula. Even though markes use the Black formula as a convention for quoting prices of caps and swaption, it was explained that it doesn't mean that market participants beleive in a constant volatility of forward rates and their lognormal distribution which the Black formula implies. Still the Black implied constant volatility is a very important source of information when a stochastic model is to be build which is used to price more complicated interest rate derivatives.

In the second part of the thesis Libor market models were built with one, two and three risk factors. It was explained in detail how the models can be calibrated with help of market prices of caps and how monte carlo simulation cas be emloyed to price more complicated payoff. To demonstrate this pricing procedure a set of swaptions was priced with use of all three models and the three sets of prices achieved were compared with market data of swaptions. The best results were delivered by a one factor model which managed to price 80% of swaptions within a difference of 5 percent. The other two modeles didn't perform significantly worse. Furthermore it was argued, that the worse results don't mean that a one factor model is better in general, it was reasoned that there are cases when a more factor model will be more useful than the one factor model.

The topic discussed in this thesis as a very actuall one, interest rate derivatives have become very popular in past couple of years by all parties involved in fixed income business. Here only the most basic procedure of model calibration and of pricing interest rate derivatives was described. The material of this topic is however much more voluminous and keeps attracting more and more attention of both business and academical spheres. I hope this thesis has provided grounds for more deep and thorough analysis.

8 Literature

1. Black, F., Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. Journal of Political Economy 81, 659-683.

2. Brace, A., Gatarek, D., Musiela, M. (1997). The Market Model of Interest Rate Dynamics. Mathematical Finance, Vol 7(2), 127 - 155.

3. Branger, N., Schlag, Ch. (2004). Zinsderivate, Modelle und Bewertung. Springer-Verlag, Berlin.

4. Collin-Dufresne, P. and Goldstein, R.S., (2002). Stochastic Correltaion and the Relative Pricing of Caps and Swaptions in a Generalized-Affine Framework. Working paper, Carnegie Mellon University.

5. Cox, J.C., Ingersoll, J.E., and Ross, S.A (1985) A Theory of the Term Structure of Interest Rates. Econometrica 53, 385 - 407.

6. Deutsch, H.P. (2004). Derivate und Interne Modelle.Schaeffer-Poeschel Verlag Stuttgart.

7. Fan,R., Gupta, A., Ritchken, P. (2002). Hedging in the Possible Presence of Unspanned Stochastic Volatility: Evidence form Swaption Markets. Journal of Finance, 58, no. 5, 2219-2248.

8. Gibson, R., Lhabitant, F. -S., Talay, D. (2001). Modelling the Term Structure of Interest Rates: A Review of the Literature.

9. Harrison, J.M., Pliska, S.R. (1981). Martingales and Stochastic Integrals in the Theory of Cotinuous Trading. Stochastic Processes and their Applications 11, 215-260.

10. Heath, D., Jarrow, R., Morton, A. (1992). Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation. Econometrica 60, 77-105.

11. Hull, J. (1993). Options, Futures and Other Derivative Securities, 2nd Edition. Prentice-Hall Int. Ed.

12. Hull, J., White, A.(1990). Pricing Interest Rate Derivative Securities. The Review of Financial Studies 3, 573-592.

13. Longstaff,F., Santa-Clara, P. and Schwartz, E. (2001). The Relative Valution of Caps and Swaptions: Theory and Empirical Evidence. Journal of Finance, 56 2067-2109.

14. Musiela, M., Rutkowski, M. (2005). Martingale Methods in Financial Modelling. Springer-Verlag, Berlin.

15. Numerical Recipies in C: The Art of Scientific Computing, (1988-1992). Cambridge University Press.

16. Rebonato, R. (2002). Modern Pricing of Interest-Rate Derivatives: The LIBOR Market Mode and Beyond. Princeton: Princeton University Press.

17. Rebonato, R. (2004). Volatility and Correlation. John Wiley & Sons, Chichester.

18. Vasicek, O. (1977). An Equilibrium Characterization of the Term Structure. Journal of Financial Economics 5, 177-188.

9 Appendix A - Ito's Lemma

Let $g : \mathbb{R} \rightarrow \mathbb{R}$ be a real function twice differentiable on $[0,T]$ and $X(t)$ a diffusion on $[0,T]$:

$$dX(t) = \mu(t, X(t))dt + \sigma(t, X(t))dW(t) \quad , \quad t \in [0,T] .$$

The task is to find dynamics of $g(X)$ on $[0,T]$ or $dg(X(t))$ where $t \in [0,T]$. First we can partition the interval $(0,t]$ as

$$(0,T] = \bigcup_{i=1}^{n} (s_{i-1}, s_i] \quad \text{with } s_0 = 0, s_n = T. \tag{103}$$

The process $g(X)$ in t is than equal

$$g(X(t)) = g(X(0)) + \sum_{i=1}^{n} [g(X(s_i)) - g(X(s_{i-i}))] . \tag{104}$$

Because g is twice differentiable, we can express $g(X(s_i))$ using Taylor expansion around $g(X(s_{i-1}))$:

$$\begin{aligned} g(X(s_i)) &= g(X(s_{i-1})) + g'(X(s_{i-1}))\,[X(s_i) - X(s_{i-1})] \\ &\quad + \frac{1}{2} g''(X(\theta))\,[X(s_i) - X(s_{i-1})]^2 , \end{aligned}$$

where $\theta = \lambda X(s_i) + (1 - \lambda)X(s_{i-1})$ with $\lambda \in [0,1]$ is such, that the value $g(X(s_i))$ is exactly met and thus all other terms of Taylor expansion can be omitted.[52] Substituting for $g(s_i)$ in equation 104 we get:

$$\begin{aligned} g(X(t)) = g(X(0)) + \sum_{i=1}^{n} \Big[& g'(X(s_{i-1}))\,[X(s_i) - X(s_{i-1})] + \\ & + \frac{1}{2} g''(X(\theta))\,[X(s_i) - X(s_{i-1})]^2 \Big] . \end{aligned} \tag{105}$$

By incriasing n we make the partitioning 103 finer and it can be shown, that in limit

$$\sum_{i=1}^{n} g'(X(s_{i-1}))\,[X(s_i) - X(s_{i-1})] \overset{n \to \infty}{\longrightarrow} \int_0^t g'(X(s))dX(s) \tag{106}$$

and

$$\sum_{i=1}^{n} g''(X(\theta))\,[X(s_i) - X(s_{i-1})]^2 \overset{n \to \infty}{\longrightarrow} \int_0^t g''(X(s))\sigma(s, X(s))^2 dt. \tag{107}$$

In the second transition the $\sigma(s, X(s))^2 dt$ term is obtained realizing, that

$$\begin{aligned} & [dX(t)]^2 = \\ & \quad = [\mu(t, X(t))dt + \sigma(t, X(t))dW(t)]^2 \\ & \quad = \mu^2 dt^2 + \mu \cdot \sigma \cdot dt \cdot dW(t) + \sigma^2 \left(dW(t)\right)^2 , \end{aligned}$$

[52]The reason why the further terms of Taylor expansion don't appear is, that they are approaching zero in limit transition. It can be easily shown using relatinships in table 108 further in text.

where $(t, X(t))$ was left out for brevity of formula and using the following facts from stochastic analysis: [53]

$$\begin{cases} (dt)^2 = 0 \\ dt \cdot dW(t) = 0 \\ [dW(t)]^2 = dt \end{cases} \tag{108}$$

Substituting 106 and 107 into 105 we obtain

$$g(X(t)) = g(X(0)) + \int_0^t g'(X(t))dX(t) + \frac{1}{2}\int_0^t g''(X(t))\sigma(t, X(t))^2 dt,$$

or in differential form

$$dg(X(t)) = g'(X(t))dX(t) + \frac{1}{2}g''(X(t))vol(dX(t)),$$

where $vol(dX(t))$ is $\sigma(s, X(s))^2 dt$. This is the Ito's Lemma for one diffusion.

Itos Lemma where g is a function of two or more diffusions is analogous, little care must be taken in Taylor's expansion where cross-derivations of second order show up. Ito's Lemma for two diffusions is thus

$$\begin{aligned} dg(X_1, X_2) &= \frac{\partial g}{\partial X_1}dX_1(t) + \frac{\partial g}{\partial X_2}dX_2(t) + \frac{1}{2}\frac{\partial^2 g}{\partial X_1^2}vol(dX_1(t)) + \\ &\quad \frac{1}{2}\frac{\partial^2 g}{\partial X_2^2}vol(dX_2(t)) + \frac{\partial^2 g}{\partial X_1 \partial X_2}\sqrt{vol(dX_1(t))}\sqrt{vol(dX_2(t))}. \end{aligned}$$

[53] See for instance Björk (2004).

10 Appendix B - Feynman-Kac's stochastic representation formula

Let's assume, that we face a boundary-value problem

$$\begin{aligned} F(t,X(t))r &= \frac{\partial F}{\partial t}(t,X(t)) + \frac{\partial F}{\partial X}\mu(t,X(t)) + \frac{1}{2}\frac{\partial^2 F}{\partial X^2}\sigma^2(t,X(t)) & (109)\\ F(T,X(T)) &= \phi(X(T)), & (110)\end{aligned}$$

where $X(t)$ is a process which follows stochastic differential equation

$$dX(t) = \mu(t,X(t))ds + \sigma(t,X(t))dW(t).$$

We can define process Z as

$$Z(t,F) = e^{-rt}F(X). \tag{111}$$

Using Ito's Lemma we can derive dynamics of $Z(t,F)$ as

$$\begin{aligned} dZ(t,F) &= \frac{\partial Z(t,F)}{\partial t}dt + \frac{\partial Z(t,F)}{\partial F}dF\\ &= -r\cdot e^{-rt}F(X)dt + e^{-rt}dF(t).\end{aligned}$$

It is clear that $Z(T)$ is equal its initial value plus increments over $[t,T]$ or

$$Z(T,F(T)) = e^{-rt}F(t,X(t)) - \int_t^T re^{-rs}F(s,X(s))ds + \int_t^T e^{-rs}dF(s,X(s))$$

and multiplying by e^{rT} from substitution 111 we get

$$\begin{aligned} F(T,X(T)) = e^{-r(t-T)}F(t,X(t)) &- e^{rT}\int_t^T re^{-rs}F(s,X(s))ds\\ &+ e^{rT}\int_t^T e^{-rs}dF(s,X(s)). \qquad (112)\end{aligned}$$

Knowing that F is a function of time t and process X, $dF(t,X(t))$ from expression 112 can be derived by Ito's Lemma. After substitution, the last term of equation 112 becomes:[54]

$$\begin{aligned} e^{rT}\int_t^T e^{-rs}dF(s,X(s)) &= e^{rT}\int_t^T e^{-rs}\left[\frac{\partial F}{\partial t} + \frac{\partial F}{\partial X}\mu + \frac{1}{2}\frac{\partial^2 F}{\partial X^2}\sigma^2\right]dt +\\ &+ e^{rT}\int_t^T e^{-rs}\frac{\partial F}{\partial X}\sigma dW(s), \qquad (113)\end{aligned}$$

where $(t,X(t))$ was left out for bravity. We can se, however, that the square bracket in the first integral of 113 is by assumption equal $F(t,X(t))r$ which causes the integral to cross out with the second integral of equation 112. The equation 112 can than be rearranged and we obtain:

$$F(t,X(t)) = e^{-r(T-t)}F(T,X(T)) - e^{rt}\int_t^T e^{-rs}\frac{\partial F}{\partial X}\sigma dW(s). \tag{114}$$

[54] see equation 17 in the thesis.

Now taking expected value of 114, the integral becomes *zero*[55] and substituting the boundary condition 110 for $F(T, X(T))$ we obtain:

$$F(t, X(t)) = e^{-r(T-t)} E[\phi(X(T))]. \tag{115}$$

[55] $\int g(t)dW(t)$ is distributed normally with expected value 0 and variance $\int g^2(t)dt$. See for instance Björk.

11 Appendix C - Caplet's Vega

The Black's formula of a caplet is (see formula 72):

$$Cpl^{Black}(t,T,S,K,\sigma) = P(t,S)\tau(T,S)Bl(K,F(t,T,S),\sigma),$$

where

$$\begin{aligned} Bl(K,F(t,T,S),\sigma) = & \; F(t,T,S)N(d) - K \cdot N(d-s) \\ d = & \; \tfrac{1}{s}\left(\ln\tfrac{F(t,T,S)}{K} + \tfrac{1}{2}s^2\right) \\ s = & \; \sigma\sqrt{T-t}. \end{aligned}$$

Thus for vega we can write:

$$\nu = \frac{\partial Cpl(\sigma)}{\partial\sigma} = P(t,S)\tau(T,S)F(t,T,S)\frac{\partial N(d)}{\partial\sigma} - P(t,S)\tau(T,S)K\frac{\partial N(d-s)}{\partial\sigma}.$$

Leaving out the arguments for brevity and writing the partial derivations as derivations of compounded function we obtain

$$\nu = P\tau F \cdot N'(d)\frac{\partial d}{\partial\sigma} - P\tau K \cdot N'(d-s)\frac{\partial(d-s)}{\partial\sigma}.$$

Because $N(.)$ is a standard normal distribution function, deriving it with respect to its argumet we obtain s standard normal density function.

$$\nu = P\tau F \cdot \frac{1}{\sqrt{2\pi}}\exp\left\{\frac{-d^2}{2}\right\}\frac{\partial d}{\partial\sigma} - P\tau K \cdot \exp\left\{\frac{-(d-s)^2}{2}\right\}\frac{\partial(d-s)}{\partial\sigma}.$$

Substituting for s in the second partial derivation and calculating the partial derivations above we obtain

$$\begin{aligned} \nu &= P\tau F \cdot \frac{1}{\sqrt{2\pi}}\exp\left\{\frac{-d^2}{2}\right\}(-1)\left(\log\frac{F}{K}\frac{1}{\sqrt{T-t}}\sigma^{-2} - \frac{\sqrt{T-t}}{2}\right) + \\ &+ P\tau K \cdot \exp\left\{\frac{-(d-s)^2}{2}\right\}\left(\log\frac{F}{K}\frac{1}{\sqrt{T-t}}\sigma^{-2} + \frac{\sqrt{T-t}}{2}\right). \end{aligned}$$

Substituting for s and putting common arguments outside of brackets we get

$$\begin{aligned} \nu &= P\tau \cdot \frac{1}{\sqrt{2\pi}}\exp\left\{\frac{-d^2}{2}\right\}\left[-F\left(\log\frac{F}{K}\frac{1}{\sqrt{T-t}}\sigma^{-2} - \frac{\sqrt{T-t}}{2}\right) - \right. \\ &- \left. K \cdot \exp\left\{d \cdot \sigma\sqrt{T-t} - \frac{\sigma^2(T-t)}{2}\right\}\left(\log\frac{F}{K}\frac{1}{\sqrt{T-t}}\sigma^{-2} + \frac{\sqrt{T-t}}{2}\right)\right]. \end{aligned}$$

Substituting for d in the exponential of the second member in square brackets the first part simplifies to F further simplification lead to the formula of *vega*

$$\begin{aligned} \nu &= P\tau \cdot \frac{1}{\sqrt{2\pi}}\exp\left\{\frac{-d^2}{2}\right\}\left[F\left(\log\frac{F}{K}\frac{1}{\sqrt{T-t}}\sigma^{-2} - \frac{\sqrt{T-t}}{2}\right) - \right. \\ &- \left. F\left(\log\frac{F}{K}\frac{1}{\sqrt{T-t}}\sigma^{-2} + \frac{\sqrt{T-t}}{2}\right)\right] \\ &= P\tau \cdot \frac{1}{\sqrt{2\pi}}\exp\left\{\frac{-d^2}{2}\right\}F\sqrt{T-t}. \end{aligned}$$